# Table of Contents

**Foreword**  
0

**Part I Introduction**  
8

**Part II MSc in Health Informatics curriculum**  
10

1. Evaluation of existing curriculum  
   - Analysis of documents ................................................................. 10  
   - Student interviews ................................................................. 11
2. Rwandan Health Informatics needs assessment  
   - Competences ................................................................. 13  
3. International standards  
   - IMIA recommendations ................................................................. 16  
   - IMIA knowledge base ................................................................. 18
4. Towards a competence based MSc in Health Informatics for Rwanda  
   - IMIA competences - knowledge base mapping ................................................................. 24  
   - IMIA competences - existing curriculum mapping ................................................................. 26  
   - IMIA knowledge - existing curriculum mapping ................................................................. 26

Development of a reference curriculum  
   - Post-graduate certificate ................................................................. 27

   - Modules  
     - Introduction to Health Informatics ................................................................. 28  
     - Health Care management and organization ................................................................. 30  
     - Knowledge management in healthcare delivery ................................................................. 32  
     - Management Information Systems ................................................................. 34  
     - Electronic Health records management and Hospital & health information Systems ................................................................. 36  
     - Health Informatics applications including; PACS, MIT, Telemedicine and mobile technologies ................................................................. 39

   - Post graduate diploma ................................................................. 42

   - Modules  
     - Softw are are based clinical decision making and support systems ................................................................. 43

---

*2014 Graham Wright, Frank Verbeke, Marc Nyssen*
Part III Regional e-Health Center of Excellence (REHCE)

1 Justification .................................................................................. 68
2 Objectives .................................................................................. 68
   Education ............................................................................... 68
      Programs ............................................................................. 68
      Methods ............................................................................. 68
   Training ................................................................................. 68
   Research ................................................................................ 69
      National & international collaboration ...................................... 69
      Publications ....................................................................... 69
3 Institution ................................................................................ 69
   Infrastructure ......................................................................... 69
      Equipment ........................................................................... 69
      Electronic learning environment .............................................. 69
      Video conferencing & capturing tools ..................................... 69
      On campus teaching instruments ......................................... 70
      Shared virtual health informatics center .................................. 70
   Staffing .................................................................................. 70
      Foreign staff dependency ...................................................... 70
      Local academic capacity building ........................................ 70
4 Finance ..................................................................................... 70
   Financial sources ................................................................... 70
      Government funding ............................................................. 70
      NGO & health partner funding .............................................. 70
      Short courses ...................................................................... 71
      Knowledge center ................................................................ 71
      Business development center (BDC) ..................................... 71
         Objectives ....................................................................... 71
         Activities .......................................................................... 71
         Training ........................................................................... 71
         Testing & evaluation .......................................................... 71
         Audit ................................................................................ 71

2014 Graham Wright, Frank Verbeke, Marc Nyssen
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software development</td>
<td>72</td>
</tr>
<tr>
<td>Solutions design</td>
<td>72</td>
</tr>
<tr>
<td>Consultancy</td>
<td>72</td>
</tr>
<tr>
<td>Staffing</td>
<td>72</td>
</tr>
<tr>
<td>Management</td>
<td>72</td>
</tr>
<tr>
<td>Experts</td>
<td>72</td>
</tr>
<tr>
<td>Technical staff</td>
<td>72</td>
</tr>
<tr>
<td>Researchers</td>
<td>72</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>73</td>
</tr>
<tr>
<td>Classrooms</td>
<td>73</td>
</tr>
<tr>
<td>Offices</td>
<td>73</td>
</tr>
<tr>
<td>Health informatics business incubator</td>
<td>73</td>
</tr>
</tbody>
</table>

**Part IV Referenced literature**

**Index**
Part I
1 Introduction

The University of Rwanda, College of Medicine and Health Sciences (CMHS) ordered a curriculum review of the MSc in Health Informatics (MSc HI) program run at the Regional e-Health Center of Excellence (REHCE). This review activity was performed from February 16th till February 25th 2014 in Kigali by a team of 3 international reviewers:
- Prof. Graham WRIGHT, Walter Sizulu University, Mthatha, South Africa
- Dr. Frank VERBEKE, Vrije Universiteit Brussel, Brussels, Belgium
- Prof. Marc NYSSSEN, Vrije Universiteit Brussel, Brussels, Belgium

The review activities consisted of:
- A study of national regulations and documents relevant for the MSc HI program
- A study of the MSc HI program documents
- A study of international standards, recommendations and literature on health informatics postgraduate education programs which have preferably been endorsed by the International Medical Informatics Association (IMIA)
- A mapping of existing MSc HI curriculum content on international standards and regulations
- A field study of needs and competences in Rwanda in the field of health informatics, mainly through a number of interviews with relevant stakeholders, including Ministry of Health (MoH) staff, NGOs, public and private health facilities, education & training institutions, health software development companies and research organizations.
- An evaluation of how the actual MSc HI curriculum was perceived by students, teaching staff and CMHS faculty.
- Development of a modified program which fills detected competence and content gaps and proved a suitable curriculum to fill those disparities; it also remedies a number of identified program delivery issues.
- Analysis of managerial & financial aspects of the REHCE and the MSc HI curriculum in particular.
- Discussion of the modified program during a stakeholder workshop.
- Integration of stakeholder feedback in a final review report.

The reviewers’ conclusions and recommendations have been presented in a validation workshop in Kigali on February 24th 2014. The feedback provided during this workshop has been integrated in the final report, which has been provided to the CMHS director of postgraduate studies on February 25th 2014.

Done in Kigali on February 25th, 2014

Marc Nyssen    Frank Verbeke    Graham Wright
Part II
2 MSc in Health Informatics curriculum

2.1 Evaluation of existing curriculum

The MSc HI has been running at REHCE since 2011. It was originally hosted at the Kigali Institute of Science and Technology and then moved to the Kigali Health Institute that was recently integrated in the CMHS. At the time of review, 2 cohorts of students had completed the post-graduate certificate and diploma courses and most of them have started preparatory activity for their master dissertation.

2.1.1 Analysis of documents

According to the MSc HI documentation the main objective of the program is to prepare students graduating from the M.Sc HI program to provide leadership in the Health Informatics field. The program has the following high-level aims:

- Develop health informatics skills to the professionals and managers in healthcare sector with hands-on skills required to operate within the ever increasing sophisticated use of health information systems.
- Equip students with generic tools to enable them respond to future developments in healthcare and information technology while grounding them in the "real world" of current initiatives in information policy, technology and management.
- Attract healthcare professionals from the Eastern & Central African Regions, to provide knowledge and skills that will help them maximize the use of information and ICTs in the healthcare sector in the region.
- Prepare professionals to respond to the changing dynamics of healthcare by providing them with the skills necessary to develop, apply, and evaluate the use of information technology in the healthcare arena.
- Produce professionals who can lead organizations in the application of information technology to improve the delivery of healthcare.

In addition, the program has the following educational aims:

- To develop proficiency in analytical skills, procedures and methods relevant to a field of health and clinical services delivery
- To develop skills to evaluate technical skills relevant to clinical or public health practice
- To develop enhanced qualitative and quantitative data analysis capabilities
- To develop awareness of the major ethical issues associated with the area of specialty and appreciation of social determinant cultural sensitivities related to these matters
- Acquisition of proficiency in the concepts, language and fundamentals of the field of study
- To support the development and application of evidence based practice
- To contribute to the development of professional knowledge and practice through education, research scholarly debate, presentation and publication
- To develop skills in research methodology and teaching practice with options to expand experience in each of these domains.

The program combines modules in computer science, health informatics, and health management. The core Health Informatics curriculum includes modules in Health Information Systems, Information Storage, Data retrieval and Management, Research Methods and Outcomes Analysis. Special areas of concentration include Electronic Health-care Records, Health Information Systems, and Health Enterprise Architecture.

Learning outcomes have been defined as follows:

A. Knowledge and Understanding
A1. The capabilities of various data, networking, and enterprise architecture in healthcare system
A2. Managerial issues and technologies related to interoperability in healthcare system
A3. The product strategies of major hardware, software, and telecommunications vendors
A4. National and global architectural standards, regulations, and future developments in healthcare systems
A5. Interaction of the firm with a competitive environment
A6. The role of Enterprise Resource Planning (ERP), Supply Chain Management (SCM), and Customer relationship management (CRM) systems as components of the enterprise architecture
A7. The impact of automation on work practices
A8. Strategic implications of emerging technologies
A9. Tools and techniques of project planning and management
A10. The organizational change process
A11. The strategic use of information technology from a business perspective at the enterprise level

B. Cognitive / Intellectual Skills / Application of Knowledge
B1. Examine issues related to the acquisitions and ongoing management of products, services, and contracts in healthcare systems
B2. Recognize, and improve processes to achieve efficiency and compliance objectives.
B3. Examine the current and potential impact of information and information technology on the aspects healthcare industry.
B4. Examine the characteristics of the information age and explore the implications of emerging ethical concerns such as information privacy, accuracy, property, and accessibility.
B5. Examine the issues which constitute a safe digital environment.

C. Communication / ICT / Numeracy / Analytic Techniques / Practical Skills
C1. Implement and manage security and disaster recovery plans and business continuity from an overall organizational perspective
C2. Develop detailed project plans, schedules, and budgets; estimate project resources; allocate/coordinate resources; and interface with management.
C3. Use project management software.
C4. Identify stakeholders; assess potential impacts of projects; and develop strategies of overcoming resistance, politics, and other human issues.
C5. Practice the processes of HCI development.

D. General transferable skills
D1. Present both technical and managerial ideas in a variety of formats (written, verbal etc).
D2. Demonstrate an effective leadership
D3. Efficiently manage time, human, technical and financial resources

Examination of the stated objectives highlighted a lack of synergy with international recommendations for Health Informatics Specialists. In addition the mapping of the module outcomes as learning outputs shows further gaps between learning outcomes included on the grid and those subsequently included in modules. In particular many aspects of research appear to have been lost in the curriculum documentation although they may have been covered in the Research Module.

2.1.2 Student interviews
Student feedback on the running MSc HI was obtained from a series of 4 individual student interviews and 1 student group discussion. Comments were provided on curriculum content, opportunities to develop practical skills, tutor and student evaluation systems and program staffing.
2.1.2.1 Content

Interviews with students highlighted that a number of modules were well received and students thought that they had a major impact on their progress in Health Informatics. Such perceptions of course may have influence such as teaching style, teacher relationships and difficulty of the subject. When asked “which modules had the biggest impact on you” the responses included:

- E-health Enterprise Architecture
- Introduction to Health Care management and organization
- Software based clinical decision making and support systems
- Applications of Mobile phone technology in Health Informatics
- Introduction to Health Care management and organization

Students thought that these modules “opened my mind”, “helpful for activities”, “professional”, “you could see the a different importance of the application in Rwanda”, “they related theory to practice”, and that “they had examples of application underpinned by theory”

A number of modules were thought to be “just theory”, “not practical” and “not important in the Rwandan situation”. The former included the modules “Introduction to ICT” and “Introduction to management Information Systems” which could be delivered with practical examples and demonstrations. The module though to be “not Important” was “Introduction to GIS Systems in Health Informatics including Epidemic Monitoring and Control”. This module also had comments about the high number of exercises during the teaching activity.

A comment made regarding a number of modules was that “student engagement was missing in much of the delivery” This was particularly highlighted with the video conferencing sessions which were said to be “not interactive”

Students also commented on the apparent lack of sequencing of the modules, which seems to have been an issue of who was available to teach rather than curriculum design.

2.1.2.2 Labs & practice

In terms of increasing the students’ exposure to practice, a number of tactics should be pursued:

- Short visits to Computer and Information experts in the field particularly at hospitals, health facilities, NGO and similar to explore specific applications and see demonstrations.
- Internships for longer periods to be able to experience information processing and analysis in the real world.
- A number of organizations exist which have the potential to provide students an environment in which to conduct their thesis work.

2.1.2.3 Internal evaluation

We have not been able to analyze any evaluation documentation mostly because an External Examiner has not been appointed for the program.

2.1.2.4 Program staffing

The comment that the department of Health Informatics is not a full department and “only has Joseph” shows that some students do not see the visiting staff as members of the department. All commented on the lack of research training particularly on project / proposal planning.

We have not been able to identify skills transfer to Rwandan staff from the International lectures.

An interview with the acting Principle of the College mainly focused on the profile of the MSc graduates and the need for capacity building. She confirmed that the College is committed to the continuation of
the course. Capacity Building was discussed and the following was agreed as a way forward for discussion.

- The successful MSc students could be encourage to have inputs into small sections of a unit within a module, working to develop their teaching skills. They Module leader would mentor them.
- Researchers from the College could be identified who could co-supervise MSc Students, working with a content specialist who would be mentored in research supervision.
- Lecturers from the College who have a health or managerial expertise would be identified who then co-teach parts of a unit alongside the HI specialist and thereby encourage cross transfer of skills.
- Experts from local NGO, Health services, Industry and the Informatics services could be identified and encouraged to provide guest lectures.

2.2 Rwandan Health Informatics needs assessment

Interviews have been organized with a number of major stakeholders, including Ministry of Health (MoH) staff, NGOs, public and private health facilities, education & training institutions, health software development initiatives and research organisations. The main focus of these interviews was to inventory competences needed for the Rwandan health informatics setting and to identify future (short and long term) employment opportunities for MSc HI graduates.

2.2.1 Competences

Information obtained from the MoH clearly showed a growing need for people with synergic skills covering Information Technologies (IT), Information Systems (IS), Information Management (IM) and Health System (HS) knowledge. Generally speaking, there is a shift from IT and IS towards IM and HS needs: internal MoH software and systems development activities are progressively phasing out in favor of private sector outsourcing. Consequently, the MoH requires new management skills relating to contracting, people management and project management.

Health facilities identified competences related to health care research (mainly for university teaching hospitals), data mining and statistical and epidemiological reporting based on Electronic Health Record (EHR) content. Public and private hospitals described the urgent need for Chief Information Officer (CIO) profiles who again must be capable of combining IT, IS and IM knowledge and skills.

Today, only few local software and systems development companies exist today in Rwanda; most of them are subsidiaries of international actors who's R&D activity has not been relocated. Locally required competences are related to system integration, customization and training activities involving predominantly IS and IM skills.

2.2.2 Rwandan Health Informatics activity

Many health informatics projects are running or have been planned in Rwanda, most of them are facing considerable staffing issues. The market for high-level health informatics specialists is real and important, economically justifying the existence of a local or regional MSc HI program.

2.2.2.1 MoH development approaches

It should be noted that the MoH policy recently shifted from internally organizing HI development towards outsourcing of this activity for a number of reasons:

- MoH's current (mainly OpenMRS related) development needs been considerably reduced
- Software development is not being considered core business of the MoH
- IT staff retention is problematic for reasons of public sector wage scales

Also, information systems implementation cannot be completely covered anymore by the MoH. More and more a multi-public sector approach must be developed integrating MoH, Ministry of Public Sector
and Labour, Ministry of Finance and Ministry of Education.

These facts have an important influence on the knowledge and skills that will be needed for the public sector health informatics related activities in the coming years.

2.2.2.2 Interoperability

Interoperability is a main topic in tomorrow's health informatics development in Rwanda. It involves a number of activities that have to be explored and developed in the near future:

- **OpenHIE**: the Open Health Information Exchange platform has been chosen as a core element of Rwanda's information systems integration plans guided by the Rwandan e-Health Enterprise Architecture plans. It mainly involves Client-, Provider- and Health facilities registries, Terminology services and Shared health record services to be shared by a series of health information management applications such as Clinical Information Systems, Laboratory Information Systems, Hospital Information Systems, Community Health Worker Systems or Pharmacy Information Systems.

- **EMR-HMIS interaction**: the extraction of aggregate data from EMR systems to feed HMIS databases in order to reduce parallel and redundant aggregate data reporting.

- **Human resource (HR) management**: implementation of iHRIS for workforce management and integration with professional council databases and HMIS through OpenHIE.

- **Mobile applications**: African healthcare offers a huge market for (smart)phone based applications, especially in rural areas. Many projects (such as RapidSMS) already run in Rwanda and their number is only expected to expand in the future.

- **Performance Based Financing (PBF)**: data validation procedures in PBF programs are time consuming and expensive and could benefit from information systems integration. Different stakeholders explore the opportunities of strengthening the PBF approach using information coming from Hospital Information Systems, HMIS (DHIS2), financial management and HR-systems.

- **Health Insurance**: linking health insurance information systems to care provider systems is a major challenge that can have a huge impact on Rwanda's health insurance implementation (management costs reduction, health care provider profiling, cost evolution projection)

2.2.2.3 Data aggregation

In Rwanda, data aggregation implementation progressively evolves from a multi-database to a one-stop-shopping-point approach, concentrating aggregate data reporting into DHIS2. This approach will in the future solve a number of problems related to:

- inconsistencies between reports from different partners on the same diseases due to different data collection timeframes
- incomplete data due to pieces of information being scattered over different databases
- duplicate reporting

Critical health and performance indicator identification and evaluation leading to the development of context based reporting instruments (e.g. different reporting for different districts) possibly integrating PBF may all be considered activities which will required many skills that can be found in a typical MSc HI program.

2.2.2.4 Human resource management

Health care workforce management includes payroll applications linking MoH activity to the Ministry of Public Services and Labour (MIFOTRA). Bridging skills combining IS, IM and financial knowledge are needed for these kind of programs.

2.2.2.5 Information continuum

The generation, documentation and re-injection of statistical, managerial and epidemiological findings from and into the health system requires high level IM and HS competences which can be provided by a
well designed MSc HI program.

2.2.2.6 Large scale EMR deployment
Rwanda plans for a very ambitious large scale EMR deployment in the next couple of years. These plans will require a massive reinforcement of the available health informatics workforce predominantly with IS and IM knowledge and skills.

2.2.2.7 Health facility financial management
Health sector financial management plans include information systems bridging between the MoH, the Ministry of Finance and health insurance organizations.

2.2.2.8 Health informatics training and education
Considering the huge needs for health informatics training to cope with the ambitious plans of the Rwandan public sector, education and training itself may offer interesting employment opportunities. Specific didactic and research skills will then have to be added to IT, IS, IM and HS proficiency.

2.2.3 Employment opportunities
Based on the above analysis combined with international and more specifically African trends, Job opportunities for MSc HI graduates may exist in many fields. During the stakeholder interviews, quantitative data was obtained for some of them:

Ministry of Health
- HMIS/DHIS implementation
- EMR deployment management
- Information systems evaluation
- Information systems cost/benefit analysis

It should be noted that MoH peripheral data clerk and data manager employment is expected to phase out rapidly, as well as internal IS developer employment.

Health facilities
- District level CIOs (a need for at least 50 jobs identified)
- National reference hospitals CIOs (5), Clinical Information Systems managers (10), Pharmacy Information Systems managers (5) and Laboratory Information Systems managers (5)
- Private health facilities require CIOs (at least 5) and IT/IS managers (5). The latter matches the MSc HI post-graduate certificate level rather than the diploma level.

Telemedicine
- Telemedicine has remained a greenfield application domain in Rwanda. Very little implementations have been deployed although a number of appealing opportunities exist. Setting up such applications will require additional entrepreneurial skills

e-Learning
- Huge needs have been identified for transforming (mainly paper based) traditional content into electronic learning solutions. Not only centers such as REHCE but also many NGOs and public sector organizations may offer jobs (10+) to graduates with combined IS, IM and didactic skills.
- The important anticipated need for MSc level HI graduates goes with the building of local teaching capacity in this domain. Such teaching capacity and expertise may help to considerably reduce the dependency on foreign input and consultancy.

Supply chain management
- This is an emerging health informatics related application field that, due to its economic impact, is
expected to offer many jobs in the health sector.

**Lab systems management**
- The regionally recognized expertise of Rwanda in laboratory practice is expected to be accompanied by a strong growth in the implementation of sophisticated lab information systems.

**Health facility financial management systems**
- Data driven financial management is key for private for-profit organizations, offering attractive jobs to high level IM/Financial specialists that are able to relate cost reduction and increased financial efficiency to effective health care. Monitoring and evaluation must be based on structured information from local health facility sources.
- The progressive shift in the public health sector from program budget- to data driven health policy also promises interesting (longer term) opportunities.

### 2.2.4 Human resources

Based on competence needs and employment perspectives, the following human resource profiles can be identified for MSc HI graduates:

- **Health information managers** (mainly IM and HS knowledge and skills needed in public sector)
- **Clinicians proficient in IT, IS and IM** (health facilities)
- **Information System Implementers** and System Integrators (e.g. for EMR-HMIS integration)
- **Information System Evaluators** (Health Informatics Interventions impact measurement and Information Systems quality evaluation)
- **Education staff** (Academic staff and Information Systems training staff / functional experts)
- **Researchers** (Health Informatics domain and other Health Care domains)

### 2.3 International standards

Studying international health informatics post-graduate education standards and recommendations was part of the review process. The reviewers have mainly studied documents from the International Medical Informatics Association (IMIA) and the American Medical Informatics Association (AMIA). Both organizations being well in line according to health informatics education standards, IMIA documents have been used as a reference.

#### 2.3.1 IMIA recommendations

The "Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics" have been used as a core reference document. An essential part of this document was the list of weighed core competences for Health Informatics specialists:

**Biomedical and Health Informatics Core Knowledge and Skills**

<p>| 1.1 Evolution of informatics as a discipline and as a profession | + |
| 1.2 Need for systematic information processing in health care, benefits and constraints of information technology in health care | ++ |
| 1.3 Efficient and responsible use of information processing tools, to support health care professionals’ practice and their decision making | ++ |
| 1.4 Use of personal application software for documentation, personal communication including Internet access, for publication and basic statistics | ++ |
| 1.5 Information literacy: library classification and systematic health related terminologies and their coding, literature retrieval methods, research methods and research paradigms | ++ |</p>
<table>
<thead>
<tr>
<th>Course Title</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.6 Characteristics, functionalities and examples of information systems in health care</strong> (e.g. clinical information systems, primary care information systems, etc.)</td>
<td>+++</td>
</tr>
<tr>
<td><strong>1.7 Architectures of information systems</strong> in health care; approaches and standards for communication and cooperation and for interfacing and integration of component, architectural paradigms (e.g. service-oriented architectures)</td>
<td>++</td>
</tr>
<tr>
<td><strong>1.8 Management of information systems</strong> in health care (health information management, strategic and tactic information management, IT governance, IT service management, legal and regulatory issues)</td>
<td>+++</td>
</tr>
<tr>
<td><strong>1.9 Characteristics, functionalities and examples of information systems to support patients and the public</strong> (e.g. patient-oriented information system architectures and applications, personal health records, sensor-enhanced information systems)</td>
<td>+</td>
</tr>
<tr>
<td><strong>1.10 Methods and approaches to regional networking and shared care</strong> (eHealth, health telematics applications and inter-organizational information exchange)</td>
<td>++</td>
</tr>
<tr>
<td><strong>1.11 Appropriate documentation and health data management principles</strong> including ability to use health and medical coding systems, construction of health and medical coding systems</td>
<td>+++</td>
</tr>
<tr>
<td><strong>1.12 Structure, design and analysis principles of the health record</strong> including notions of data quality, minimum data sets, architecture and general applications of the electronic patient record/electronic health record</td>
<td>+++</td>
</tr>
<tr>
<td><strong>1.13 Socio-organizational and socio-technical issues</strong>, including workflow/process modelling and reorganization</td>
<td>++</td>
</tr>
<tr>
<td><strong>1.14 Principles of data representation and data analysis</strong> using primary and secondary data sources, principles of data mining, data warehouses, knowledge management</td>
<td>++</td>
</tr>
<tr>
<td><strong>1.15 Biomedical modelling and simulation</strong></td>
<td>+</td>
</tr>
<tr>
<td><strong>1.16 Ethical and security issues</strong> including accountability of health care providers and managers and BMHI specialists and the confidentiality, privacy and security of patient data</td>
<td>++</td>
</tr>
<tr>
<td><strong>1.17 Nomenclatures, vocabularies, terminologies, ontologies and taxonomies in BMHI</strong></td>
<td>++</td>
</tr>
<tr>
<td><strong>1.18 Informatics methods and tools to support education</strong> (incl. flexible and distance learning), use of relevant educational technologies, incl. Internet and World Wide Web</td>
<td>+</td>
</tr>
<tr>
<td><strong>1.19 Evaluation and assessment</strong> of information systems, including study design, selection and triangulation of (quantitative and qualitative) methods, outcome and impact evaluation, economic evaluation, unintended consequences, systematic reviews and meta-analysis, evidence-based health informatics</td>
<td>+++</td>
</tr>
</tbody>
</table>

**Medicine, Health and Biosciences, Health System Organisation**

<table>
<thead>
<tr>
<th>Course Title</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2.1 Fundamentals of human functioning</strong> and biosciences (anatomy, physiology, microbiology, genomics, and clinical disciplines such as internal medicine, surgery etc.)</td>
<td>+</td>
</tr>
<tr>
<td><strong>2.2 Fundamentals of what constitutes health</strong>, from physiological, sociological, psychological, nutritional, emotional, environmental, cultural, spiritual perspectives and its assessment</td>
<td>+</td>
</tr>
<tr>
<td><strong>2.3 Principles of medical decision making</strong> and diagnostic and therapeutic strategies</td>
<td>++</td>
</tr>
<tr>
<td><strong>2.4 Organisation of health institutions</strong> and of the overall health system, interorganizational aspects, shared care</td>
<td>+++</td>
</tr>
<tr>
<td><strong>2.5 Policy and regulatory frameworks</strong> for information handling in health care</td>
<td>+</td>
</tr>
<tr>
<td><strong>2.6 Principles of evidence-based medicine</strong> and evidence-based nursing</td>
<td>+</td>
</tr>
<tr>
<td><strong>2.7 Health administration, health economics</strong>, health quality management and resource management, patient safety initiatives, public health services and outcome measurement</td>
<td>++</td>
</tr>
</tbody>
</table>
### Informatics/Computer Science, Mathematics, Biometry

<table>
<thead>
<tr>
<th>Competence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>3.1 Basic informatics terminology</strong></td>
<td>like data, information, knowledge, hardware, software, computer, networks, information systems, information systems management</td>
</tr>
<tr>
<td><strong>3.2 Ability to use personal computers</strong></td>
<td>text processing and spread sheet software, easy-to-use database management systems</td>
</tr>
<tr>
<td><strong>3.3 Ability to communicate electronically</strong></td>
<td>including electronic data exchange, with other health care professionals, internet/intranet use</td>
</tr>
<tr>
<td><strong>3.4 Methods of practical informatics/computer science</strong></td>
<td>especially on programming languages, software engineering, data structures, database management systems, information and system modelling tools, information systems theory and practice, knowledge engineering, (concept) representation and acquisition, software architectures</td>
</tr>
<tr>
<td><strong>3.5 Methods of theoretical informatics/computer science</strong></td>
<td>e.g. complexity theory, encryption/security</td>
</tr>
<tr>
<td><strong>3.6 Methods of technical informatics/computer science</strong></td>
<td>e.g. network architectures and topologies, telecommunications, wireless technology, virtual reality, multimedia</td>
</tr>
<tr>
<td><strong>3.7 Methods of interfacing and integration</strong></td>
<td>of information system components in health care, interfacing standards, dealing with multiple patient identifiers</td>
</tr>
<tr>
<td><strong>3.8 Handling of the information system life cycle</strong></td>
<td>analysis, requirement specification, implementation and/or selection of information systems, risk management, user training</td>
</tr>
<tr>
<td><strong>3.9 Methods of project management and change management</strong></td>
<td>i.e. project planning, resource management, team management, conflict management, collaboration and motivation, change theories, change strategies</td>
</tr>
<tr>
<td><strong>3.10 Mathematics</strong></td>
<td>algebra, analysis, logic, numerical mathematics, probability theory and statistics, cryptography</td>
</tr>
<tr>
<td><strong>3.11 Biometry and epidemiology</strong></td>
<td>including study design</td>
</tr>
<tr>
<td><strong>3.12 Methods for decision support</strong></td>
<td>and their application to patient management, acquisition, representation and engineering of medical knowledge; construction and use of clinical pathways and guidelines</td>
</tr>
<tr>
<td><strong>3.13 Basic concepts and applications of ubiquitous computing</strong></td>
<td>e.g. pervasive, sensor-based and ambient technologies in health care, health enabling technologies, ubiquitous health systems and ambient assisted-living</td>
</tr>
<tr>
<td><strong>3.14 Usability engineering, human-computer interaction</strong></td>
<td>usability evaluation, cognitive aspects of information processing</td>
</tr>
</tbody>
</table>

This competences list served as one of the benchmarks for evaluating the Rwandan MSc HI curriculum. All of the competences identified by the Rwandan stakeholders were covered by the IMIA recommendation.

### 2.3.2 IMIA knowledge base

In 2011, IMIA also published a knowledge base structuring the knowledge of the Health Informatics domain into 14 topics which were subdivided into 245 knowledge units. This unit list served in the review process as a tool for measuring the coverage of Health Informatics knowledge by the MSc HI curriculum.

**Computer Science for Health Informatics (ICT for Health)**
- Biomedical equipment interfaces
- Collaborative Internet architectures
- Computer literacy (ECDL)
• Computer systems
• Computer-aided design
• Computing methodologies
• Data structures
• Decision Support Systems performance evaluation
• Decision support tools
• Demystify IT for users
• Design databases
• Design of IS and IT systems for Health
• Differentiate between technical, syntactic and semantic interoperability
• EPR: architecture, content, views and use
• Futurology
• Human Computer Interaction (HCI) principles
• Information and systems requirements
• Information retrieval
• Knowledge management
• Messaging standards
• Mobile communications
• Natural language processing
• Networking
• Open systems
• Picture archiving systems (PACs)
• Prototype system for a department
• System engineering
• Systems design and the application of systems theory to design

Health & Social Care processes
• Advantages and disadvantages of existing systems for primary clinical uses
• Application of telehealth solutions to the elderly population
• Appointments and Admissions management
• Cancer registration
• Clinical consultation
• Clinical guidelines
• Clinical information needs
• Clinical pathways and disease management
• Consumer health and patient information
• Decision making processes in clinical decision making
• Diagnostic process
• Effective implementation strategy for a clinical system
• Healthcare process
• Home monitoring
• How delivery of care differs in primary and secondary care
• How Health Informatics affects outcomes
• Issues of consumerism in Health Informatics
• Patient information needs and range of delivery methods
• Patient journey
• Patient outcomes and how to measure them
• Specialty-specific clinical information systems
• Stages of communication from source to recipient and what influences
• Systems currently used by clinicians to gather clinical information
• The interface between the technology and caring philosophies
• The method of primary to secondary to tertiary care referral and the flows of information
• Treatment process
**Health (care) records**
- Audit trails
- Case note tracking
- Clean clinical data
- Clinical communication from message to meaning
- Community information systems
- Criteria for paperlite systems
- Data structures
- Decision support
- Define EHR
- Document management
- Electronic healthcare records
- History of methods of gathering information in the clinical workplace
- Hospital information systems
- Information Storage and Retrieval
- National care-record systems
- Patient : clinician consultation
- Patient Access to records
- Patient and provider identification
- Patient referral
- The importance of maintaining the integrity of the Master Patient Index and Patient Numbering Systems
- Usefulness of clinical record keeping standards

**Health and Social care Industry**
- Assess Health Informatics readiness of an organization
- Financial management
- Healthcare models: economic and financial
- Healthcare supply chain
- Integrating healthcare and social care
- Interoperability
- Procurement
- Provide Business requirements for system
- Understand the health care organization

**Health informatics standards**
- Clinical coding systems
- Clinical interface protocol
- Interoperability standards
- Messaging standards
- Ontologies and data definitions
- Quality management
- Quality of message design process standard
- Specific clinical message
- Standards for coding terminology and communication
- Technical infrastructures
- Technical standards - Design and Implementation
- Unique architectures and standards development
- Unique Identifiers Design and development

**Knowledge Domains and Knowledge Discovery**
- Access to information
• Benefits realization
• Clinical content production
• Clinical Data structures
• Clinical decision support
• Data mining
• Evaluate integrity of decision support system
• Evaluate the impact of a health informatics intervention
• Information sources
• Knowledge of performance and process improvement
• Measuring (evaluating) performance of organizations
• Produce knowledge domain constraint models
• Research methods of information science and healthcare
• Statistics and numerical data

Legal and Ethical
• Current relevant legislation
• Data Protection Acts and other privacy legislation
• Data security, privacy, confidentiality, access, integrity and standards
• Ethics and governance
• Health Informatics ethics
• How Health Informatics benefits patients
• Implement safe and secure IS and IT systems
• Justification of e-Health
• Legislation
• Patient consent
• Patient safety
• Policy regulatory requirements around medical record documentation
• Principles of Freedom of Information and other legislation
• Principles of security in a health context
• The impact of breaches in confidentiality on patient care

People in organizations
• Capacity management
• Change management
• Comprehends psychology and sociology
• Contract management
• Contrast management Theories
• Educational resources for healthcare professionals
• Explains health informatics
• Healthcare professional roles
• Implement new information systems
• Manage projects and services
• Organizational and cultural change
• Organizational issues
• Project and Contract management tools
• Project governance
• Project planning
• The needs of patients, clients and professionals
• Use communication, teaching and learning strategy to present information to patients and clients

Politics and policy
• Analyse differences in strategy when comparing e-Health and IT programmes
• Describe likely trends in Health Policy and IT
Differentiates professional organizations
Evaluation
Information services for the public
Information Systems Strategies - Alignment with External forces
Local healthcare strategy
National IT policies, strategies and programs
Strategy and policy implementation
Understand the economic forces driving the healthcare industry

Technologies for health
Ambulance systems
Assistive technology
Automatic data processing
Care at home
Data acquisition
Data communications and messaging
Databases
e-Healthcare
Electronic booking
Electronic prescribing and medicine management
Home monitoring
Infrastructure
Internet, intranets and associative technologies
Laboratory information systems
Mobile technology
Picture archiving and communications systems
Point-of-care testing
Radio frequency identification
Robotics
Smart cards
Software engineering and development
Telecare and telemedicine
Telemonitoring
Voice interface systems
Web technologies
Wireless technology

Terminology, classification and grouping
Clinical coding, terminology and systems
Data management and storage
Data quality
Disease registers
Evaluate competence of a terminology for a clinical task
Evaluate different terminology systems that describe health
Limitations of coding systems
Natural language processing
Ontologies - Data definitions
Standards for coding, terminology and communication
The significance of accurate data
Types of coding, terming, classification
Unified medical language system

Uses of clinical information
• Analyse (clinical) data
• Clinical audit
• Data analysis and statistical presentation
• Data for population management
• Data mining
• Data quality
• Data structures
• Data warehousing
• Differentiate relative and actual risk
• Epidemiology
• Examples of information systems used to gather clinical data for secondary care purposes
• Geographical Information systems
• Identify data sources to support clinical risk management
• Improvements in data quality
• Interpret clinical data
• Resource management
• Techniques to maximize data quality
• The clinical audit cycle
• The significance of accurate data
• The uses of clinical information in secondary and primary care

Using informatics to support clinical healthcare governance
• Adverse Event Management
• Clinical audit and effectiveness
• Clinical governance
• Information governance
• Organizational Learning
• Patient Information i.e. Health Education
• Patient safety
• Performance indicators and management
• Risk Management

Computer Systems applications in Health (Toolkit)
• Ability to undertake audit
• Appraise options for IS and IT solutions
• Build a model of a hospital department
• Compare the effects of alternative system designs
• Construct systems
• Data modeling to support analysis
• Develop and implement information and ICT strategy
• Evaluate health system information flow and system analysis
• Evaluate healthcare
• Evaluate IS and IT
• Evaluate proposed systems
• Evaluation of implementations
• Formulating questions
• How to do Computational modeling
• Manage information systems - the management of IS
• Modeling of processes
• Political influencing
• Process Analysis and redesign
• Process optimization
• Reconstruction of Clinical Guidelines in a form for a Clinical Decision Support system
2.4 Towards a competence based MSc in Health Informatics for Rwanda

Based on the results obtained from the Rwandan Health Informatics needs assessment, the evaluation of the existing curriculum and against the background of international health informatics education standards, the reviewers started the process of developing a reference curriculum for Rwanda. This reference curriculum was then compared to the existing one resulting in a competence- and content gap analysis, which lead to a series of recommendations for reaching the target curriculum.

2.4.1 IMIA competences - knowledge base mapping

In a first exercise, 40 competences identified in the "IMIA recommendations for health informatics education" have been mapped on 0 or more knowledge units in the IMIA knowledge base. This resulted in a total of 400 mappings, clearly demonstrating the knowledge overlap for many of the competences. Only for 1 competence no corresponding knowledge unit could be found: "2.1 Fundamentals of human functioning and biosciences (anatomy, physiology, microbiology, genomics, and clinical disciplines such as internal medicine, surgery etc.)".

<table>
<thead>
<tr>
<th>Competence</th>
<th>Knowledge units</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Evolution of informatics as a discipline and as a profession</td>
<td>4</td>
</tr>
<tr>
<td>1.10 Methods and approaches to regional networking and shared care (eHealth, health telematics applications and inter-organizational information exchange)</td>
<td>8</td>
</tr>
<tr>
<td>1.11 Appropriate documentation and health data management principles including ability to use health and medical coding systems</td>
<td>11</td>
</tr>
<tr>
<td>1.12 Structure, design and analysis principles of the health record including notions of data quality, minimum data sets, architecture and general applications of the electronic patient record/electronic health record</td>
<td>9</td>
</tr>
<tr>
<td>1.13 Socio-organizational and socio-technical issues, including workflow/process modelling and reorganization</td>
<td>9</td>
</tr>
<tr>
<td>1.14 Principles of data representation and data analysis using primary and secondary data sources, principles of data mining, data warehouses, knowledge management</td>
<td>9</td>
</tr>
<tr>
<td>1.15 Biomedical modelling and simulation</td>
<td>5</td>
</tr>
<tr>
<td>1.16 Ethical and security issues including accountability of health care providers and managers and BMHI specialists and the confidentiality, privacy and security of patient data</td>
<td>15</td>
</tr>
<tr>
<td>1.17 Nomenclatures, vocabularies, terminologies, ontologies and taxonomies in BMHI</td>
<td>11</td>
</tr>
<tr>
<td>1.18 Informatics methods and tools to support education (incl. flexible and distance learning), use of relevant educational technologies, incl. Internet and World Wide Web</td>
<td>5</td>
</tr>
<tr>
<td>1.19 Evaluation and assessment of information systems, including study design, selection and triangulation of (quantitative and qualitative) methods, outcome and impact evaluation, economic evaluation, unintended consequences, systematic reviews and meta-</td>
<td>22</td>
</tr>
<tr>
<td>1.2 Need for systematic information processing in health care, benefits and constraints of information technology in health care</td>
<td>14</td>
</tr>
<tr>
<td>1.3 Efficient and responsible use of information processing tools, to support health care</td>
<td>13</td>
</tr>
<tr>
<td>1.4 Use of personal application software for documentation, personal communication including Internet access, for publication and basic statistics</td>
<td>6</td>
</tr>
<tr>
<td>1.5 Information literacy: library classification and systematic health related terminologies and their coding, literature retrieval methods, research methods and research paradigms</td>
<td>11</td>
</tr>
<tr>
<td>1.6 Characteristics, functionalities and examples of information systems in health care (e.g. clinical information systems, primary care information systems, etc.)</td>
<td>18</td>
</tr>
<tr>
<td>1.7 Architectures of information systems in health care; approaches and standards for communication and cooperation and for interfacing and integration of component, architectural paradigms (e.g. service-oriented architectures)</td>
<td>11</td>
</tr>
<tr>
<td>1.8 Management of information systems in health care (health information management, strategic and tactic information management, IT governance, IT service management, legal and regulatory issues)</td>
<td>11</td>
</tr>
<tr>
<td>1.9 Characteristics, functionalities and examples of information systems to support patients and the public (e.g. patient-oriented information system architectures and applications, personal health records, sensor-enhanced information systems)</td>
<td>15</td>
</tr>
<tr>
<td>2.1 Fundamentals of human functioning and biosciences (anatomy, physiology, microbiology, genomics, and clinical disciplines such as internal medicine, surgery etc.)</td>
<td>0</td>
</tr>
<tr>
<td>2.2 Fundamentals of what constitutes health, from physiological, sociological, psychological, nutritional, emotional, environmental, cultural, spiritual perspectives and its assessment</td>
<td>4</td>
</tr>
<tr>
<td>2.3 Principles of medical decision making and diagnostic and therapeutic strategies</td>
<td>11</td>
</tr>
<tr>
<td>2.4 Organisation of health institutions and of the overall health system, interorganizational aspects, shared care</td>
<td>12</td>
</tr>
<tr>
<td>2.5 Policy and regulatory frameworks for information handling in health care</td>
<td>10</td>
</tr>
<tr>
<td>2.6 Principles of evidence-based medicine and evidence-based nursing</td>
<td>4</td>
</tr>
<tr>
<td>2.7 Health administration, health economics, health quality management and resource management, patient safety initiatives, public health services and outcome measurement</td>
<td>16</td>
</tr>
<tr>
<td>3.1 Basic informatics terminology like data, information, knowledge, hardware, software, computer, networks, information systems, information systems management</td>
<td>3</td>
</tr>
<tr>
<td>3.10 Mathematics: algebra, analysis, logic, numerical mathematics, probability theory and statistics, cryptography</td>
<td>9</td>
</tr>
<tr>
<td>3.11 Biometry and epidemiology, including study design</td>
<td>14</td>
</tr>
<tr>
<td>3.12 Methods for decision support and their application to patient management, acquisition, representation and engineering of medical knowledge; construction and use of clinical pathways and guidelines</td>
<td>9</td>
</tr>
<tr>
<td>3.13 Basic concepts and applications of ubiquitous computing (e.g. pervasive, sensor-based and ambient technologies in health care, health enabling technologies, ubiquitous health systems and ambient assisted-living)</td>
<td>9</td>
</tr>
<tr>
<td>3.14 Usability engineering, human-computer interaction, usability evaluation, cognitive aspects of information processing</td>
<td>7</td>
</tr>
<tr>
<td>3.2 Ability to use personal computers, text processing and spread sheet software, easy-to-use database management systems</td>
<td>4</td>
</tr>
<tr>
<td>3.3 Ability to communicate electronically, including electronic data exchange, with other health care professionals, internet/intranet use</td>
<td>8</td>
</tr>
<tr>
<td>3.4 Methods of practical informatics/computer science, especially on programming languages, software engineering, data structures, database management systems, information and system modelling tools, information systems theory and practice, knowledge engi</td>
<td>7</td>
</tr>
<tr>
<td>3.5 Methods of theoretical informatics/computer science, e.g. complexity theory,</td>
<td>3</td>
</tr>
</tbody>
</table>
3.6 Methods of technical informatics/computer science, e.g. network architectures and topologies, telecommunications, wireless technology, virtual reality, multimedia

3.7 Methods of interfacing and integration of information system components in health care, interfacing standards, dealing with multiple patient identifiers

3.8 Handling of the information system life cycle: analysis, requirement specification, implementation and/or selection of information systems, risk management, user training

3.9 Methods of project management and change management (i.e. project planning, resource management, team management, conflict management, collaboration and motivation, change theories, change strategies)

### 2.4.2 IMIA competences - existing curriculum mapping

Each module of the existing MSc HI curriculum has been mapped on a number of IMIA competences (replacing the original program objectives). This showed that a total set of 38 IMIA competences (out of 40) were covered by the existing curriculum, which was a remarkably good score. Only the following competences were missing:

- 2.1 Fundamentals of human functioning and biosciences (anatomy, physiology, microbiology, genomics, and clinical disciplines such as internal medicine, surgery etc.)
- 2.2 Fundamentals of what constitutes health, from physiological, sociological, psychological, nutritional, emotional, environmental, cultural, spiritual perspectives and its assessment

It was concluded that in terms of competences, the actual MSc HI curriculum responded to the Rwandan stakeholder expectations.

### 2.4.3 IMIA knowledge - existing curriculum mapping

Based on the mapping between existing MSc HI modules and IMIA competences combined with the previously performed mapping between IMIA competences and IMIA knowledge units, an automatic mapping between existing MSc HI modules and IMIA knowledge units was generated in order to evaluate the coverage of the Health Informatics knowledge domain by the existing program. This demonstrated that 26 (out of 245) knowledge units were not covered by the existing program.

### 2.4.4 Development of a reference curriculum

The automatically generated mapping table between existing curriculum modules and IMIA knowledge units was then cleaned for redundancy (assigning each knowledge unit only to 1 single module) and for completeness (modifying module definitions in order to cover previously missed knowledge units). This provided a set of modules with 100% knowledge coverage whilst avoiding any redundant teaching.

In a final phase, a number of modules have been (partially) merged in order to come to an equally distributed module weight of 10 credits for each module, with 6 modules in the post-graduate certificate year and 6 modules in the post-graduate diploma year.

All modules of the developed curriculum have been completed with references to recent literature. Some of the referenced papers may be chosen by the module manager to be mandatory literature during lectures, assignments or examinations and other papers may remain optional yet useful for developing critical scientific reading skills. It is encouraged that such reading skills be peer-reviewed within each module.
2.4.4.1 Post-graduate certificate

The post-graduate certificate program essentially focuses on IT and IS after an introduction to the Health Informatics domain. The development of practical skills and essential knowledge of the Rwandan health informatics context are core components of the program.

2.4.4.1.1 Modules
2.4.4.1.1 Introduction to Health Informatics

1. **Code** EHI 611

2. **Title** Introduction to Health Informatics

3. **Year** Certificate.

4. **Credits** 10 Credits

5. **Pre-requisite modules** None

6. **Allocation of study & teaching hours**
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   This module is an introduction to the discipline of Health Informatics and the scope of the program. It covers a number of basic concepts in Health Informatics and ensures that the students from different background have the opportunity to share perceptions of Health and Medical Informatics.

8. **Learning outcomes (competences)**
   1.1 Evolution of informatics as a discipline and as a profession
   1.4 Use of personal application software for documentation, personal communication including Internet access, for publication and basic statistics
   3.1 Basic informatics terminology like data, information, knowledge, hardware, software, computer, networks, information systems, information systems management
   3.2 Ability to use personal computers, text processing and spread sheet software, easy-to-use database management systems
   3.3 Ability to communicate electronically, including electronic data exchange, with other health care professionals, internet/intranet use
   3.4 Methods of practical informatics/computer science, especially on programming languages, software engineering, data structures, database management systems, information and system modelling tools, information systems theory and practice, knowledge engineering, (concept) representation and acquisition, software architectures
   3.6 Methods of technical informatics/computer science, e.g. network architectures and topologies, telecommunications, wireless technology, virtual reality, multimedia

9. **Indicative content (units)**
   Access to information
   Computer literacy (ECDL)
   Computer systems
   Computing methodologies
   Data management and storage
   Databases
   Demystify IT for users
   Explains health informatics
   History of methods of gathering information in the clinical workplace
   Human Computer Interaction (HCI) principles
   Information sources
   Information Storage and Retrieval
   Internet, intranets and associative technologies
Networking
Web technologies
Wireless technology
Prototype system for a department

10. **Learning & teaching strategy**
Through an integrated program of lectures, seminars, workshops and tutorials, students engage with course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies.

11. **Assessment strategies**
Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.
Collaborative work: oral/visual presentations and written reports

12. **Strategy for feedback and student support during module**
A personal tutor should be identified for each student during this initial module. This tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.
Student should complete a formal feedback to the staff at the end of the module.

13. **Indicative resources**
MBANANGA, N. 2006. Introduction to health informatics. *Pretoria: Notoro Publish*
2.4.4.1.2 Health Care management and organization

1. **Code**  EHI612

2. **Title**  Health Care management and organization

3. **Year**  Certificate

4. **Credit**  10 Credits

5. **Pre-requisite modules**
   EHI611 Introduction to Health Informatics

6. **Allocation of study & teaching hours**
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   This module covers the concepts of Healthcare Management in relation to Health Informatics. The health Informatics concepts underpinning management, organisational culture and socio-technical aspects are explored.

8. **Learning outcomes (competences)**
   1.5 Information literacy: library classification and systematic health related terminologies and their coding, literature retrieval methods, research methods and research paradigms
   1.11 Appropriate documentation and health data management principles including ability to use health and medical coding systems, construction of health and medical coding systems
   1.13 Socio-organizational and socio-technical issues, including workflow/process modelling and reorganization
   1.15 Biomedical modelling and simulation
   2.4 Organisation of health institutions and of the overall health system, interorganizational aspects, shared care
   2.7 Health administration, health economics, health quality management and resource management, patient safety initiatives, public health services and outcome measurement

9. **Indicative content (units)**
   Health Care management and organization
   Assess Health Informatics readiness of an organisation
   Contract management Theories
   Differentiates professional organizations
   Financial management
   Healthcare models: economic and financial
   Healthcare professional roles
   Integrating healthcare and social care
   Modelling of processes
   Resource management
   Standards for coding terminology and communication
   Understand the health care organisation

10. **Learning & teaching strategy**
    Through an integrated program of lectures, seminars, workshops and tutorials, students engage with
course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies.

11. **Assessment strategies**
   Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.
   Collaborative work: oral/visual presentations and written reports.

12. **Strategy for feedback and student support during module**
   The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.
   Student should complete a formal feedback to the staff at the end of the module.

13. **Indicative resources**
2.4.4.1.3 Knowledge management in healthcare delivery

1. **Code** EHI613

2. **Title** Knowledge management in healthcare delivery

3. **Year** Certificate

4. **Credits** 10 Credits

5. **Pre-requisite modules**
   - EHI611 Introduction to Health Informatics
   - EHI612 Health Care management and organization

6. **Allocation of study & teaching hours**
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   This module covers the Health Informatics topics of the use of knowledge within a number of areas. In particular the processing of Data into Information, and then into Knowledge. The use of such knowledge is explored in Decision Making and Education.

8. **Learning outcomes (competences)**
   1.3 Efficient and responsible use of information processing tools, to support health care professionals’ practice and their decision making
   1.5 Information literacy: library classification and systematic health related terminologies and their coding, literature retrieval methods, research methods and research paradigms
   1.11 Appropriate documentation and health data management principles including ability to use health and medical coding systems, construction of health and medical coding systems
   1.13 Socio-organizational and socio-technical issues, including workflow/process modelling and reorganization
   1.14 Principles of data representation and data analysis using primary and secondary data sources, principles of data mining, data warehouses, knowledge management
   1.17 Nomenclatures, vocabularies, terminologies, ontologies and taxonomies in BMHI
   1.18 Informatics methods and tools to support education (incl. flexible and distance learning), use of relevant educational technologies, incl. Internet and World Wide Web
   2.3 Principles of medical decision making and diagnostic and therapeutic strategies
   2.6 Principles of evidence-based medicine and evidence-based nursing

9. **Indicative content (units)**
   - Clinical coding, terminology and systems
   - Clinical pathways and disease management
   - Educational resources for healthcare professionals
   - Knowledge management
   - Natural language processing
   - Ontologies and data definitions
   - Organisational Learning (Pun and Nathai-Balkissoon, 2011)
   - Produce knowledge domain constraint models

10. **Learning & teaching strategy**
    Through an integrated program of lectures, seminars, workshops and tutorials, students engage with
course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies.

11. Assessment strategies
Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.
Collaborative work: oral/visual presentations and written reports.

12. Strategy for feedback and student support during module
The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.
Student should complete a formal feedback to the staff at the end of the module.

13. Indicative resources
2.4.4.1.1.4 Management Information Systems

1. **Code** EHI614

2. **Title** Management Information Systems

3. **Year** Certificate

4. **Credits** 10 Credits

5. **Pre-requisite modules**
   - EHI611 Introduction to Health Informatics
   - EHI612 Health Care management and organization
   - EHI613 Knowledge management in healthcare delivery

6. **Allocation of study & teaching hours**
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   This module explores the connection between information systems (IS) and business performance. It also explores the issues of security, Unique Identifiers and MOI.

8. **Learning outcomes (competences)**
   1.2 Need for systematic information processing in health care, benefits and constraints of information technology in health care
   1.8 Management of information systems in health care (health information management, strategic and tactic information management, IT governance, IT service management, legal and regulatory issues)
   3.3 Ability to communicate electronically, including electronic data exchange, with other health care professionals, internet/intranet use
   3.7 Methods of interfacing and integration of information system components in health care, interfacing standards, dealing with multiple patient identifiers

9. **Indicative content (units)**
   - Management Information Systems
   - Audit trails
   - Data communications and messaging
   - Data quality
   - Information governance
   - Manage information systems - the management of IS
   - Principles of security in a health context
   - The importance of maintaining the integrity of the Master Patient Index and Patient Numbering Systems
   - Unique Identifiers Design and development

10. **Learning & teaching strategy**
    Through an integrated program of lectures, seminars, workshops and tutorials, students engage with course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies

11. **Assessment strategies**
    Individual work: essays, portfolios, presentations and projects, including communication artifacts with
reflective reports.
Collaborative work: oral/visual presentations and written reports

12. **Strategy for feedback and student support during module**

The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.

Student should complete a formal feedback to the staff at the end of the module.

13. **Indicative resources**


2.4.4.1.5 Electronic Health records management and Hospital & health information Systems

1. **Code** EHI615

2. **Title** Electronic Health records management and Hospital & health information Systems

3. **Year** Certificate

4. **Credits** 10 Credits

5. **Pre-requisite modules**
   - EHI611 Introduction to Health Informatics
   - EHI612 Health Care management and organization
   - EHI613 Knowledge management in healthcare delivery
   - EHI614 Management Information Systems

6. **Allocation of study & teaching hours**
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   This module assist students to understand the complexities of managing individual and community based health information. They will be able to understand different approaches of health record modeling; purpose based structuring of health information and the principles of integrated health information management.

8. **Learning outcomes (competences)**
   1.6 Characteristics, functionalities and examples of information systems in health care (e.g. clinical information systems, primary care information systems, etc.)
   1.7 Architectures of information systems in health care; approaches and standards for communication and cooperation and for interfacing and integration of component, architectural paradigms (e.g. service-oriented architectures)
   1.9 Characteristics, functionalities and examples of information systems to support patients and the public (e.g. patient- oriented information system architectures and applications, personal health records, sensor-enhanced information systems)
   1.11 Appropriate documentation and health data management principles including ability to use health and medical coding systems, construction of health and medical coding systems
   1.12 Structure, design and analysis principles of the health record including notions of data quality, minimum data sets, architecture and general applications of the electronic patient record/electronic health record
   1.15 Biomedical modelling and simulation
   3.14 Usability engineering, human-computer interaction, usability evaluation, cognitive aspects of information processing

9. **Indicative content (units)**
   Appointments and Admissions management
   Build a model of a hospital department
   Case note tracking
   Clinical Data structures
   Clinical information needs
   Define EHR
Electronic booking
Electronic healthcare records
Electronic prescribing and medicine management
EPR: architecture, content, views and use
Hospital information systems
Laboratory information systems
Limitations of coding systems
National care-record systems
Specialty-specific clinical information systems
Specific clinical message
Systems currently used by clinicians to gather clinical information
The needs of patients, clients and professionals
Usefulness of clinical recordkeeping standards

10. Learning & teaching strategy
Through an integrated program of lectures, seminars, workshops and tutorials, students engage with course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies.

11. Assessment strategies
Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.
Collaborative work: oral/visual presentations and written reports

12. Strategy for feedback and student support during module
The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.
Student should complete a formal feedback to the staff at the end of the module.

13. Indicative resources
MADDEN, A. 2014. Health informatics and the importance of coding. *Anaesthesia & Intensive Care*
**Medicine,** 15, 62-63.
2.4.1.6 Health Informatics applications including; PACS, MIT, Telemedicine and mobile technologies

1. **Code** EHI616

2. **Title** Health Informatics applications including; PACS, MIT, Telemedicine and mobile technologies.

3. **Year** Certificate

4. **Credits** 10 Credits

5. **Pre-requisite modules**
   - EHI611 Introduction to Health Informatics
   - EHI612 Health Care management and organization
   - EHI613 Knowledge management in healthcare delivery
   - EHI614 Management Information Systems
   - EHI615 Electronic Health records management and Hospital & health information Systems

6. **Allocation of study & teaching hours**
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   This module explores the concepts of ubiquitous computer and technologies to provide assertive support in the population. It explores the use of telemedicine and eHealth solutions for medical support. This is underpinned by an exploration of wireless technologies.

8. **Learning outcomes (competences)**
   1.6 Characteristics, functionalities and examples of information systems in health care (e.g. clinical information systems, primary care information systems, etc.)
   1.9 Characteristics, functionalities and examples of information systems to support patients and the public (e.g. patient-oriented information system architectures and applications, personal health records, sensor-enhanced information systems)
   1.10 Methods and approaches to regional networking and shared care (eHealth, health telematics applications and inter-organizational information exchange)
   2.4 Organisation of health institutions and of the overall health system, interorganizational aspects, shared care
   3.6 Methods of technical informatics/computer science, e.g. network architectures and topologies, telecommunications, wireless technology, virtual reality, multimedia
   3.7 Methods of interfacing and integration of information system components in health care, interfacing standards, dealing with multiple patient identifiers
   3.13 Basic concepts and applications of ubiquitous computing (e.g. pervasive, sensor-based and ambient technologies in health care, health enabling technologies, ubiquitous health systems and ambient assisted-living)
   3.14 Usability engineering, human-computer interaction, usability evaluation, cognitive aspects of information processing

9. **Indicative content (units)**
   Application of telehealth solutions to the elderly population
   Assistive technology
   Biomedical equipment interfaces
   Care at home
   Clinical interface protocol
eHealthcare  
Home monitoring  
How delivery of care differs in primary and secondary care  
Infrastructure  
Mobile technology  
Patient Access to records  
Patient information needs and range of delivery methods  
Patient referral  
Picture archiving and communications systems  
Picture archiving systems (PACs)  
Point-of-care testing  
Robotics  
Technical infrastructures  
Telecare and telemedicine  
Telemonitoring  
The interface between the technology and caring philosophies

10. Learning & teaching strategy
Through an integrated program of lectures, seminars, workshops and tutorials, students engage with course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies.

11. Assessment strategies
Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.  
Collaborative work: oral/visual presentations and written reports

12. Strategy for feedback and student support during module
The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.  
Student should complete a formal feedback to the staff at the end of the module.

13. Indicative resources
Computational Statistics, 5, 432-447.
2.4.4.2 Post graduate diploma

The post-graduate diploma year is supposed to add IM knowledge and skills, leading the student to a broader comprehension of the complete Health Informatics domain and its place in the health system ("the big picture"). It also adds essential project management skills and elements of critical thinking enabling well-considered evaluation of different options. Research skills and scientific reasoning have also been formally included in the program in order to better cope with some research related shortcomings discovered in the existing curriculum.

2.4.4.2.1 Modules
2.4.4.2.1.1 Software based clinical decision making and support systems

1. **Code** EHI621

2. **Title** Software based clinical decision making and support systems

3. **Year** Diploma

4. **Credits** 10 Credits

5. **Pre-requisite modules**
   - All Certificate level modules

6. **Allocation of study & teaching hours**
   - 1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   - 2) 20 hours of self study
   - 3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   This module explores the concepts of information to support health profession in their decision making, and therapeutic strategies. It covers clinical pathways and guidelines and Biomedical modelling and simulation. Underpinning some of the application are the principles of data representation and cognitive aspects of information processing.

8. **Learning outcomes (competences)**
   - 1.3 Efficient and responsible use of information processing tools, to support health care professionals’ practice and their decision making
   - 1.9 Characteristics, functionalities and examples of information systems to support patients and the public (e.g. patient-oriented information system architectures and applications, personal health records, sensor-enhanced information systems)
   - 1.14 Principles of data representation and data analysis using primary and secondary data sources, principles of data mining, data warehouses, knowledge management
   - 1.15 Biomedical modelling and simulation
   - 2.3 Principles of medical decision making and diagnostic and therapeutic strategies
   - 2.6 Principles of evidence-based medicine and evidence-based nursing
   - 3.10 Mathematics: algebra, analysis, logic, numerical mathematics, probability theory and statistics, cryptography
   - 3.12 Methods for decision support and their application to patient management, acquisition, representation and engineering of medical knowledge; construction and use of clinical pathways and guidelines
   - 3.14 Usability engineering, human-computer interaction, usability evaluation, cognitive aspects of information processing

9. **Indicative content (units)**
   - Analyse (clinical) data
   - Clinical content production
   - Clinical decision support
   - Clinical guidelines
   - Data mining
   - Decision making processes in clinical decision making
   - Decision Support Systems performance evaluation
   - Decision support tools
   - Diagnostic process
Futurology
How to do Computational modelling
Interpret clinical data
Patient Information i.e. Health Education
Reconstruction of Clinical Guidelines in a form for a Clinical Decision Support system
Treatment process

10. **Learning & teaching strategy**
Through an integrated program of lectures, seminars, workshops and tutorials, students engage with course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies.

11. **Assessment strategies**
Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.
Collaborative work: oral/visual presentations and written reports

12. **Strategy for feedback and student support during module**
The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.
Student should complete a formal feedback to the staff at the end of the module.

13. **Indicative resources**


2.4.4.2.1.2 Public Health Informatics including, Patient Information Kiosks, Website, and Public Health Systems for Epidemiology, Epidemic Control and GIS

1. **Code** EHI622

2. **Title** Public Health Informatics including, Patient Information Kiosks, Website, and Public Health Systems for Epidemiology, Epidemic Control and GIS.

3. **Year** Diploma

4. **Credits** 10 Credits

5. **Pre-requisite modules**
   All Certificate level modules plus EHI 621

6. **Allocation of study & teaching hours**
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   This module explores the concepts of data representation and information analysis. It will explore clinical information system and epidemiology and public health. Information tools to support education for the public and health professional are explore together with the cognitive aspects of information processing.

8. **Learning outcomes (competences)**
   1.6 Characteristics, functionalities and examples of information systems in health care (e.g. clinical information systems, primary care information systems, etc.)
   1.9 Characteristics, functionalities and examples of information systems to support patients and the public (e.g. patient-oriented information system architectures and applications, personal health records, sensor-enhanced information systems)
   1.10 Methods and approaches to regional networking and shared care (eHealth, health telematics applications and inter-organizational information exchange)
   1.14 Principles of data representation and data analysis using primary and secondary data sources, principles of data mining, data warehouses, knowledge management
   1.15 Biomedical modelling and simulation
   1.18 Informatics methods and tools to support education (incl. flexible and distance learning), use of relevant educational technologies, incl. Internet and World Wide Web
   3.11 Biometry and epidemiology, including study design
   3.14 Usability engineering, human-computer interaction, usability evaluation, cognitive aspects of information processing

9. **Indicative content (units)**
   Community information systems
   Consumer health and patient information
   Data acquisition
   Data for population management
   Data modelling to support analysis
   Data warehousing
   Differentiate relative and actual risk
   Epidemiology
Examples of information systems used to gather clinical data for secondary care purposes

- Geographical Information systems
- Improvements in data quality
- Information services for the public
- Issues of consumerism in Health Informatics
- Techniques to maximize data quality
- Cancer registration
- Use communication, teaching and learning strategy to present information to patients and clients

10. Learning & teaching strategy

Through an integrated program of lectures, seminars, workshops and tutorials, students engage with course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies.

11. Assessment strategies

- Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.
- Collaborative work: oral/visual presentations and written reports

12. Strategy for feedback and student support during module

- The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.
- Student should complete a formal feedback to the staff at the end of the module.

13. Indicative resources


2.4.4.2.1.3 Social-cultural, Legal and Economic impact of Health Informatics

1. **Code** EHI623

2. **Title** Social-cultural, Legal and Economic impact of Health Informatics

3. **Year** Diploma

4. **Credits** 10 Credits

5. **Pre-requisite modules**
   All Certificate level modules plus EHI 621 and EHI 622

6. **Allocation of study & teaching hours**
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   This module explores ethical, legal, and social issues arising in the use of computer-based technology and information systems in the delivery of health care. It also includes Health Informatics ethics and regulatory frameworks.

8. **Learning outcomes (competences)**
   1.3 Efficient and responsible use of information processing tools, to support health care professionals’ practice and their decision making
   1.13 Socio-organizational and socio-technical issues, including workflow/process modelling and reorganization
   1.16 Ethical and security issues including accountability of health care providers and managers and BMHI specialists and the confidentiality, privacy and security of patient data
   1.19 Evaluation and assessment of information systems, including study design, selection and triangulation of (quantitative and qualitative) methods, outcome and impact evaluation, economic evaluation, unintended consequences, systematic reviews and meta-
   2.4 Organisation of health institutions and of the overall health system, interorganizational aspects, shared care
   2.5 Policy and regulatory frameworks for information handling in health care
   3.14 Usability engineering, human-computer interaction, usability evaluation, cognitive aspects of information processing

9. **Indicative content (units)**
   Data Protection Acts and other privacy legislation
   Data security, privacy, confidentiality, access, integrity and standards
   Describe likely trends in Health Policy and IT
   Ethics and governance
   Health Informatics ethics
   Information Systems Strategies - Alignment with External forces
   Justification of eHealth
   Legislation
   Measuring (evaluating) performance of organisations
   National IT policies, strategies and programmes
   Organisational and cultural change
Patient consent
Patient safety
Policy regulatory requirements around medical record documentation
Political influencing
Principles of Freedom of Information and other legislation
Strategy and policy implementation
Comprehends psychology and sociology
The impact of breaches in confidentiality on patient care

10. Learning & teaching strategy
Through an integrated program of lectures, seminars, workshops and tutorials, students engage with
course-related material individually, in groups or collectively as classes. This is supported through a
variety of means, including handouts, on-line learning, workshops and case studies

11. Assessment strategies
   Individual work: essays, portfolios, presentations and projects, including communication artifacts with
   reflective reports.
   Collaborative work: oral/visual presentations and written reports

12. Strategy for feedback and student support during module
The personal tutor would be the main support for learning in general. The Module Leader will be able
for specific feedback on the subject covered and point student to further reading and other learning
resources.
Student should complete a formal feedback to the staff at the end of the module.

13. Indicative resources
2.4.4.2.1.4 Project Management and IT introduction in Health Care Delivery; case studies in Health Informatics

1. **Code** EHI624

2. **Title** Project Management and IT introduction in Health Care Delivery; case studies in Health Informatics

3. **Year** Diploma

4. **Credits** 10 Credits

5. **Pre-requisite modules**
   All Certificate level modules plus EHI 621, EHI 622 and EHI 623

6. **Allocation of study & teaching hours**
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   In this module students will learn the principles of project management as well as of information management within projects. Through exploration of health information projects, students will gain a real-world understandings of how to manage biomedical informatics projects. They will take into account the socio-technical aspects of implementation.

8. **Learning outcomes (competences)**
   1.2 Need for systematic information processing in health care, benefits and constraints of information technology in health care
   1.8 Management of information systems in health care (health information management, strategic and tactic information management, IT governance, IT service management, legal and regulatory issues)
   1.13 Socio-organizational and socio-technical issues, including workflow/process modelling and reorganization
   1.19 Evaluation and assessment of information systems, including study design, selection and triangulation of (quantitative and qualitative) methods, outcome and impact evaluation, economic evaluation, unintended consequences, systematic reviews and meta-
   2.4 Organisation of health institutions and of the overall health system, interorganizational aspects, shared care
   3.8 Handling of the information system life cycle: analysis, requirement specification, implementation and/or selection of information systems, risk management, user training
   3.9 Methods of project management and change management (i.e. project planning, resource management, team management, conflict management, collaboration and motivation, change theories, change strategies)

9. **Indicative content (units)**
   Appraise options for IS and IT solutions
   Benefits realisation
   Capacity management
   Change management
   Clinical governance
   Contract management
   Effective implementation strategy for a clinical system
10. Learning & teaching strategy
Through an integrated program of lectures, seminars, workshops and tutorials, students engage with course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies.

11. Assessment strategies
Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.
Collaborative work: oral/visual presentations and written reports.

12. Strategy for feedback and student support during module
The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.
Student should complete a formal feedback to the staff at the end of the module.

13. Indicative resources

2.4.4.2.1.5 Research, Monitoring and Evaluation in Health Informatics

1. **Code** EHI625

2. **Title** Research, Monitoring and Evaluation in Health Informatics.

3. **Year** Diploma

4. **Credits** 10 Credits

5. **Pre-requisite modules**
   
   All Certificate level modules plus EHI 621, EHI 622 and EHI 623

6. **Allocation of study & teaching hours**
   
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**
   
   This module is an introduction to research in the areas of health informatics and covers a wide range of methods and techniques. These range from epidemiology data and analysis and mapping to measures of patient outputs following medical interventions. Evaluation and Monitoring of programs of care and clinical audit are also included.

8. **Learning outcomes (competences)**
   
   1.19 Evaluation and assessment of information systems, including study design, selection and triangulation of (quantitative and qualitative) methods, outcome and impact evaluation, economic evaluation, unintended consequences, systematic reviews and meta-
   2.7 Health administration, health economics, health quality management and resource management, patient safety initiatives, public health services and outcome measurement
   3.9 Methods of project management and change management (i.e. project planning, resource management, team management, conflict management, collaboration and motivation, change theories, change strategies)
   3.10 Mathematics: algebra, analysis, logic, numerical mathematics, probability theory and statistics, cryptography
   3.11 Biometry and epidemiology, including study design

9. **Indicative content (units)**
   
   Ability to undertake audit
   Clean clinical data
   Clinical audit
   Clinical audit and effectiveness
   Data analysis and statistical presentation
   Evaluate competence of a terminology for a clinical task
   Evaluate different terminology systems that describe health
   Evaluate health system information flow and system analysis
   Evaluate healthcare
   Evaluate integrity of decision support system
   Evaluate IS and IT
   Evaluate the impact of a health informatics intervention
   Formulating questions
   How Health Informatics affects outcomes
   How Health Informatics benefits patients
Patient outcomes and how to measure them
Performance indicators and management
Research methods of information science and healthcare
Statistical methods
Statistics and numerical data
The clinical audit cycle
The significance of accurate data

10. Learning & teaching strategy
Through an integrated program of lectures, seminars, workshops and tutorials, students engage with course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies

11. Assessment strategies
Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.
Collaborative work: oral/visual presentations and written reports

12. Strategy for feedback and student support during module
The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.
Student should complete a formal feedback to the staff at the end of the module.

13. Indicative resources
GILLAM, S. & SIRIWARDENA, A. N. 2013. Frameworks for improvement: clinical audit, the
plandostudyact cycle and significant event audit. *Quality in primary care*, 21, 123-130.


2.4.4.2.1.6 E-health Enterprise Architecture

1. **Code**  EHI626

2. **Title**  E-health Enterprise Architecture

3. **Year**  Diploma

4. **Credits**  10 Credits

5. **Pre-requisite modules**  
   All Certificate level modules plus EHI 621, EHI 622, EHI 623, EHI 624 and EHI 625

6. **Allocation of study & teaching hours**  
   1) 30 contact hours (lectures, tutorials, labs, discussions, role games, case studies etc.)
   2) 20 hours of self study
   3) 50 hours work on off-campus assignment

7. **Brief description of aims and content**  
   This module explores the details of healthcare information technology (HIT) interoperability and standards. The evolution of technology in healthcare, along with the impact on clinical information systems, will be studied. The benefits of integrating healthcare information systems will be investigated, as will the challenges of integrating systems across disparate organizations, healthcare disciplines, and technologies. The value proposition of a standards-based approach to integration will be presented. Students will learn the process of HIT integration projects, and how that parallels the development process of interoperability standards. The module also discusses the ZACHMAN, the Open Group Architecture, and The Health Metric Network Frameworks. Students will gain experience in navigating through standards documents and tools. Students will utilize the skills and knowledge gained to design a standards-based interoperability project addressing a real-world need.

8. **Learning outcomes (competences)**  
   1.7 Architectures of information systems in health care; approaches and standards for communication and cooperation and for interfacing and integration of component, architectural paradigms (e.g. service-oriented architectures)
   1.8 Management of information systems in health care (health information management, strategic and tactic information management, IT governance, IT service management, legal and regulatory issues)
   1.10 Methods and approaches to regional networking and shared care (eHealth, health telematics applications and inter-organizational information exchange)
   1.14 Principles of data representation and data analysis using primary and secondary data sources, principles of data mining, data warehouses, knowledge management
   1.16 Ethical and security issues including accountability of health care providers and managers and BMHI specialists and the confidentiality, privacy and security of patient data
   2.4 Organisation of health institutions and of the overall health system, interorganizational aspects, shared care
   3.3 Ability to communicate electronically, including electronic data exchange, with other health care professionals, internet/intranet use
   3.4 Methods of practical informatics/computer science, especially on programming languages, software engineering, data structures, database management systems, information and system modelling tools, information systems theory and practice, knowledge engineering, (concept) representation and acquisition, software architectures
   3.7 Methods of interfacing and integration of information system components in health care, interfacing standards, dealing with multiple patient identifiers
   3.8 Handling of the information system life cycle: analysis, requirement specification, implementation
and/or selection of information systems, risk management, user training
3.9 Methods of project management and change management (i.e. project planning, resource management, team management, conflict management, collaboration and motivation, change theories, change strategies)

9. **Indicative content (units)**
- Collaborative Internet architectures
- Compare the effects of alternative system designs
- Construct systems
- Data structures
- Design of IS and IT systems for Health
- Develop and implement information and ICT strategy
- Differentiate between technical, syntactic and semantic interoperability
- Implement new information systems
- Implement safe and secure IS and IT systems
- Information and systems requirements
- Interoperability
- Knowledge of performance and process improvement
- Local healthcare strategy
- Messaging standards
- Open systems
- Patient and provider identification
- Procurement
- Provide Business requirements for system
- Systems design and the application of systems theory to design
- Systems implementation
- Systems integration
- The method of primary to secondary to tertiary care referral and the flows of information
- Understand the economic forces driving the healthcare industry
- Unique architectures and standards development

10. **Learning & teaching strategy**
Through an integrated program of lectures, seminars, workshops and tutorials, students engage with course-related material individually, in groups or collectively as classes. This is supported through a variety of means, including handouts, on-line learning, workshops and case studies.

11. **Assessment strategies**
- Individual work: essays, portfolios, presentations and projects, including communication artifacts with reflective reports.
- Collaborative work: oral/visual presentations and written reports

12. **Strategy for feedback and student support during module**
The personal tutor would be the main support for learning in general. The Module Leader will be able for specific feedback on the subject covered and point student to further reading and other learning resources.
- Student should complete a formal feedback to the staff at the end of the module.

13. **Indicative resources**


2.4.4.3 Thesis

The thesis dissertation still accounts for 60 credits in the reviewed curriculum. Student and teaching staff interviews revealed that thesis development has been problematic for most of the students. Therefore it has been foreseen to better prepare students for this part during the certificate and diploma years by:

- Explicitly including research methods and scientific writing skills in the module on "Research, Monitoring and Evaluation in Health Informatics"
- Including mandatory participation in seminars and conferences in Health Informatics (from ordinary attendance to submission and presentation of a paper)
- Programming 1 or 2 training sessions on drafting a thesis proposal for obtaining ethical clearance

2.4.4.4 Program learning outcomes

The program learning outcomes have been completely aligned with IMIA recommendations (see above), with the exception of "2.1 Fundamentals of human functioning and biosciences (anatomy, physiology, microbiology, genomics, and clinical disciplines such as internal medicine, surgery etc.):

Biomedical and Health Informatics Core Knowledge and Skills
1.1 Evolution of informatics as a discipline and as a profession
1.2 Need for systematic information processing in health care, benefits and constraints of information technology in health care
1.3 Efficient and responsible use of information processing tools, to support health care professionals’ practice and their decision making
1.4 Use of personal application software for documentation, personal communication including Internet access, for publication and basic statistics
1.5 Information literacy: library classification and systematic health related terminologies and their coding, literature retrieval methods, research methods and research paradigms
1.6 Characteristics, functionalities and examples of information systems in health care (e.g. clinical information systems, primary care information systems, etc.)
1.7 Architectures of information systems in health care; approaches and standards for communication and cooperation and for interfacing and integration of component, architectural paradigms (e.g. service-oriented architectures)
1.8 Management of information systems in health care (health information management, strategic and tactic information management, IT governance, IT service management, legal and regulatory issues)
1.9 Characteristics, functionalities and examples of information systems to support patients and the public (e.g. patient-oriented information system architectures and applications, personal health records, sensor-enhanced information systems)
1.10 Methods and approaches to regional networking and shared care (eHealth, health telematics applications and inter-organizational information exchange)
1.11 Appropriate documentation and health data management principles including ability to use health and medical coding systems, construction of health and medical coding systems
1.12 Structure, design and analysis principles of the health record including notions of data quality, minimum data sets, architecture and general applications of the electronic patient record/electronic health record
1.13 Socio-organizational and socio-technical issues, including workflow/process modelling and reorganization
1.14 Principles of data representation and data analysis using primary and secondary data sources, principles of data mining, data warehouses, knowledge management
1.15 Biomedical modelling and simulation
1.16 Ethical and security issues including accountability of health care providers and managers and BMHI specialists and the confidentiality, privacy and security of patient data
1.17 Nomenclatures, vocabularies, terminologies, ontologies and taxonomies in BMHI
1.18 Informatics methods and tools to support education (incl. flexible and distance learning), use of relevant educational technologies, incl. Internet and World Wide Web
1.19 Evaluation and assessment of information systems, including study design, selection and triangulation of (quantitative and qualitative) methods, outcome and impact evaluation, economic evaluation, unintended consequences, systematic reviews and meta-analysis, evidence-based health informatics

Medicine, Health and Biosciences, Health System Organisation
2.2 Fundamentals of what constitutes health, from physiological, sociological, psychological, nutritional, emotional, environmental, cultural, spiritual perspectives and its assessment
2.3 Principles of medical decision making and diagnostic and therapeutic strategies
2.4 Organisation of health institutions and of the overall health system, interorganizational aspects, shared care
2.5 Policy and regulatory frameworks for information handling in health care
2.6 Principles of evidence-based medicine and evidence-based nursing
2.7 Health administration, health economics, health quality management and resource management, patient safety initiatives, public health services and outcome measurement

Informatics/Computer Science, Mathematics, Biometry
3.1 Basic informatics terminology like data, information, knowledge, hardware, software, computer, networks, information systems, information systems management
3.2 Ability to use personal computers, text processing and spreadsheet software, easy-to-use database management systems
3.3 Ability to communicate electronically, including electronic data exchange, with other health care professionals, internet/intranet use
3.4 Methods of practical informatics/computer science, especially on programming languages, software engineering, data structures, database management systems, information and system modelling tools, information systems theory and practice, knowledge engineering, (concept) representation and acquisition, software architectures
3.5 Methods of theoretical informatics/computer science, e.g. complexity theory, encryption/security
3.6 Methods of technical informatics/computer science, e.g. network architectures and topologies, telecommunications, wireless technology, virtual reality, multimedia
3.7 Methods of interfacing and integration of information system components in health care, interfacing standards, dealing with multiple patient identifiers
3.8 Handling of the information system life cycle: analysis, requirement specification, implementation and/or selection of information systems, risk management, user training
3.9 Methods of project management and change management (i.e. project planning, resource management, team management, conflict management, collaboration and motivation, change theories, change strategies)
3.10 Mathematics: algebra, analysis, logic, numerical mathematics, probability theory and statistics, cryptography
3.11 Biometry and epidemiology, including study design
3.12 Methods for decision support and their application to patient management, acquisition, representation and engineering of medical knowledge; construction and use of clinical pathways and guidelines
3.13 Basic concepts and applications of ubiquitous computing (e.g. pervasive, sensor-based and ambient technologies in health care, health enabling technologies, ubiquitous health systems and ambient assisted-living)
3.14 Usability engineering, human-computer interaction, usability evaluation, cognitive aspects of information processing
2.4.4.5 Student profile

The postgraduate program can admit students from all over the world, specifically targeting those from the great lake region, who are highly motivated and have an aptitude for, and interest in, solving health-related problems with the use of modern technology tools. Students must be sufficiently proficient in English.

2.4.4.6 Admission criteria

Applicants for admission to the postgraduate program in health informatics must be graduates from medical, nursing and health related backgrounds or from engineering and applied sciences disciplines, and they should be expected to satisfy the general admission requirements as specified in the Postgraduate Framework & Regulations. Applicants should be able to demonstrate potential within three particular domains:

- **Academically**, applicants will typically possess a degree in science, technology or a medical/public health area.
- **Technically**, applicants will be required to have some experience of the use of IT in their workplace. They will be computer literate in basic computer applications such as word processing, email and use of the Internet. Optimally they will have access to computers at work or at home.
- **Contextually**, applicants should ideally have a couple of years’ experience working in a healthcare (related) environment. Applicants should typically work within the health services although the program may also attract those working in the commercial sector whose role requires them to interact significantly with the health sector. In some circumstances, applicants with less than two years’ health care experience can be accepted, provided that the quality of their experience is regarded as sufficiently high for them to understand the healthcare and technology working environments.

2.4.4.7 Teaching strategy

1 credit is equivalent to 10 working hours for the student. A 10 credit module is equivalent to 100 student hours, composed on average of 30 contact hours, 50 off-campus assignment hours and 20 hours of self-study. Each module shall be taught within a period of 4 weeks combining any of the following elements:

2.4.4.7.1 On campus contact hours

On campus teaching will be a mixture of an average of 30 contact hours, including on average 20 hours of lectures and 10 hours of tutorials, case studies, discussions and labs.

2.4.4.7.2 Off-campus work

Off-campus assignment work is composed of different components which add up on average to 50 student working hours:

- Assignments (problem solving assignments, seminars, laboratory/project work, structured exercises, group work and individual work)
- Field visits
- Seminars and conferences (attendance and presentations)

2.4.4.7.3 Self study

It is expected that on average an extra 20 student hours will be needed for preparing classes and digesting knowledge related to the topic.

2.4.4.8 Assessment strategy

The following assessment strategy is proposed for the MSc HI curriculum:

**Post graduate certificate & diploma**
Assignments including seminars & conferences: 60%
Formal examination: 40%

Thesis
• Research dissertation: 100%

2.4.5 Budget

The budget foreseen for running the MSc HI program has been calculated as follows:

<table>
<thead>
<tr>
<th>Estimated income per year</th>
<th>Unit price</th>
<th>#</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student fees</td>
<td>$2,500,00</td>
<td>40</td>
<td>$ 100,000,00</td>
</tr>
<tr>
<td>MoH, MINEDUC, NGOs, short courses, BD</td>
<td>$100,040,00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total income</td>
<td></td>
<td></td>
<td>$ 200,040,00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Estimated costs per year</th>
<th>Unit price</th>
<th>#</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>On campus teaching</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foreign volunteer teaching staff/module</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Travel</td>
<td>$1,750,00</td>
<td>8</td>
<td>$ 14,000,00</td>
</tr>
<tr>
<td>Lodging</td>
<td>$1,680,00</td>
<td>8</td>
<td>$ 13,440,00</td>
</tr>
<tr>
<td>Expenses</td>
<td>$700,00</td>
<td>8</td>
<td>$ 5,600,00</td>
</tr>
<tr>
<td>Local teaching staff</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salaries/module</td>
<td>$1,000,00</td>
<td>4</td>
<td>$ 4,000,00</td>
</tr>
<tr>
<td>Classroom/day</td>
<td>$50,00</td>
<td>200</td>
<td>$ 10,000,00</td>
</tr>
<tr>
<td>Video projector</td>
<td>$2,000,00</td>
<td>1</td>
<td>$ 2,000,00</td>
</tr>
<tr>
<td>Blackboard/whiteboard</td>
<td>$500,00</td>
<td>1</td>
<td>$ 500,00</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>$ 49,540,00</td>
</tr>
<tr>
<td>Assisting academic staff</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salaries/year (parttime)</td>
<td>$12,000,00</td>
<td>2</td>
<td>$ 24,000,00</td>
</tr>
<tr>
<td>Personal computer equipment</td>
<td>$1,000,00</td>
<td>2</td>
<td>$ 2,000,00</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>$ 26,000,00</td>
</tr>
<tr>
<td>Field activities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seminars participation/student</td>
<td>$150,00</td>
<td>40</td>
<td>$ 6,000,00</td>
</tr>
<tr>
<td>Transportation costs/student</td>
<td>$150,00</td>
<td>40</td>
<td>$ 6,000,00</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>$ 12,000,00</td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dissertation supervision</td>
<td>$500,00</td>
<td>20</td>
<td>$ 10,000,00</td>
</tr>
<tr>
<td>Program supervision</td>
<td>$5,000,00</td>
<td>1</td>
<td>$ 5,000,00</td>
</tr>
<tr>
<td>Electronic learning environment</td>
<td>$5,000,00</td>
<td>1</td>
<td>$ 5,000,00</td>
</tr>
<tr>
<td>Course material conversion</td>
<td>$20,000,00</td>
<td>1</td>
<td>$ 20,000,00</td>
</tr>
<tr>
<td>Video conferencing &amp; capturing equipment</td>
<td>$5,000,00</td>
<td>1</td>
<td>$ 5,000,00</td>
</tr>
<tr>
<td>Internet connectivity/month</td>
<td>$250,00</td>
<td>12</td>
<td>$ 3,000,00</td>
</tr>
<tr>
<td>Office renting</td>
<td>$2,500,00</td>
<td>12</td>
<td>$ 30,000,00</td>
</tr>
<tr>
<td>Office costs</td>
<td>$1,500,00</td>
<td>12</td>
<td>$ 18,000,00</td>
</tr>
<tr>
<td>Seminars &amp; conferences</td>
<td>$5,000,00</td>
<td>1</td>
<td>$ 5,000,00</td>
</tr>
<tr>
<td>External examinators</td>
<td>$10,000,00</td>
<td>1</td>
<td>$ 10,000,00</td>
</tr>
<tr>
<td>Demonstration application server</td>
<td>$1,500,00</td>
<td>1</td>
<td>$ 1,500,00</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>$ 112,500,00</td>
</tr>
</tbody>
</table>
Total costs $ 200,040,00
Part III
3 Regional e-Health Center of Excellence (REHCE)

The MSc HI informatics curriculum is one of the education instruments in the hands of the REHCE. The functioning of the REHCE has a clear impact on the success of the MSc HI program. Therefore, the reviewers also looked at some of the core elements, which influence and could potentially improve the program results.

3.1 Justification

A number of fundamental community objectives justify the existence of REHCE within the CMHS:
- Improvement of patient care
- Improvement of health service provisioning
- Optimization of health management business processes
- Health services cost control
- Health related research promotion

3.2 Objectives

The center has essentially 3 objectives:

3.2.1 Education

3.2.1.1 Programs

Education programs consist of post-graduate certificate (IS and IT oriented) and diploma (IM-oriented adding project management skills) programs. It is suggested that these program primarily focus on real world needs and systems that have been identified in the Rwandan context. Fundamental health informatics research is considered to be out of scope at the moment of reviewing the MSc HI curriculum.

3.2.1.2 Methods

Education methods consist of:
- **On campus teaching** in the form of lectures which are preferably completed with labs, tutorials and hands-on training sessions.
- **Off-campus practical education** including internships, assignments, field visits and participation in health informatics seminars and conferences in order to demonstrate the purpose of learning to the students.
- The use of an **electronic learning system** for hosting educational content, assignments work, dissertations and research documents.
- **Local assisting academic staff** should provide assistance to students with lecture content digestion, assignments execution, dissertation development and research activities.

3.2.2 Training

Training activities can focus on IS targeted short courses which may serve multiple purposes and clients:
- Continuing Professional Development: development of practical IT and IS skills for health workers
- NGOs: fulfilling application oriented training needs
- Vertical health programs: getting application skills in line with program needs
- Health facilities: development of basic IT skills for hospital staff and developing skills on specific hospital software applications
3.2.3  **Research**

Research activities remained almost unexploited within REHCE at the moment of the MSc HI curriculum review. The following approach for improving research output is suggested:

3.2.3.1  **National & international collaboration**

- Collaboration with NGOs, health facilities (mainly reference level) and health insurance organizations for development of masters dissertations.
- Collaboration with NGOs, health facilities (mainly reference level) and health insurance organizations for undertaking off-campus assignments
- Organizing post-graduate/PhD research initiatives in collaboration with other international universities, the MoH, NGOs and national reference health facilities (which explicitly offered this opportunity)

3.2.3.2  **Publications**

Publications are essential for academic survival and integration in the international Health Informatics community. Therefore affiliation of REHCE to relevant Health Informatics Associations (HELINA, IMIA) and international universities active in the field of Health Informatics is strongly advised. Active participation of REHCE faculty and students in national and international Health Informatics conferences is an important step towards academic networking and collaborative publication.

3.3  **Institution**

On an institutional level, the practical organization of the REHCE activities must cover a number of infrastructure and staffing requirements.

3.3.1  **Infrastructure**

Classroom infrastructure has been evaluated to be rather minimalistic during the last year of the program. Classrooms and furniture should be adapted to the teaching methods used (small group work, individual work, lectures, IT labs...) which has not always been the case in the past.

Based on the fact that several well-equipped teaching health facilities are hosting CMHS/REHCE students during their internship, may justify the sharing of teaching and training infrastructure with these facilities (CHUK, CHUB, KFH), somehow alleviating the infrastructural burden for the CMHS.

3.3.1.1  **Equipment**

A minimum set of equipment must be foreseen for running a MSc HI program:

3.3.1.1.1  **Electronic learning environment**

The use of the complete set of features of the electronic learning environment should be mandatory. Therefore, training of teaching staff and students may need to be programmed in order to integrate:
- Course materials
- A software repository of applications used in the program
- Student and teacher evaluation tools including self-assessment, formal exams, assignments and course quality scoring
- Collaboration tools moderated by local assisting academic staff

3.3.1.1.2  **Video conferencing & capturing tools**

Video conference tools should be easy to set up in low bandwidth conditions and allow teacher-student interaction. Most often, course content needs to be adapted in order to make it usable in a remote teaching setting. Simple and cheap solutions such as WebEx may be worth testing.
3.3.1.3 On campus teaching instruments

On campus teaching requires also a minimum of teaching instruments:

- Stable wireless intranet/internet connectivity with appropriate access to power plugs (almost all students own a laptop computer that can be used in many of the modules)
- A blackboard with all necessary accessories
- A video projector with a projection screen

3.3.1.2 Shared virtual health informatics center

The setup of a shared virtual health informatics center for the different nursing schools integrated in the CMHS can provide an answer to some but not all of the education and research needs of CMHS. For the time being, virtual teaching is evaluated to be complementary rather than substitutive.

3.3.2 Staffing

3.3.2.1 Foreign staff dependency

REHCE teaching staff heavily relies on foreign experts. Only 20% of the teaching staff is local and 80% was recruited in the US, India and Belgium.

3.3.2.2 Local academic capacity building

The building of local academic capacity is key for the sustainability of the center. During the interviews the reviewers were able to identify several locally available qualified persons which could be excellent candidates for REHCE teaching. The lack of local academic experience has multiple causes:

- A very young workforce in the Health Informatics field
- Health Informatics itself is a young discipline
- Academic capacity building is a lengthy process

In order to progressively build local capacity, the reviewers suggest the progressive integration of local academics into the program (in chronological order):

- Teaching of individual knowledge units in modules
- Providing assistance to students with assignments
- Assist students with their research activities (dissertations)
- Teaching of a complete module
- Doing their own Health Informatics research

Local academic staff should also be assigned to students as personal tutors that can then assist students with language issues, scientific thinking problems and other general issues.

3.4 Finance

Financial resources for REHCE can be derived from different sources:

3.4.1 Financial sources

3.4.1.1 Government funding

Mainly MoH and MINEDUC, who already have committed to the funding of the MSc HI program.

3.4.1.2 NGO & health partner funding

A number of NGOs and health programs who's activities heavily rely on the availability of solid health information management solutions may be interested in co-funding the REHCE in order to support capacity building in the health informatics domain.
3.4.1.3 **Short courses**

The organization of short courses targeting specific application training needs from stakeholders such as health facilities, NGOs, the MoH and health programs offers perspectives from remunerated teaching activities. The growing needs for Continuing Professional Development (CPD) in healthcare is also an opportunity for training health workers to become proficient in the usage of IS. In the past, the REHCE already successfully organized a series of such short courses for the Kigali University Teaching Hospital.

Besides the teaching of short courses, the REHCE can also play a fee-based role in accreditation of health informatics training content and courses.

3.4.1.4 **Knowledge center**

Based on the accumulated tacit and explicit knowledge in the Health Informatics field, the REHCE should become in the future a national and regional reference for Health Informatics domain knowledge, serving information requests from institutional and private health care stakeholders.

3.4.1.5 **Business development center (BDC)**

Based on its unique position in Rwanda and in the East-African region, the REHCE has the potential of developing a number of business activities besides education and teaching. Such activities can of course reinforce the financial autonomy of the center and should concentrate on the development of an incubation environment for innovative Rwandan Health Informatics startups.

Bridging REHCE education activity to local and regional business development may provide a very strong case for the center and for Rwanda as an already very visible partner in the EAC.

3.4.1.5.1 **Objectives**

The main objectives of the BDC can be summarized as follows:

- Job creation in the broad field of health informatics
- Health system innovation through research and evaluation
- Health Informatics related knowledge creation and retention
- Creation and incubation of innovative Health Informatics spin-offs in Rwanda

3.4.1.5.2 **Activities**

For reaching the objectives, the following activities must be developed:

3.4.1.5.2.1 **Training**

The organization of targeted short courses and stakeholder-specific training in the Health Informatics field falls into this category. This activity must not be in conflict with education tasks that have been assigned to the center. Typical eligible training domains are:

- Health information management applications
- Data management capabilities
- Information security & privacy implementation

3.4.1.5.2.2 **Testing & evaluation**

Based on the availability of staff with specialized domain knowledge, REHCE may serve as an independent partner for testing and evaluation of new information technology and systems. Accreditation and certification can thereby be part of the assignment.

3.4.1.5.2.3 **Audit**

Many healthcare stakeholders will be interested in independent auditing of information technology and systems that have been delivered within the context of public tenders. Tender specification conformance
validation is thereby one of the key tasks that can be assigned to REHCE. Audit activities can thereby as well concentrate on "Solutions and Systems" as on "Business Processes and Organization".

3.4.1.5.2.4 Software development

Development of early prototypes and user interfaces can also be considered part of the business development domain of REHCE. Such developments incorporate an important academic and research component and therefore require appropriately trained staff that is not always available in software development, implementation or integration companies.

3.4.1.5.2.5 Solutions design

The design of high level IT and IS based solutions for health system problems can be another task for REHCE. Such activity typically requires solid biomedical modeling skills as well as a good knowledge of health care organization, legislation and social factors, business processing re-engineering and systems integration which all may be found in future REHCE MSc HI graduates and teaching staff.

3.4.1.5.2.6 Consultancy

Finally, the many Health Informatics related consultancy tenders in sub-Saharan Africa offer an important low-overhead opportunity for REHCE for business generation. Rwanda already has a number of repudiated enterprise level health information management solutions which can serve as an example for many other countries in the region. The Rwandan experience can therefore serve as solid a consultant's reference for development partners that wish to develop Health Informatics capacity and solutions in other African areas. Consultancy activities can be developed
- for national and international tenders
- for providing implementation assistance
- in the form of project management and coordination services
- for e-Health Enterprise Architecture development

3.4.1.5.3 Staffing

REHCE BD staffing should require minimal overhead. Recruitment can mainly be project-based and experts may be contracted on freelance basis.

3.4.1.5.3.1 Management

The BD unit essentially requires a CEO, CFO and secretary services as permanent management staff members.

3.4.1.5.3.2 Experts

Health Informatics domain experts can be recruited from REHCE graduates or as independent experts which are then managed in national and international freelance pools.

3.4.1.5.3.3 Technical staff

Technical staff may be limited to a local system administrator (taking care of BD unit networks and hardware) and a number of developers that are recruited on a project basis (software development should not be at the core of the BD unit activity for reasons of disloyal competition risks).

3.4.1.5.3.4 Researchers

Research activity must preferably be developed in collaboration with international academic institutions with a focus on applied doctoral and post-doctoral research for which funding can be secured.
3.4.1.5.4 Infrastructure

REHCE infrastructure must be in line with the developed activities and includes:

3.4.1.5.4.1 Classrooms

Mainly for hosting the training activities.

3.4.1.5.4.2 Offices

Necessary for center management purposes and accommodating testing and evaluation, software development, solutions design and consultancy activities.

3.4.1.5.4.3 Health informatics business incubator

For REHCE staff and graduates with entrepreneurial skills that can identify potential Health Informatics business cases, REHCE should provide a subsidized business incubation infrastructure such as secretary services, meeting rooms, labs and offices. REHCE can thereby participate in startup capital or limit its activity to startup facilitation and business or technical mentoring.
Part IV
4 Referenced literature
Recommendations of the International Medical Informatics Association (IMIA) on Education in Biomedical and Health Informatics - 1st Revision

International Medical Informatics Association,
Working Group on Health and Medical Informatics Education

Prepared by: John Mantas, Elske Ammenwerth, George Demiris, Arie Hasman, Reinhold Haux, William Hersh, Evelyn Hovenga, KC Lun, Heimar Marin, Fernando Martin-Sanchez, Graham Wright; IMIA Recommendations on Education Task Force

Summary
Objective: The International Medical Informatics Association (IMIA) agreed on revising the existing international recommendations in health informatics / medical informatics education. These should help to establish courses, course tracks or even complete programs in this field, to further develop existing educational activities in the various nations and to support international initiatives concerning education in biomedical and health informatics (BMHI), particularly international activities in educating BMHI specialists and the sharing of courseware.
Method: An IMIA task force, nominated in 2006, worked on updating the recommendation’s first version. These updates have been broadly discussed and refined by members of IMIA’s National Member Societies, IMIA’s Academic Institutional Members and by members of IMIA’s Working Group on Health and Medical Informatics Education.
Results and conclusions: The IMIA recommendations centre on educational needs for health care professionals to acquire knowledge and skills in information processing and information and communication technology. The educational needs are described as a three-dimensional framework. The dimensions are: 1) professionals in health care (e.g. physicians, nurses, BMHI professionals), 2) type of specialisation in BMHI (IT users, BMHI specialists), and 3) stage of career progression (bachelor, master, doctorate). Learning outcomes are defined in terms of knowledge and practical skills for health care professionals in their role (a) as IT user and (b) as BMHI specialist. Recommendations are given for courses/course tracks in BMHI as part of educational programs in medicine, nursing, health care management, dentistry, pharmacy, public health, health record administration, and informatics/computer science as well as for dedicated programs in BMHI (with bachelor, master or doctor degree).
To support education in BMHI, IMIA offers to award a certificate for high quality BMHI education. It supports information exchange on programs and courses in BMHI through its Working Group on Health and Medical Informatics Education.

Keywords
Medical informatics; health informatics, biomedical informatics, education, recommendations, International Medical Informatics Association, IMIA.
# Table of Contents

1 Introduction .......................................................................................................................... 1  
  1.1 Why Do We Need Biomedical and Health Informatics Education? ........................ 1  
  1.2 IMIA Recommendations for Biomedical and Health Informatics Education ........ 3  
2 General Considerations ......................................................................................................... 4  
  2.1 Key Principles of the IMIA Recommendations ............................................................ 4  
  2.2 Structural Outline of the IMIA Recommendations ......................................................... 5  
3 Recommendations for Learning Outcomes in Biomedical and Health Informatics ........ 6  
4 Recommendations for Courses/Course Tracks in Biomedical and Health Informatics as Part of Educational Programs ................................................................. 14  
  4.1 General Remarks ............................................................................................................ 14  
  4.2 Recommendations for Biomedical and Health Informatics Courses as Part of Medical, Nursing, Health Care Management, Dentistry, Pharmacy, and Public Health Programs ........................................................................................................... 14  
     Courses/Course Tracks for IT Users .................................................................................. 14  
     Courses/Course Tracks for BMHI Specialists ................................................................. 15  
  4.3 Recommendations for Biomedical and Health Informatics Courses as Part of Health Record Administration Programs ................................................................. 15  
     Courses/Course Tracks for IT Users ................................................................................ 15  
     Courses/Course Tracks for Biomedical and Health Informatics Specialists ................. 16  
  4.4 Recommendations for Biomedical and Health Informatics Courses as Part of Informatics/Computer Science Programs ................................................................. 16  
     Courses/Course Tracks for Biomedical and Health Informatics Specialists ................. 16  
5 Recommendations for Dedicated Educational Programs in Biomedical and Health Informatics ......................................................................................................................... 17  
  5.1 Recommendations for Bachelor Programs in Biomedical and Health Informatics .... 17  
  5.2 Recommendations for Master and Doctoral Programs in Biomedical and Health Informatics ................................................................................................................. 18  
6 Recommendations for Continuing Education .................................................................... 19  
  6.1 Continuing Education in Biomedical and Health Informatics .................................... 19  
  6.2 Life-Long Learning ......................................................................................................... 19  
7 Other Recommendations ...................................................................................................... 19  
  7.1 How to Commence with Biomedical and Health Informatics Education ................... 19  
  7.2 Modes of Education ....................................................................................................... 20  
  7.3 Qualified Teachers ......................................................................................................... 20  
  7.4 Recognized Qualifications ............................................................................................. 20  
8 IMIA Support for Programs and Courses in Biomedical and Health Informatics ........... 20  
  8.1 IMIA Certification .......................................................................................................... 20  
  8.2 International Programs, International Exchange of Students and Teachers ............... 21  
9 Information Exchange on Programs and Courses in Biomedical and Health Informatics supported by IMIA ................................................................................................. 21  
10 Concluding Remarks ......................................................................................................... 21  
References .............................................................................................................................. 22
1 Introduction

The first version of the “Recommendations of the International Medical Informatics Association (IMIA) on education in health and medical informatics” [1] has been widely used. Because of the tremendous progress in and the evolution of the field of health and biomedical informatics, the content of these recommendations are, however, not fully up-to-date. As there is still a continuous need for such recommendations, a revision was necessary.

The revised version, presented here, is based on the original recommendations and the fundamental work following those recommendations ([2-9] and table 1 in section 1.2).

There is increasing evidence that health information technology (HIT) improves health, health care, public health, and biomedical research. A number of recent systematic reviews have documented the evidence in favour of clinical decision support [10], information and communication technology (IT) interventions [11], and telemedicine [12]. This has led to widespread adoption of HIT around the world [13]. In addition to a growing range of research and application fields in biomedical and health informatics (BMHI) [14, 15], there is also growth in related areas of BMHI, such as clinical research informatics [16-18] and bioinformatics [19]. The growth of HIT has also led to the recognition of the need for educational programs to train professionals to develop, implement, and evaluate these systems. While this need has been recognized worldwide [20], there have been few international efforts with some notable exceptions [21, 22].

1.1 Why Do We Need Biomedical and Health Informatics Education?

Despite the documented benefits, there are still barriers to HIT in clinical settings, including a mismatch of return on investment between those who pay and those who benefit, challenges to ameliorate workflow in clinical settings, lack of standards and interoperability, and concerns about privacy and confidentiality [23-25]. Another barrier, lesser studied and quantified but increasingly recognized, is the lack of characterization of the workforce and its training needed to most effectively implement HIT systems [26, 27].

An additional challenge is that there are various definitions of the field of BMHI [26, 28, 29]. Furthermore, the field has difficulty agreeing on the adjective in front of the word informatics (i.e., medical, biomedical, and/or health) as well as whether a practitioner should be called an informaticist or informatician (this paper uses the latter). We also do not know where pure IT ends and informatics begins. For example, the individual who installs applications on a desktop computer in a hospital probably does not need formal training in informatics, although the CIO and his or her project leaders certainly do. This has led to calls for BMHI to become a professional discipline [30] and for it to acquire the attributes of a profession, such as a well-defined set of competencies, certification of fitness to practice, shared professional identity, lifelong commitment, and a code of ethics [31].

A number of efforts to establish formal certification have been and are underway. E.g., in Germany, medical informatics courses are mandatory for medical students since the 1970s ([32]). In the US, certification in nursing informatics has been available for over a decade [33]. More recently, the American Medical Informatics Association (AMIA) has proposed a plan for a physician board sub-specialization, which will be followed for other non-physician doctorally prepared individuals [34, 35].

Despite these challenges, a number of organizations have deemed competence in BMHI to be important. Competencies in the field have been developed for a variety of disciplines within BMHI. In table 1 some major publications on the development of such competencies are listed as examples.
Table 1: Some major publications on competencies in biomedical and health informatics.

<table>
<thead>
<tr>
<th>Organization (Reference)</th>
<th>Year</th>
<th>Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td>University of Heidelberg and the University of Applied Sciences Heilbronn medical informatics program [36, 37]</td>
<td>1972</td>
<td>dedicated BMHI programs</td>
</tr>
<tr>
<td>German Association for Medical Informatics, Biometry and Epidemiology (gmds) and German Society for Computer Science (Reisensburg Conference) [38, 39]</td>
<td>1973</td>
<td>BMHI and computer science</td>
</tr>
<tr>
<td>Association for Computing Machinery [40]</td>
<td>1978</td>
<td>BMHI and computer science</td>
</tr>
<tr>
<td>Association of the Medical Colleges [41]</td>
<td>1984</td>
<td>BMHI and medicine</td>
</tr>
<tr>
<td>gmds [42]</td>
<td>1992</td>
<td>BMHI in general</td>
</tr>
<tr>
<td>Concerted Action on Education and Training in Health Care Informatics (EDUCTRA) [43]</td>
<td>1992</td>
<td>BMHI for health professionals</td>
</tr>
<tr>
<td>Council of Europe Committee of Ministers [44]</td>
<td>1995</td>
<td>BMHI for health professionals</td>
</tr>
<tr>
<td>NIGHTINGALE project [45, 46]</td>
<td>1996</td>
<td>nursing informatics</td>
</tr>
<tr>
<td>IT-EDUCTRA project [43]</td>
<td>1996</td>
<td>BMHI for health professionals</td>
</tr>
<tr>
<td>English National Board for Nursing, Midwifery and Health Visiting [47]</td>
<td>1996</td>
<td>BMHI for nurses, midwives and health visitors</td>
</tr>
<tr>
<td>Association of American Medical Colleges [49]</td>
<td>1999</td>
<td>BMHI and medicine</td>
</tr>
<tr>
<td>International Medical Informatics Association [1]</td>
<td>1999</td>
<td>BMHI in general</td>
</tr>
<tr>
<td>Schleyer [50]</td>
<td>1999</td>
<td>dental informatics</td>
</tr>
<tr>
<td>UK National Health Service [51]</td>
<td>2001</td>
<td>BMHI in general</td>
</tr>
<tr>
<td>Staggers et al [52]</td>
<td>2001</td>
<td>nursing informatics</td>
</tr>
<tr>
<td>Covvey et al [53]</td>
<td>2001</td>
<td>BMHI in general</td>
</tr>
<tr>
<td>Mantas and Hasman [54]</td>
<td>2001</td>
<td>nursing informatics</td>
</tr>
<tr>
<td>O’Caroll et al. [55]</td>
<td>2002</td>
<td>BMHI for public health (PH) professionals</td>
</tr>
<tr>
<td>Curran [56]</td>
<td>2003</td>
<td>nursing informatics</td>
</tr>
<tr>
<td>American College of Medical Informatics [57]</td>
<td>2004</td>
<td>bioinformatics</td>
</tr>
<tr>
<td>American Health Information Management Association (AHIMA) [58]</td>
<td>2004</td>
<td>health information management</td>
</tr>
<tr>
<td>Hovenga and Mantas [20]</td>
<td>2004</td>
<td>BMHI in general</td>
</tr>
<tr>
<td>Hovenga and Garde [59]</td>
<td>2006</td>
<td>health informatics</td>
</tr>
<tr>
<td>Ivanitskaya et al. [60]</td>
<td>2006</td>
<td>health information literacy</td>
</tr>
<tr>
<td>AHIMA [61]</td>
<td>2007</td>
<td>health information management (HIM)</td>
</tr>
<tr>
<td>Canadian Health Informatics Association [62]</td>
<td>2007</td>
<td>health informatics</td>
</tr>
<tr>
<td>Medical Library Association [63]</td>
<td>2007</td>
<td>health science librarians</td>
</tr>
<tr>
<td>Pigott et al. [64]</td>
<td>2007</td>
<td>health informatics</td>
</tr>
<tr>
<td>American Medical Informatics Association (AMIA) and Oregon Health and Science University [65]</td>
<td>2008</td>
<td>BMHI in general</td>
</tr>
<tr>
<td>AMIA and AHIMA [66]</td>
<td>2008</td>
<td>BMHI and HIM</td>
</tr>
<tr>
<td>Karras et al. [67]</td>
<td>2008</td>
<td>PH informatics</td>
</tr>
<tr>
<td>AMIA [68]</td>
<td>2009</td>
<td>clinical informatics</td>
</tr>
</tbody>
</table>
1.2 IMIA Recommendations for Biomedical and Health informatics Education

There are many opportunities worldwide for obtaining education in this field. In some countries, there are extensive educational components in BMHI at different levels of education and for the different health care professions. Increasingly dedicated BMHI programs exist (i.e. organized, structured sets of course offerings aimed at preparing participants for specific career paths and culminating in a BMHI degree, diploma or leaving certificate). Many other countries have not, or at least not sufficiently, established such opportunities until now, with all the consequences concerning the quality and effectiveness of health care. Lists of educational programs have been made available at a variety of Web sites, e.g.:

- BMHI programs world wide [69] (University of Freiburg, Germany),
- BMHI programs in [70] and outside [71] North America (AMIA).

In 2007 IMIA, the International Medical Informatics Association [72] endorsed and published its strategic plan, “Towards IMIA 2015” [73-76]. Education in BMHI was listed among its six core subject-focused ‘sectors’. Recognizing that the original version of the IMIA recommendations for education had become outdated, in 2006 a task force was established under the auspices of IMIA’s Working Group on Health and Medical Informatics Education “to consider and undertake any necessary work to update the IMIA Recommendations on Education in Health and Medical Informatics” ([77]). These recommendations will continue to stimulate the further development of existing educational activities in the various nations and to support international initiatives concerning education in BMHI.

Because a variety of educational and health care systems exist all over the world, programs, courses and course tracks in BMHI may vary in different countries. In spite of this variability, basic similarities in BMHI education can be identified and used as a framework for recommendations. Such recommendations are also necessary for enabling an international exchange of students and teachers and for establishing international programs.

The IMIA recommendations, presented here, should be regarded as a framework for national initiatives in BMHI education, and for constituting international programs and exchange of students and teachers in this field. They shall also encourage and support the sharing of courseware.
2  General Considerations

2.1  Key Principles of the IMIA Recommendations

In order to provide good quality health care, training and education in biomedical and health medical informatics is needed:

H for various *Health* care professions,
E in different modes of *Education*,
A with different, *Alternate* types of specialisation in BMHI, and
L at various *Levels* of education, corresponding to respective stages of career progression. There must be
T qualified *Teachers* to provide BMHI courses, which lead to
H recognised qualifications for biomedical and *Health* informatics positions.

In more detail, ‘HEALTH’ means:

H Practically all *professionals in health care* should, during their studies, be confronted with BMHI education: e.g. physicians, nurses, pharmacists, dentists, health care managers, health record administrators, and also health and medical informaticians who are graduates from specialised programs in BMHI. Computer scientists/informaticians and other scientists (e.g. engineers), who intend to work in the fields of medicine and health care, also need BMHI education.

E Various *education methodologies* are needed to provide the required theoretical knowledge, practical skills and mature attitudes. In addition to traditional classroom-based models, there are many different models of flexible, distance and supported open learning to be considered. The explosive growth of the Internet and World Wide Web, video-conferencing, document sharing and social networking platforms and applications are additionally having great impacts on all educational methodologies, and in particular will favour flexible and distance learning including both synchronous and asynchronous communication between instructors and students. Inter-university collaborations might also facilitate curricular choice.

A *Alternate routes to different types of specialisation* in BMHI will depend on career choice. The majority of health care professionals (e.g. physicians, nurses) need to know how to efficiently and responsibly use information and communication technology, but only a few will choose to have accredited specialisation in this field. They should, however, also be able to acquire an additional specialist qualification in BMHI as part of their chosen career development. BMHI specialisations may be different to suit the various types of health care professionals. Finally, it should also be possible to acquire specialist qualifications in BMHI via specific BMHI programs leading to accreditation at different levels, e.g. master or PhD.

L Every profession in health care even at an early stage needs some core BMHI knowledge. Different *levels of education*, respectively stages of career progression, (bachelor, master, and doctor) have different BMHI education needs according to experience, professional role and responsibility. A junior professional uses information differently compared to a senior professional. There are specialised BMHI university programs but BMHI instruction could also be integrated within other professional educational programs (medicine, nursing, informatics/computer science etc.). Thus educational components will vary in depth and breadth to suit specific student groups. Subsequent continuous education programs in BMHI also need to be available.

T The content and delivery of BMHI courses and programs must be of good quality. *Teachers* of BMHI courses must have adequate and specific competence in this field.

H There must be recognised qualifications in *BMHI* for positions in this field. Accreditation of educational content and competence in BMHI is required, to eventually have recognition on an international basis.
2.2 Structural Outline of the IMIA Recommendations

The IMIA recommendations centre on educational needs for health care professionals to acquire knowledge and skills in information processing and information and communication technology as it is needed and used in medicine and health care. The educational needs are described as a three-dimensional framework with dimensions ‘professional in health care’, ‘type of specialisation in BMHI’ and ‘stage of career progression’ (figure 1).

![Knowledge and skills needed in biomedical and health informatics](image)

**Figure 1: Structural outline of the IMIA recommendations on education in biomedical and health informatics (BMHI): knowledge and skills needed in BMHI by students in health care, to get a specialisation in BMHI and to reach a career progression lead to learning outcomes in programs, and have to be transformed into educational components with appropriate depth and breadth.**

Figure 1 points out that if one is studying a certain discipline (e.g. medicine to receive a bachelor degree), then the IMIA recommendations suggest that in their study, all these students should get a minimum of education in BMHI, so that they are able to efficiently use information and communication technology (IT users).

On the other side, candidates may want to prepare for careers as BMHI specialists. The study of BMHI for these specialists is somewhat different. Here we have to interpret figure 1 in the sense, that learning outcomes (also being given in table 2 and further explained in sections 4 and 5) are defined to get a bachelor, master or doctor degree in BMHI. Per definition this predefined a BMHI specialist. There are obviously different ways to become a qualified BMHI specialist.
3 Recommendations for Learning Outcomes in Biomedical and Health Informatics

Interesting differences exist both within and between countries and programmes with regard to structures of curricula and expected learning outcomes. Several initiatives have been launched in the last years to define some standardized content of BMHI curricula, aiming at developing sample informatics curricula.

A clear trend in curriculum design is the integration of disciplines closely related to the core field of BMHI, such as biomedical engineering, medical information sciences, molecular biology or nanosciences. Those fields share knowledge, methods and tools with BMHI. Figure 2 highlights and describes the most important related disciplines. For further discussion of this issue, see e.g. also [78-80].

The learning outcomes, presented in section 3, will integrate those overlapping areas as optional elements, depending on the focus of the respective programme. This assures that graduates from BMHI programmes know at least the basics of those related disciplines, and to give various programmes the flexibility to focus on one or more of those overlapping areas, depending on the cultural, scientific and technical context of the institution.

Figure 2: Biomedical and health informatics and related fields. Overlapping areas: A – medical information science, B – medical chemoinformatics, C – clinical informatics, D – medical (translational) bioinformatics, E – public health informatics, F – medical nanoinformatics, G – medical imaging and devices.
For education in BMHI two kinds of major learning outcomes can be identified. They specify the

1. **Learning outcomes for all health care professionals in their role as IT users:** Enabling health care professionals to efficiently and responsibly use information and knowledge processing methodology and information and communication technology. These learning outcomes need to be included in all undergraduate curricula, leading to a health care professional qualification. On the other hand there are:

2. **Learning outcomes for BMHI specialists:** Preparing graduates for careers in BMHI in academic, health care (e.g. hospital, primary care), government or industrial settings. These learning outcomes need to be included in all curricula, leading to a qualification as specialist in BMHI.

Obviously, between the specialisation of a health care professional as IT users and a health care professional as a BMHI specialist, *various levels concerning depth and breadth of learning outcomes exist*. Some programmes may focus on either health care professionals or on health informatics specialists. Other programmes may focus on a kind of intermediary level, where students are educated to communicate with physicians and nurses as IT users on one side and health informatics specialists on the other side.

*The learning outcomes define the levels of knowledge and skills needed.* The desired outcomes determine the educational components either in courses/course tracks in BMHI as part of educational programs or as dedicated programs in BMHI.

Table 2 contains the list of learning outcomes, recommended by IMIA. The knowledge and skill levels are classified into four domain areas:

1. BMHI core knowledge and skills,
2. medicine, health and biosciences, health system organisation,
3. informatics/computer science, mathematics, biometry,
4. optional modules in BMHI and from related fields.

In order to achieve the learning outcomes mentioned above, their educational components should be considered for inclusion into the respective educational programs. Knowledge and skills, which are described, in part four are optional and are recommended e.g. if the research profile of the university/school offering a program includes these fields and if it fits well into the program.
Table 2-1: Recommended and optional learning outcomes in terms of levels of knowledge and skills for professionals in health care either in their role as IT users or as BMHI specialists. Additional recommendations, specific for a certain educational program, will be added in sections 4 and 5. Recommended level of knowledge and skills:

+ = introductory. ++ = intermediate. +++ = advanced.

<table>
<thead>
<tr>
<th>Knowledge/Skill - Domain</th>
<th>- Level</th>
<th>IT user</th>
<th>BMHI specialist</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Evolution of informatics as a discipline and as a profession</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>1.2 Need for systematic information processing in health care, benefits and constraints of information technology in health care</td>
<td>++</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>1.3 Efficient and responsible use of information processing tools, to support health care professionals’ practice and their decision making</td>
<td>++</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>1.4 Use of personal application software for documentation, personal communication including Internet access, for publication and basic statistics</td>
<td>++</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>1.5 Information literacy: library classification and systematic health related terminologies and their coding, literature retrieval methods, research methods and research paradigms</td>
<td>+</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>1.6 Characteristics, functionalities and examples of information systems in health care (e.g. clinical information systems, primary care information systems, etc.)</td>
<td>+</td>
<td>+++</td>
<td></td>
</tr>
<tr>
<td>1.7 Architectures of information systems in health care; approaches and standards for communication and cooperation and for interfacing and integration of component, architectural paradigms (e.g. service-oriented architectures)</td>
<td></td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>1.8 Management of information systems in health care (health information management, strategic and tactic information management, IT governance, IT service management, legal and regulatory issues)</td>
<td>+</td>
<td>+++</td>
<td></td>
</tr>
<tr>
<td>1.9 Characteristics, functionalities and examples of information systems to support patients and the public (e.g. patient-oriented information system architectures and applications, personal health records, sensor-enhanced information systems)</td>
<td></td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>1.10 Methods and approaches to regional networking and shared care (eHealth, health telematics applications and inter-organizational information exchange)</td>
<td>+</td>
<td>++</td>
<td></td>
</tr>
</tbody>
</table>

*to be continued*
### Table 2-2: Recommended and optional learning outcomes in terms of levels of knowledge and skills for professionals in health care either in their role as IT users or as BMHI specialists.

<table>
<thead>
<tr>
<th>Knowledge/Skill - Domain</th>
<th>- Level</th>
<th>IT user</th>
<th>BMHI specialist</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>(1) Biomedical and Health Informatics Core Knowledge and Skills (continued)</em></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.11 Appropriate documentation and health data management principles including ability to use health and medical coding systems, construction of health and medical coding systems</td>
<td>+</td>
<td>+++</td>
<td></td>
</tr>
<tr>
<td>1.12 Structure, design and analysis principles of the health record including notions of data quality, minimum data sets, architecture and general applications of the electronic patient record/electronic health record</td>
<td>+</td>
<td>+++</td>
<td></td>
</tr>
<tr>
<td>1.13 Socio-organizational and socio-technical issues, including workflow/process modelling and reorganization</td>
<td>+</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>1.14 Principles of data representation and data analysis using primary and secondary data sources, principles of data mining, data warehouses, knowledge management</td>
<td>++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.15 Biomedical modelling and simulation</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.16 Ethical and security issues including accountability of health care providers and managers and BMHI specialists and the confidentiality, privacy and security of patient data</td>
<td>+</td>
<td>++</td>
<td></td>
</tr>
<tr>
<td>1.17 Nomenclatures, vocabularies, terminologies, ontologies and taxonomies in BMHI</td>
<td>++</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.18 Informatics methods and tools to support education (incl. flexible and distance learning), use of relevant educational technologies, incl. Internet and World Wide Web</td>
<td>+</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.19 Evaluation and assessment of information systems, including study design, selection and triangulation of (quantitative and qualitative) methods, outcome and impact evaluation, economic evaluation, unintended consequences, systematic reviews and meta-analysis, evidence-based health informatics</td>
<td>++</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*to be continued*
Table 2-3: Recommended and optional learning outcomes in terms of levels of knowledge and skills for professionals in health care either in their role as IT users or as BMHI specialists.

<table>
<thead>
<tr>
<th>Knowledge/Skill - Domain</th>
<th>IT user</th>
<th>BMHI specialist</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2) Medicine, Health and Biosciences, Health System Organisation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.1 Fundamentals of <strong>human functioning</strong> and biosciences (anatomy, physiology, microbiology, genomics, and clinical disciplines such as internal medicine, surgery etc.)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2.2 Fundamentals of <strong>what constitutes health</strong>, from physiological, sociological, psychological, nutritional, emotional, environmental, cultural, spiritual perspectives and its assessment</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2.3 Principles of <strong>medical decision making</strong> and diagnostic and therapeutic strategies</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>2.4 <strong>Organisation of health institutions</strong> and of the overall health system, interorganizational aspects, shared care</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>2.5 <strong>Policy and regulatory frameworks</strong> for information handling in health care</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>2.6 Principles of <strong>evidence-based medicine</strong> and evidence-based nursing</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>2.7 <strong>Health administration, health economics</strong>, health quality management and resource management and patient safety initiatives, public health services and outcome measurement</td>
<td></td>
<td>++</td>
</tr>
</tbody>
</table>

*to be continued*
Table 2-4: Recommended and optional learning outcomes in terms of levels of knowledge and skills for professionals in health care either in their role as IT users or as BMHI specialists.

<table>
<thead>
<tr>
<th>Knowledge/Skill – Domain</th>
<th>IT user</th>
<th>BMHI specialist</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(3) Informatics/Computer Science, Mathematics, Biometry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1 Basic informatics terminology like data, information, knowledge, hardware, software, computer, networks, information systems, information systems management</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>3.2 Ability to use personal computers, text processing and spread sheet software, easy-to-use database management systems</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>3.3 Ability to communicate electronically, including electronic data exchange, with other health care professionals, internet/intranet use</td>
<td>++</td>
<td>+++</td>
</tr>
<tr>
<td>3.4 Methods of practical informatics/computer science, especially on programming languages, software engineering, data structures, database management systems, information and system modelling tools, information systems theory and practice, knowledge engineering, (concept) representation and acquisition, software architectures</td>
<td></td>
<td>+++</td>
</tr>
<tr>
<td>3.5 Methods of theoretical informatics/computer science, e.g. complexity theory, encryption/security</td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>3.6 Methods of technical informatics/computer science, e.g. network architectures and topologies, telecommunications, wireless technology, virtual reality, multimedia</td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>3.7 Methods of interfacing and integration of information system components in health care, interfacing standards, dealing with multiple patient identifiers</td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>3.8 Handling of the information system life cycle: analysis, requirement specification, implementation and/or selection of information systems, risk management, user training</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>3.9 Methods of project management and change management (i.e. project planning, resource management, team management, conflict management, collaboration and motivation, change theories, change strategies)</td>
<td>+</td>
<td>+++</td>
</tr>
</tbody>
</table>

to be continued
Table 2-5: Recommended and optional learning outcomes in terms of levels of knowledge and skills for professionals in health care either in their role as IT users or as BMHI specialists.

<table>
<thead>
<tr>
<th>Knowledge/Skill – Domain</th>
<th>IT user</th>
<th>BMHI specialist</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(3) Informatics/Computer Science, Mathematics, Biometry (continued)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.10 Mathematics: algebra, analysis, logic, numerical mathematics, probability theory and statistics, cryptography</td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>3.11 Biometry and epidemiology, including study design</td>
<td></td>
<td>++</td>
</tr>
<tr>
<td>3.12 Methods for decision support and their application to patient management, acquisition, representation and engineering of medical knowledge; construction and use of clinical pathways and guidelines</td>
<td>+</td>
<td>+++</td>
</tr>
<tr>
<td>3.13 Basic concepts and applications of ubiquitous computing (e.g. pervasive, sensor-based and ambient technologies in health care, health enabling technologies, ubiquitous health systems and ambient assisted-living)</td>
<td></td>
<td>+</td>
</tr>
<tr>
<td>3.14 Usability engineering, human-computer interaction, usability evaluation, cognitive aspects of information processing</td>
<td></td>
<td>++</td>
</tr>
</tbody>
</table>

_to be continued_
Table 2-6: Recommended and optional learning outcomes in terms of levels of knowledge and skills for professionals in health care either in their role as IT users or as BMHI specialists.

<table>
<thead>
<tr>
<th>Knowledge/Skill – Domain</th>
<th>IT user</th>
<th>- Level BMHI specialist</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(4) Optional Modules in BHMI and from Related Fields</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.1 Biomedical imaging and signal processing</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>4.2 Health-enabling technologies, ubiquitous health systems and ambient-assisted living</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>4.3 Medical bioinformatics and computational biology</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>4.4 Medical chemoinformatics</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>4.5 Medical information sciences</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>4.6 Medical nanoinformatics</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>4.7 Medical robotics</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>4.8 Public health informatics</td>
<td>+</td>
<td>–</td>
</tr>
</tbody>
</table>

All health care professional graduates should, in their role as IT users, have the levels of knowledge and skills mentioned for IT users. Analogously, those professionals in health care, being BMHI specialists, should have the levels of knowledge and skills specified for them.

The levels of knowledge and skills mentioned may particularly work well for developed, industrialised countries, with high levels of access to, and use of, information technology, and which have highly developed health care infrastructures. Developing countries may at the beginning have the need to adapt them with regard to the level of technology. The principles of BMHI, however, can still be taught, developed and applied in the absence of high levels of information and communication technology.

Recommendations, either specific for certain courses or course tracks in BMHI as part of educational programs or specific for dedicated educational programs in BMHI are mentioned in sections 4 and 5.
4 Recommendations for Courses/Course Tracks in Biomedical and Health Informatics as Part of Educational Programs

4.1 General Remarks

Educational course components in BMHI should be tailored to the student's advancement and where possible be made relevant for and used to support a given stage of student progression. For example, teaching about the patient record for students of medicine should be introduced after the student has gained some clinical experience, but not too late so that students can benefit from this knowledge in the latter stages of their clinical training.

Due to the afore-mentioned large variety, there exist different perspectives for BMHI education. For BMHI specialists we especially can distinguish between a more informatics-based and a more health care-based approach to BMHI education, with a variety of combinations in-between. It is important to recognise the need for teamwork as all health informatics projects require input from more than one person each with their own unique skill set, so that the team as whole is able to address all project aspects in a cohesive and coordinated manner.

The objective of an informatics-based approach to BMHI is to focus on the machine processing of data, information and knowledge in health care and medicine with a strong emphasis on the need for advanced knowledge and skills of BMHI, of workflow, people and organisational aspects, of mathematics, as well as of theoretical, practical and technical informatics/computer science, especially semantic interoperability, ontology based software engineering and its relationship with effective and safe data, information and knowledge processing and representation. Health care problems, however, can be treated cooperatively with physicians and other health care professionals. In such an approach to BMHI education, knowledge and skills of informatics/computer science predominate, but must be applied with a sound knowledge of the business of providing healthcare services.

The objective of a health care-based approach to BMHI is again to focus on the machine processing of data, information and knowledge in health care and medicine requiring, apart from knowledge in BMHI, also knowledge in medicine or other health sciences to such an extent that can only be imparted within the scope of a medical or health science education. In such an approach to BMHI education, clinical knowledge and skills predominate but these must be applied within the BMHI context.

The recommendations, given in section 4.2 and 4.3 for BMHI specialists are for health care-based approaches to BMHI. The recommendations in sections 4.4 and 5.2 are oriented towards an informatics-based approach. With respect to educational progression, especially for a bachelor, master, and doctoral degree, the general distinctions in depth and breadth should be considered as mentioned in section 5.

For specifying a student’s workload, the European Credit Transfer and Accumulation System (ECTS) is used [81]. In ECTS a full academic year’s student workload corresponds to 60 ECTS credits.

4.2 Recommendations for Biomedical and Health Informatics Courses as Part of Medical, Nursing, Health Care Management, Dentistry, Pharmacy, and Public Health Programs

Courses/Course Tracks for IT Users

In order to achieve the levels of knowledge and skills in BMHI as recommended in section 3 for IT users, the total student workload for educational components in BMHI should comprise at least 4 ECTS credits. 4 ECTS credits can correspond, e.g., to approx. 40 hours of lectures, exercises and practical training at universities. A course, charged with 4 ECTS credits, may e.g. consist of a 3 hours/week lecture, given in one semester with 14 weeks of lecturing.
Specific examples from the work of the respective health professionals should be used. Emphasis should particularly be given to practical and simulated training.

The additional recommendations of this section 3 may also apply to the programs of other professions in health care such as medical laboratory technicians, medical librarians, radiology technicians, dieticians, occupational therapists etc. or for the programs allied health/clinical researchers studied. These people also need to know about the potentials and the risks of information processing in health care and be able to efficiently use methods and tools of information processing and information and communication technology.

**Courses/Course Tracks for BMHI Specialists**

In order to achieve the levels of knowledge and skills in BMHI, as recommended in section 3 for specialists, the student workload associated with these educational components in BMHI should be at least 60 ECTS credits, i.e. one year of full time studies. This is similar to dedicated master programs in BMHI.

In addition to the ‘core’ knowledge and skills obtained in each program, the relative amount of student workload for the three knowledge and skills areas inside the health/medical informatics course track should approximately be as indicated in table 3.

Table 3: Recommended student workload in ECTS credits for the three knowledge and skill areas of health/medical informatics course tracks inside programs of medicine and other health sciences.

<table>
<thead>
<tr>
<th>Knowledge/Skill Area</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) BMHI core knowledge and skills</td>
<td>Medicine, Nursing, Health Care Management, Dentistry, Pharmacy, Public Health</td>
</tr>
<tr>
<td>(2) medicine, health and biosciences, health system organisation</td>
<td>5</td>
</tr>
<tr>
<td>(3) informatics/computer science, mathematics, biometry</td>
<td>15</td>
</tr>
<tr>
<td>Σ</td>
<td>60</td>
</tr>
</tbody>
</table>

For all health care professionals domain area (2) should focus on health system organisation, area (3) on practical informatics and project management. For nurses it should be possible that specialisation can be included in a post registration nursing curriculum. For health care managers, knowledge and practical skills of information systems architectures, incl. characteristics required to achieve semantic interoperability and information systems/network management should particularly comprise work and information flow supporting enterprise functions for administration, controlling, quality management and executive decision making.

4.3 **Recommendations for Biomedical and Health Informatics Courses as Part of Health Record Administration Programs**

Within the past decade the discipline of health record administration (also denoted as health information management) has often enhanced its scope from document handling to managing health care information. Also the scope of practice has changed considerably.

For educating health record administrators, two different levels of education are recommended:
- A first level should cover introductory concepts and principles and assumes an introductory skill level. Students at this level take e.g. a two- or three year course of study at a college level resulting (e.g. in the U.S.) in an associate’s degree. The focus for these students needs to be on data, meta-data, coding rules, classification systems and their relationship with health informatics.

- At a second level a deeper understanding of knowledge and more advanced terminology skills, problem solving and critical thinking skills in more depth is assumed. Students at this level take e.g. a three- or four-year course of study resulting in a bachelor degree where health information management skills and knowledge is integrated with more advanced informatics knowledge and skills. Further studies may follow.

**Courses/Course Tracks for IT Users**

Health record administration students at the mentioned first level can be regarded as IT users. The recommendations on levels of knowledge and skills are the same as for IT users, mentioned in section 4.2. Particular emphasis should be given to information literacy, health terminology, coding and classification systems, the electronic health record, and evaluation methodology. There should be introductory knowledge and skills in the knowledge/skill-domain medicine, health and biosciences, health systems organisation.

**Courses/Course Tracks for Biomedical and Health Informatics Specialists**

Students of health record administration programs, respectively health information management programs, which lead to bachelor and master degrees should have the knowledge and skills of BMHI specialists, as mentioned in section 4.2. Again, special emphasis should be given to information literacy, meta-data, health terminology, coding and classification systems, the electronic health record, and evaluation methodology and to the relationship between these concepts and the use of various informatics technologies.

### 4.4 Recommendations for Biomedical and Health Informatics Courses as Part of Informatics/Computer Science Programs

**Courses/Course Tracks for Biomedical and Health Informatics Specialists**

In order to achieve the levels of knowledge and skills in BMHI, recommended in section 3 for specialists, the length of studies for educational components in BMHI should be at least 60 ECTS credits, i.e. one year of full time studies.

In addition to the ‘core’ knowledge and skills of informatics/computer science, the relative amount of student workload for the three knowledge and skills areas inside the health/medical informatics course track should approximately be as indicated in table 4.

*Table 4: Recommended student workload in ECTS credits for the three knowledge and skill areas of a health/medical informatics course track inside informatics/computer science programs.*

<table>
<thead>
<tr>
<th>Knowledge/Skill Area</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMHI core knowledge and skills</td>
<td>Informatics/Computer Science</td>
</tr>
<tr>
<td>medicine, health and biosciences, health system organisation</td>
<td>15</td>
</tr>
<tr>
<td>informatics/computer science, mathematics, biometry</td>
<td>5</td>
</tr>
<tr>
<td>Σ</td>
<td>60</td>
</tr>
</tbody>
</table>

The student workload in (3) comprises knowledge and skills in biometry, semantic interoperability and evaluation methods. Applying methods and tools of informatics in health care institu-
tions, and for concrete problems in diagnosis, therapy, nursing and health care management should be emphasised. It is essential to include ontology based software engineering and the need to separate knowledge from system configuration, as these concepts are fundamental to achieving semantic interoperability and safe clinical decision support systems. This assists informatics or computer science students to gain more knowledge about the health care environment. Health information systems management should include the development and implementation of software and hardware components of health information systems. In medical signal and image processing technical and informatics aspects should particularly be considered.

5 Recommendations for Dedicated Educational Programs in Biomedical and Health Informatics

The aim of all dedicated programs in BMHI is to prepare graduates for careers in academic/research settings, health care facilities or organizations, governmental or international public health entities or industrial settings.

5.1 Recommendations for Bachelor Programs in Biomedical and Health Informatics

The curriculum of a program leading to a bachelor degree in BMHI should be application-related aiming to prepare students for a professional career in the field. In addition, the curriculum should offer the background and theoretical foundation necessary to pursue advanced graduate studies in this or related fields.

The objective of this undergraduate education is to equip students with specialised knowledge in the field of BMHI and the skills to apply the acquired knowledge in a variety of practical situations. The intention is to provide a deep understanding of the state-of-the-art of the discipline and the ability to translate expertise gained in the field into practical application of tools and concepts. Compared to the comprehensive formal methodological foundation of a master program, it is the practice-oriented application that predominates the undergraduate curriculum. Given the diversity of the discipline, students at the bachelor program level need to understand the breadth of the field and become familiar with the spectrum of BMHI (capturing all sub-domains such as bioinformatics, clinical informatics, public health informatics etc.). The challenge herein is to provide knowledge and skills that students can apply in practice while recognizing that areas of interest could be explored further and in more-depth at the graduate educational level.

In order to achieve the levels of knowledge and skills recommended in section 3, and to achieve a broad depth and breadth of all educational components, the length of study for the instructional component of a bachelor program in BMHI should be at least three years. This corresponds to a student workload of 180 ECTS credits.

For an informatics-based approach to BMHI, the relative amount of student workload for the four knowledge and skills areas for the bachelor program should be approximately as indicated in table 5. This composition can be varied from very strong technical IT skill acquisition to less IT skill and a stronger health application focus, depending on the desired learning outcomes.
Table 5: Recommended student workload in ECTS credits for the four knowledge and skill areas of a biomedical and health informatics bachelor program.

<table>
<thead>
<tr>
<th>Knowledge/Skill Area</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BMHI (bachelor)</td>
</tr>
<tr>
<td>(1) BMHI core knowledge and skills</td>
<td>50</td>
</tr>
<tr>
<td>(2) medicine, health and biosciences, health system organisation</td>
<td>20</td>
</tr>
<tr>
<td>(3) informatics/computer science, mathematics, biometry</td>
<td>110</td>
</tr>
<tr>
<td>Σ</td>
<td>180</td>
</tr>
</tbody>
</table>

5.2 Recommendations for Master and Doctoral Programs in Biomedical and Health Informatics

For programs leading to a master or doctoral degree, it is the comprehensive formal methodological foundation for BMHI that dominates the instructional component of the program.

The objective is to provide scientific education that captures the theoretical foundations of the field, provides specialized knowledge and equips students with both practical skills and analytical approaches that will allow them to further the knowledge base of the discipline. Graduates will be able to master both the practical methods and tools and the leadership of independent research.

Unlike undergraduate bachelor programs, these graduate and post-graduate programs emphasize a formal penetration into the knowledge and foundation of the discipline and promote methodological expertise and independent analysis. Graduates are expected to contribute to the field and lead its scientific advancement.

In order to achieve the levels of knowledge and skills in BMHI as recommended in section 3, and in order to achieve the desired broad depth and breadth of the educational components, the length of study should be at least one year full time for a master degree, corresponding to at least 60 ECTS credits. Two years of study, corresponding to 120 ECTS credits, should be preferred. Ph.D. studies or Ph.D. work should usually last three to four years.

The relative amount of study time for the three knowledge and skills areas for the master program should approximately be as indicated in table 6.

Table 6: Recommended student workload in ECTS credits for the four knowledge and skill areas of a two year (one year) biomedical and health informatics master program.

<table>
<thead>
<tr>
<th>Knowledge/Skill Area</th>
<th>Program</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>BMHI (master)</td>
</tr>
<tr>
<td>(1) BMHI core knowledge and skills</td>
<td>80 (40)</td>
</tr>
<tr>
<td>(2) medicine, health and biosciences, health system organisation</td>
<td>20 (10)</td>
</tr>
<tr>
<td>(3) informatics/computer science, mathematics, biometry</td>
<td>40 (20)</td>
</tr>
<tr>
<td>Σ</td>
<td>120 (60)</td>
</tr>
</tbody>
</table>

It is expected that master students have successfully finished either (a) a bachelor program in BMHI, (b) a bachelor or master program in medicine, biology, public health, health administration or another health science, or (c) in computer science/ information science. For cases (b) and (c) additional complementary courses (for b in informatics/computer science and for c in health and biosciences, health systems) should be offered.

For programs leading to a doctoral degree, independent comprehensive research should be carried out by the student in addition to the instructional requirements already mentioned.
Knowledge and skills should also have additional depth and breadth and students may choose to gain additional insight into elective fields that are at the core of their research.

6 Recommendations for Continuing Education
6.1 Continuing Education in Biomedical and Health Informatics
A certificate of ‘Health Informatics’, ‘Medical Informatics’ and/or ‘Biomedical Informatics’ should be offered in recognition of having acquired sufficient competence in BMHI from an academic, educational and/or practical perspective relative to specific tasks or roles within the health industry. Furthermore, for physicians, who usually have well-established forms of continuing education, there should be offered the possibility of receiving, in addition to their medical degree, the supplementary qualification of ‘Health Informatics’, ‘Medical Informatics’ and/or ‘Biomedical Informatics’. This additional qualification can be issued by any national medical or health professional association or University. The same holds true for nurses, for whom various forms of continuing education are very well-established in many countries.

In order to offer courses in BMHI for continuing education, it is recommended that specific entities are established to provide such courses. These entities might be inside universities or, as academies of health/medical informatics established by any national association in BMHI or provided by an independent private entity provided that in all cases the people responsible for course curriculum, content and educational delivery are suitably qualified.

6.2 Life-Long Learning
Working in the field of BMHI and even using information and communication technology requires life long learning. Therefore opportunities for continuing education should be offered for BMHI specialists as well as IT users of the various health professions. The ability of ‘learning to learn’ will become of particular importance.

7 Other Recommendations
7.1 How to Commence with Biomedical and Health Informatics Education
BMHI affects all health care professionals. To commence education in this field, IMIA recommends that education in BMHI for all types of health care professionals, including the different types of specialization and levels of education is considered. The first step is to consider the level of practice of the individual. As noted in table 7, work in informatics depends upon whether someone has a technical/informatics or clinical background. Within informatics, one may practice more at the applied level in operational settings or may be an academic who teaches and/or performs research. For countries without formal informatics educational programs, additional steps are necessary. In this instance, teachers have to be educated (‘teach the teachers’) first, or, e.g., retiring BMHI faculty in developed countries could also perform this service. Next, courseware has to be prepared (or adapted) and institutes for health informatics or medical informatics must be established within universities, usually inside medical or health sciences faculties.
Table 7 - Categories of informatics practice, adapted from [53].

<table>
<thead>
<tr>
<th>Level of Practice</th>
<th>Type of Work</th>
<th>Example Job Titles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Academic</td>
<td>Individual who does research and/or teaching in an academic center</td>
<td>Professor, Scientist or Researcher</td>
</tr>
<tr>
<td>Applied</td>
<td>Individual who works in an operational informatics setting for a majority of his or her working time</td>
<td>Chief Information Officer, Chief Medical or Nursing Information Officer, Project Manager, Developer, Trainer</td>
</tr>
<tr>
<td>Clinical</td>
<td>Clinicians and others who apply informatics in their work</td>
<td>Physician, Nurse, Health administrator</td>
</tr>
</tbody>
</table>
recommendations and the assessment of the quality of the program will be examined by a committee usually consisting of four members of IMIA’s Working Group on Health and Medical Informatics Education or other persons, experienced in BMHI education, and will be approved by the IMIA President and the Chairperson of IMIA’s Working Group on education.

After approval, a written certificate, signed by the IMIA President, the Chairperson of IMIA’s Working Group on Health and Medical Informatics Education, will be given to the respective organisation.

Requests should be submitted to the Executive Director of IMIA.

8.2 International Programs, International Exchange of Students and Teachers
IMIA encourages and recommends international activities in educating BMHI specialists. IMIA also recommends the international exchange of students and of teachers in this field. It encourages the establishment of international programs to support this and to exchange courseware. Programs should be built up in a modular way, and international credit transfer systems such as the European Credit Transfer and Accumulation System (ECTS, [81]) should be used in the respective national programs to support these international perspectives.

9 Information Exchange on Programs and Courses in Biomedical and Health Informatics supported by IMIA
IMIA’s Working Group on Health and Medical Informatics Education is a group devoted to BMHI education. Its activities include to disseminate and exchange information on BMHI programs and courses and to support BMHI courses and exchange of students and teachers. The Working Group intends to advance the knowledge of (1) how informatics is taught in the education of health care professionals around the world, (2) how in particular BMHI is taught to students of computer science/informatics, and (3) how it is taught within dedicated curricula in BMHI. Working Group members contribute to various Web-based catalogues of BMHI programs, including those, mentioned in section 1.2.

In addition, IMIA’s Working Group on Health and Medical Informatics Education operates a mailing list to facilitate communication between all persons interested in BMHI education worldwide. For subscription, instructions on the Working Group’s web site should be followed. IMIA encourages the development and sharing of courseware of high quality for courses in BMHI. This will help to further establish courses in this field. It also encourages the use of its IMIA Working Group on education’s web site and list server for the dissemination of information about such courseware.

10 Concluding Remarks
These recommendations provide a framework for individual curriculum development. Individual countries may wish to develop more detailed or better defined curricula guidelines to suit their specific needs and educational system. This could include specific minimum level competencies required for each level and knowledge/skill domain. Such national efforts are expected to inform future reviews of these guidelines.

IMIA, understanding its role as a leader in the scientific progress of the expanding field of BMHI, is also developing as part of its strategic plan [73] a knowledge base [83], which can be of real significance for educators in the field as well as to healthcare professionals.

IMIA’s Working Group on Health and Medical Informatics Education may in the near future develop teaching credentialing criteria to serve as a guide for teachers wishing to participate in BMHI education.
Acknowledgements

These revised recommendations are the result of the task force team established by the IMIA Working Group on Health and Medical Informatics Education in 2006 ([77]). The draft document is being communicated and discussed by the working group members during its Working Conference 2008 at Buenos Aires. The recommendations are subject to final approval by the IMIA General Assembly.

Significant contributions to the recommendations came from Elske Ammenwerth (Austria), George Demiris (USA), Arie Hasman (The Netherlands), Reinhold Haux (Germany), William Hersh (USA), Evelyn Hovenga (Australia), K.C. Lun (Singapore), John Mantas (Greece), Heimar Marin (Brazil), Fernando Martin-Sanchez (Spain), Graham Wright (UK). The team would like to thank Professor Dr. Reinhold Haux, President of IMIA for his continued support in developing the revised version of the Educational Recommendations. We also thank Dr. Peter Murray, Executive Director of IMIA for monitoring and supporting the team’s efforts.

We would like to thank all contributors of the original recommendations for their pioneering work as mentioned in [1].

The recommendations have been edited by John Mantas (co-chairman of IMIA on Health and Medical Informatics Education) chairing the task force team.

References


Health Informatics Competency Profiles for the NHS. London: National Health Service Information Authority; 2001.


Ivanitskaya L, O’Boyle I, Casey AM. Health information literacy and competencies of information age students: results from the interactive online Research Readiness Self-Assessment (RRSA). J Med Internet Res. 2006; 8: e6.


Address of Correspondence:
Professor John Mantas
Chair of Task Force
Email: jmantas@cc.uoa.gr

Members of the Task Force:
John Mantas (Chair), Elske Ammewerth, George Demiris, Arie Hasman, Reinhold Haux, William Hersh, Evelyn Hovenga, KC Lun, Heimar Marin, Fernando Martin-Sanchez, Graham Wright

International Medical Informatics Association (IMIA)
Working Group on Health and Medical Informatics Education

http://www.imia.org
Chair:
Professor William Hersh
Chairman, Department of Medical Informatics and Clinical Epidemiology
Oregon health and Science University
USA
Email: hersh@ohsu.edu
Co-Chair:
Professor John Mantas
Director, Health Informatics Laboratory
University of Athens
Greece
Email: jmantas@cc.uoa.gr
The Development of the IMIA Knowledge Base

Background: The discipline of health or medical informatics is relatively new in that the literature has existed for only 40 years. The British Computer Society (BCS) health group was of the opinion that work should be undertaken to explore the scope of medical or health informatics. Once the mapping work was completed the International Medical Informatics Association (IMIA) expressed the wish to develop it further to define the knowledge base of the discipline and produce a comprehensive internationally applicable framework. This article will also highlight the move from the expert opinion of a small group to the analysis of publications to generalise and refine the initial findings, and illustrate the importance of triangulation.

Objectives: The aim of the project was to explore the theoretical constructs underpinning the discipline of health informatics and produce a cognitive map of the existing understanding of the discipline and develop the knowledge base of health informatics for the IMIA and the BCS.

Method: The five-phase project, described in this article, undertaken to define the discipline of health informatics used four forms of triangulation.

Results: The output from the project is a framework giving the 14 major headings (Subjects) and 245 elements, which together describe the current perception of the discipline of health informatics.

Conclusion: This article describes how each phase of the project was strengthened, through using triangulation within and between the different phases. This was done to ensure that the investigators could be confident in the confirmation and completeness of data, and assured of the validity and reliability of the final output of the ‘IMIA Knowledge Base’ that was endorsed by the IMIA Board in November 2009.

Background
The author undertook this research over a four-year period with a number of collaborators in five discrete phases, which utilised quantitative and qualitative approaches. The discipline of health or medical informatics is relatively new in that the literature has existed for only 40 years. The British Computer Society (BCS) health group was of the opinion that work should be undertaken to explore the scope of medical or health informatics. A qualitative approach was used to gather expert opinion and construct a cognitive map of the discipline of health informatics. Once the mapping work was completed the International Medical Informatics Association (IMIA) expressed a wish to develop it further to define the knowledge base of the discipline and produce a comprehensive internationally applicable framework. Various data extraction methods were then used to identify the most commonly used keywords in the health informatics published literature followed by a consensus method to produce a final framework and knowledge base. This mixed method approach was adopted as a pragmatic means to address the development of what the discipline considered the current knowledge base and thus a reflection of the thoughts and publications of the discipline. The work was overseen by an International Research Advisory Board and refereed by Professor Lorenzi on behalf of the IMIA Board and General Assembly.

Research problem
The discipline of health informatics had not been formally defined and many definitions of the discipline have emerged in the literature. Not only was there a lack of agreed definition in that the discipline was variously called: health informatics, medical informatics, clinical informatics and latterly bioinformatics, but the scope of the discipline had not been adequately defined. Some of the consequences included misunderstandings regarding standards and use of terminology, lack of consistency within educational curriculum and a lack of a framework for defining skills and workforce requirements.
Objectives

The aim of the project was to explore the theoretical constructs underpinning the discipline of health informatics and produce a cognitive map (Eden & Ackermann 2004) of the existing understanding of the discipline.

Subsequent aims of the project were to develop the knowledge base of health informatics, which was seen as central to the IMIA strategy (Murray 2008; Lorenzi 2007), and to undertake the task of exploring the current perceptions of the Health Informatics community as to the scope of the discipline.

Method

The project’s international advisory board of health informatics experts provided advice on the methods that were used and facilitated access to source materials. The mixed methods used in the project were:

- a consensus conference using a cognitive mapping exercise
- workshops to verify international interpretation
- extraction of keywords from the entire published index papers on health informatics using computer software packages and techniques
- workshop to examine keywords and exclude terms
- voting in of keywords by international volunteers using a voting system based in an Excel spreadsheet.

The aim was to obtain different perspectives (data) on the issue of mapping the discipline of health informatics with the belief that the analysis would provide confidence and confirmation that the data was complete and the final outcomes from all the phases of the project were not just artefacts of one particular method of data collection or analysis. This process of data gathering and systematic analysis reflects the principles of grounded theory where the researcher begins with an area of study and allows the theory to emerge from the data. In this project the area of study was the discipline of health informatics and the knowledge base was derived from data systematically gathered and analysed through the research processes undertaken (Strauss & Corbin 1998).

Wolf (2010) in a recent article says the final consideration in using a mixed method approach is ‘to consider thoroughly whether to engage in triangulation, and if doing so, to use tailor-made triangulation strategies fitted to the research questions and interests’.

The project undertaken to define the discipline of health informatics used all four forms of triangulation (Denzin 1970) and this article describes how each phase of the project triangulated with the other phases for confirmation and completeness of data, and validation and verification of the project outputs.

Triangulation is a strategy to ‘overcome the intrinsic bias that comes from single methods, single observer and single theory studies’ (Patton 1990). Its objective is the confirmation and completeness of data through cross checking data from several sources to seek out consistencies in the data (Begley 1996; O’Donoghue & Punch 2003). Many researchers also advocate triangulation as a means of resolving the quantitative and/or qualitative question through integrating the two approaches in one study and contributing to methodological rigor in order to validate the findings (Begley 1996; Cohen et al. 1994).

Denzin (1970) identified four forms of triangulation: data, investigator, theoretical and methodological. Data triangulation involves gathering data using different sampling strategies, so that segments of data are collected at various times, social situations and with different people. Investigator triangulation requires the use of more than one researcher in collecting and interpreting data. Using more than one theoretical position for data interpretation is called theoretical triangulation, whereas the most common form of triangulation, methodological triangulation, refers to the use of more than one method of data collection.

The ability to generalise findings to wider groups is one of the most common tests of validity for quantitative research. Triangulation is typically a strategy for improving the validity and reliability of research findings. Patton (2002:247) advocates the use of triangulation stating ‘triangulation strengthens a study by combining methods’. However, the idea that triangulation is simply the combination of different methods of investigation is a restricted one, and researchers need to increase their use of the other less frequently employed forms of triangulation. When using triangulation of methods, researchers should also reflect on whether the use of within-method triangulation would be advantageous to their project. Within-method triangulation involves using dissimilar aspects of the same method in one study; for example, a questionnaire might contain two different scales to measure emotions. Between-method triangulation involves using different research methods, for example a questionnaire and observation to collect data (Bryman 2003; Begley 1996). Sequential use of quantitative and qualitative methods may also be more effective for some projects rather than simultaneous use, which do not permit the development and refinement of the methodologies. The deliberate use of multiple data sources and methods to cross-check and validate findings, should pervade all projects and lead to the objective of confirmation. Triangulation should be chosen intentionally, and a description of its rationale, planning and implementation is essential in project reports to give authority to triangulation and the project outcomes (Begley 1996).

The project explored the theoretical constructs underpinning the discipline of health informatics. The early project work was situated within a theoretical educational framework. Bloom’s taxonomy affords a hierarchical scheme for categorising levels of complexity for objectives within educational settings (Bloom et al. 1984). It also overlays well against other academic levels, such as the progression from undergraduate to postgraduate levels (Furst 1981; Seddon 1978).
Bloom classified three domains of educational activity (Forehand 2005):

1. cognitive, describing knowledge and mental skills
2. affective, describing attitude, feelings and emotions
3. psychomotor, describing manual or physical skills.

Bloom identified six levels of educational objectives within the cognitive domain; from the lowest level, knowledge, through comprehension, application, analysis, synthesis and evaluation (Forehand 2005; Anderson et al. 2001). The first phase of the project was a mapping exercise that was based on these concepts.

The five phases of the project

Consensus conference

The 2005 Consensus conference was an intensive 24-hour workshop involving small group and plenary discussions, with participants and researchers in residence overnight. There were 24 invited participants drawn from a sample frame that had professions down one axis and organisations across the other. Organisations included health providers, family medicine, AssIST the 11 professions union, a number of United Kingdom (UK) health informatics groups and the IMIA; the world body for health and medical informatics. Most of the participants were from the UK whilst others came from Europe, Australia, South Africa and the USA. The conference aimed to capture all the elements of the discipline of health informatics and also the broad themes or subject areas into which these elements could be grouped. Within small groups, participants listed the main subject areas or themes from their own curricula, knowledge and experience. Then again within small groups they identified smaller elements of the subject areas. Finally, in a whole group activity, participants assigned each element to a subject area and a level from Bloom’s cognitive domain (Forehand 2005) where possible. The discussions resulted in a first data set comprising 221 elements, grouped into 13 themes that varied in size, with the smallest containing six elements and the largest 37. It was recognised that the largest theme, the ‘Toolkit’, which consists of IT skills and knowledge of IT processes, would likely be divided following further discussions, which subsequently happened during a 24 hour workshop in Belfast.

This consensus conference therefore used group activities as its research methodology to produce lists of elements grouped into themes. There were six researchers involved in facilitating the group and plenary activities, thus adding investigator triangulation to reduce a single researcher bias.

Workshops to verify international interpretation

Workshops were conducted in 2005 at two major health informatics conferences, the European Federation for Medical Informatics (MIE 2005) in Geneva and the American Medical Informatics Association (AMIA 2005) in Washington DC. They were short workshops and hence only explored the overall concept and the clinical informatics theme. Participants commented that there were no major issues with either the methodology used in phase one or the initial outcomes that should modify the direction of the project. These workshops therefore used investigator triangulation in that three of the original six investigators were present at the European workshop and two at the American workshop. The investigators were therefore a subset of the original research team employing both data triangulation in that data was gathered using a different sampling strategy, in other words, those international conference participants who chose to attend the workshop and methodological triangulation as the method here was not to create themes and elements but rather take that data and refine it through smaller and shorter validation workshops. Another workshop to validate the outputs was held in Belfast in 2007 after the January 2007 workshop in London highlighted the size of the toolkit. This meeting focused on refining the technical and computing themes previously developed in phase one and successfully affirmed the two technical themes ‘Computer Science for Health Informatics (ICT for Health) and Computer Systems Applications in Health (toolkit)’. Thus the large toolkit theme was logically separated and participants from computer science who had expressed concern that the single large theme did not reflect the computer science heading system were the main re-shapers of the two new themes. The resulting themes are:

- computer science for health informatics (ICT for Health)
- health and social care processes
- health (care) records
- health and social care industry
- health informatics standards
- knowledge domains and knowledge discovery
- legal and ethical
- people in organisations
- politics and policy
- technologies for health
- terminology, classification and grouping
- uses of clinical information
- using informatics to support clinical healthcare governance
- computer systems applications in health (Toolkit).

Extraction of keywords from the available published index papers on health informatics using computer software packages and techniques

Scopus is the largest abstract and citation database of research literature and quality web sources with smart tools to track, analyse and visualise research. A search of Scopus was undertaken using a set of keywords that are descriptors of Informatics. The project’s International Advisory Board agreed that the following key words should be used:

- health informatics
- medical informatics
- clinical informatics
- nursing informatics
- pharmacy informatics
- dental informatics.
The keywords within each article of the Reference Manager 11 database were exported as a series of tiles and then imported one at a time to an Excel spreadsheet as in the raw data format the total number of keywords extracted exceeded the number of rows available in an Excel worksheet.

After processing the data to count the number of occurrences of each keyword a master list of some 10 000 different keywords were identified, many of which were just English terms rather than health informatics specific, for example the authors place of abode and conference venue or country of study. The use of keywords in many publications depends on author choice and often reflects the wish to have the article seen as being in a particular theme or subject area. This is particularly so with those conferences that identify themes for the submission of papers.

This activity produced a new set of data and so triangulated with phase one of the project that also produced raw data. In itself it was preparatory work for the next two phases of the project.

Workshop to examine and exclude keywords

The next phase of the project refined and reduced the raw data by removing keywords not directly associated with health informatics. The list of keywords were given to information specialists, grouped into teams of three, at a workshop in London, UK in January 2007. The groups considered each word and excluded any that were not thought to be a health informatics term. Each word was tagged with the number of occurrences it had in the search. At the same time, keywords were assessed to see if they would fit into the existing cognitive map from the phase one workshop (Table 1).

The participants in the workshop reduced the list of 10 000 words to 444. The number of occurrences found in the literature search ranked each keyword on the spreadsheet and small focus groups excluded words unconnected with health informatics. The remaining 444 words appeared to be connected with areas of health informatics as opposed to being just English words and phrases used to describe the content of the papers.

Voting in of keywords by international volunteers using a voting system based in an Excel spreadsheet

An Excel spreadsheet was constructed with a list of the keywords from which participants were invited to chose (vote in) those that were associated with health informatics. The complete spreadsheet together with the instructions and examples of how to vote was emailed to the International Advisory Board, the IMIA working groups, the BCS specialist groups, and the European Federation for Medical Informatics (EFMI) working groups.

The voting was conducted with all of the keywords listed on the spreadsheet and a choice box next to each. The 444 keywords were divided into groups and each group was given a range of letters, A to G, H to M, N to R, and S to Z. Participants were asked to complete the group that contained the initial letter of their surname. Thus, as an example, Heather Carter voted on the columns A to G and Peter Ross voted on columns N to R.

Participants voted for about 100 words in their group. They were asked to vote for the keywords they thought were health informatics terms and classify them according to which phase one theme they thought the keyword belonged with by putting the number of the theme next to the word on the spreadsheet. Keywords that were consistently chosen were added to the original phase one cognitive map. These final two phases used methodological triangulation to refine the data and match it with the output of the first two phases: the phase one workshop and the international interpretation workshops.

Results

The final spreadsheet, which forms the basis of the IMIA Knowledge Base, was constructed from the outcomes of the original phase one workshop, the subsequent phase to check international interpretation, a review and content analysis of the literature, and a two-phase refinement following the extraction of keywords from the entire electronic published papers on health informatics. The different phases to the project in all took:

- data from different sources (people and electronic papers) – data triangulation
- used different research methodologies (workshops, electronic searches, electronic analysis, electronic voting) – methodological triangulation
- information from different investigators (one primary investigator, with five secondary investigators) – investigator triangulation
- from different theoretical positions (grounded theory, educational theory) – theoretical triangulation.

Conclusion

Through using mixed modes of research within and between the different phases of the project the investigators and subsequently the IMIA Board and General Assembly can be confident in the confirmation and completeness of the data through cross confirmation and validation from more than one data source.

<table>
<thead>
<tr>
<th>Theme</th>
<th>Uses of clinical information</th>
<th>Number of times tagged in search</th>
</tr>
</thead>
<tbody>
<tr>
<td>Element</td>
<td>Data analysis and statistical presentation</td>
<td></td>
</tr>
<tr>
<td>Keyword</td>
<td>Automatic data processing</td>
<td>78</td>
</tr>
<tr>
<td>Keyword</td>
<td>Analysis</td>
<td>635</td>
</tr>
</tbody>
</table>

Triangulation strengthened the project and ensured the validity and reliability of the project outcomes. The endorsement of the ‘IMIA Knowledge Base’ took place at the IMIA Board and General Assembly meetings of IMIA in July 2010. The final report and spreadsheet are available on the IMIA website in the section on IMIA Endorsed Documents (Wright 2009).

The initial outputs from phase one have been used in a number of ways including to help formulate an undergraduate biomedical informatics degree programme (Pritchard-Copley et al. 2006) and as a framework to classify scientific papers for the European Federation for Medical Informatics (EFMI) conferences.

Acknowledgements

The author wishes to acknowledge the invaluable contributions of the CHIRAD health informatics team and in particular Dr Helen Betts and Dr Peter Murray, the BCS, the IMIA, and the many colleagues in the international health and medical informatics communities who contributed to the various stages of the project.

Author competing interests

I declare that I have no financial or personal relationship(s) which may have inappropriately influenced me in writing this paper.

References


Funst, E., 1981, ‘Bloom’s taxonomy of educational objectives for the cognitive domain: Philosophical and educational issues’, Review of educational research 51, 441. doi:10.3302/00346543051004441


O’Donoghue, T. & Punch, K., 2003, Qualitative educational research in action: Doing and reflecting, Routledge, Abingdon, Oxon, UK.

Patton, M., 2002, Qualitative research and evaluation methods, Sage, Newbury Park, CA.


As seen in Fig. 2, the growth of WSN is rapid and fast. The projected sales of sensors are going to increase rapidly. Similarly Fig. 3 shows the world revenue forecast and growth rate for healthcare, medical and biometrics markets. We can see that sensor networks have a great future ahead with tremendous growth rate.

Sensor networks applications in healthcare have potential for large impacts [18]. These can be realized through real-time, continuous vital monitoring to give immediate alerts of changes in patient status. The data can also be relayed to the hospital or correlate with patient records and so on. Home monitoring applications for chronic and elderly patients which can be used to collect periodic or continuous data and be uploaded to a physician and can allow long-term care and trend analysis. It can also reduce length of hospital stay. Manual tracking of patient status is difficult. Collection of long-term databases of clinical data can be used in future diagnosis.

Human lives are directly involved in these application scenarios. The impacts will certainly influence the life of a person. It is well known that any wireless system has some inherent technical vulnerabilities and limitations. Many of the sensor networks applications in the healthcare are heavily relied on technologies that can pose security threats like eavesdropping and denial of services. There are concerns of health hazards for the implanted sensor devices. The above concerns have far reaching social implications. The social implications and issues that are directly related to the above mentioned application scenarios can be categorized into three major areas—security, privacy and legal issues. Besides these, there can be more issues such as economic and political issues. Hence before sensor networks applications in healthcare become a widely accepted concept, psychological, socio-political and a number of challenging system design issues should be taken care of. If resolved successfully, these systems will open a whole range of possible new applications that can significantly influence our lives [4].

In this paper, we discuss the security and privacy issues of wireless sensor networks application within healthcare perspective. Table 1 shows the Comparison between medical BAN and general WSN [27].

This paper has been further organized in the following manner. In section II, we discuss related works. In section III, we discuss the security issues. In section IV, we discuss the privacy issues and then finally in section V, the conclusion.

Related works

Research in healthcare applications of sensor devices are being under progress all over the world. Many projects are developed or in developing stage. A number of recent projects have focused on wearable health devices [15]. These projects are funded by both government agencies and private organizations.
Real life projects and applications

Some of the major indoor/outdoor application projects are mentioned here. These applications work on both real time and non-real time modes.

**HealthGear [12]** It is a product of Microsoft Research. It consists of a set of physiological sensors connected via Bluetooth to a cell phone. It is basically a wearable real-time health system for monitoring and analyzing physiological signals.

**MobiHealth [24]** It is a mobile healthcare project funded by the European Commission. It allows patients to be fully mobile while undergoing continuous health monitoring by utilizing UMTS and GPRS networks.

**Ubimon [26]** It is from the Department of Computing, Imperial College, London. The aim of this project is to address the issues related to using wearable and implantable sensors for distributed mobile monitoring. Two areas under consideration are the management of patients with arrhythmic heart disease and the follow-up monitoring of post operative care in patients who have had surgery.

**CodeBlue [20]** It is a research project at Harvard University, US. It integrates sensor nodes and other wireless devices into a disaster response setting. It is designed to work across various network densities and a wide range of wireless devices. From a tiny small sensor mote to more powerful devices such as PDSs, PCs can be combined in CodeBlue.
**Table 1** Comparison between medical BAN and general WSN

<table>
<thead>
<tr>
<th></th>
<th>Medical BAN</th>
<th>General WSN</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Common features</strong></td>
<td>Limited resources: battery, computation, memory, energy efficiency Diversity coexistence environment low/modest data rate, low/modest duty cycle Dynamic network scale, plug-and-play, heterogeneous devices ability, dense distribution</td>
<td></td>
</tr>
<tr>
<td><strong>Sensor/ actuator</strong></td>
<td>Single-function device Fast relative movement in small range device lifetime, days, &lt;10 years (implant sensor) Safe (low SAR) and quality first</td>
<td>Multi-function device Rare or slow movement in large range network lifetime and device lifetime, months, &lt;10 years Cost sensitive</td>
</tr>
<tr>
<td><strong>Dependability</strong></td>
<td>Reliability (first), guaranteed QoS Strongly security (except emergency)</td>
<td>Expected QoS, redundancy-based reliability Required security</td>
</tr>
<tr>
<td><strong>Networking</strong></td>
<td>Small scale star network No redundancy in device Random node distribution</td>
<td>Large scale hierarchical network redundant distribution Random node distribution</td>
</tr>
<tr>
<td><strong>Traffic</strong></td>
<td>Periodical real time (dominant), burst (priority) Uni-directional traffic M:1 communication</td>
<td>Burst (dominant), periodical Uni-directional or bi-directional traffic M:1 or point-point communication</td>
</tr>
<tr>
<td><strong>channel</strong></td>
<td>Specific medical channel, ISM band Body surface or through body</td>
<td>ISM band Obstacle is unknown</td>
</tr>
</tbody>
</table>

**eWatch** [7] It is a wearable sensor and notification platform developed for context aware computing research. It fits into a wrist watch form making it highly available, instantly viewable, and socially acceptable. eWatch provides tactile, audio and visual notification while sensing and recording light, motion, sound and temperature.

**The vital jacket** [21] A mobile device which is an intelligent wearable garment that is able to continuously monitor electrocardiogram (ECG) waves and Heart Rate for different fitness, high performance sports, security and medical applications. Here data can be sent via Bluetooth to a PDA and stored in a memory card at the same time.

IEEE 802.15.6 also known as Task Group 6 (or TG6) was formed in November 2007 to address the issues and standardize WBAN. The call for proposal was issued in January 2009 and heard till May 2009. The standard intends to address both medical/healthcare applications and other non-medical applications with diverse requirements [5]. The MAC layer in the standard intends to define short range, wireless communication in and around the body area. The standard aims to support a low complexity, low cost, ultra-low power and highly reliable wireless communication for use in close proximity to, or inside, a human body (but not limited to humans) to satisfy an evolutionary set of entertainment and healthcare products and services. The project will also address the coexistence issue with other WBAN and similar networks [5].

The major focus of these projects is to provide affordable services as well as cost effectiveness and power consumption of the devices.

**Related works in security and privacy**

The concerns for privacy and security have been investigated by some authors. The focus is normally on security related issues in general wireless sensor networks. But these issues as a whole for application scenarios in healthcare perspective have not yet been covered extensively. Many authors have suggested these issues as important. Authors in [5] discussed these issues in the e-Health monitoring applications. Authors in [9] also have discussed some of these issues for personal health monitoring. We have found that most published works address the security issues for sensor networks applications. These include works by authors such as [6] and [11]. Security issues are major concern raised by most authors. Privacy issue or other social implications are not discussed extensively regarding this field. We have mentioned the works done by various authors related to particular issues in the subsequent sections of this paper.

**Security issues**

Security is one of the most important aspects of any system. People have different perspective regarding security and hence it is defined in many ways. In general words, security is a concept similar to safety of the system as a whole. The US department of commerce site [23] has defined security as a condition that results from the establishment and maintenance of protective measures that ensure a state of inviolability from hostile acts or influences.
The communications in sensor networks applications in healthcare are mostly wireless in nature. This may result in various security threats to these systems. These threats and attacks could pose serious problems to the social life of an individual who is using the wireless sensor devices. In some cases such as tracking the location of a patient or person if compromised may lead to grave consequences. People with malicious intent may use the private data to harm the person.

Security issues in healthcare applications of sensor networks have been always part of active research. Security issues in general wireless sensor networks are a major area of research in recent times. Some works include [16, 17]. Similarly many people such as [5, 8, 11] have specifically addressed security issues with respect to healthcare applications. We have tried to highlight and discuss some threats and attacks in the following section along with some counter measures.

Threats and attacks

Security breach in healthcare applications of sensor networks is a major concern. It is also worth to mention that since healthcare applications of sensor networks are almost similar to WSN application environment, most of the security issues are also similar and hence comparable. The security issues can be divided into two broad levels: system security and information security. We have discussed these in the subsection of this discussion. Authors in [11] classified the threats and attacks [11] into two major categories—passive and active. A passive attack may occur while routing the data packets in the system. The attackers may change the destination of packets or make routing inconsistent. The attackers may also steal the health data by eavesdropping to the wireless communication media. Active threats are more harmful than their passive counter parts. Criminal minded people may find the location of the user by eavesdropping. This may lead to life threatening situation.

The normal trend of sensor device design is that they have little external security features and hence prone to physical tempering. This increases the vulnerability of the devices and poses tougher security challenges. Similarly vital data transmission from WBAN networks through GPRS or similar networks can be stolen by eavesdropping.

Authors in [5] have mentioned attacks in health monitoring in detail manner viz. eavesdropping and modification of medical data, forging of alarms on medical data, denial of service, location and activity tracking of users, physical tampering with devices and jamming attacks. People with bad intent may use the information for harmful activities. The attacks which can occur in any healthcare system using wireless sensor networks are shown in Table 2.

a. Data modification—The attacker can delete or replace part or all of eavesdropped information and send the modified information back to original receiver to achieve some illegal purpose. Health data are vital. Modifying them may result in system failure and cause disaster for a person.

b. Impersonation attack—If an attacker eavesdrops a wireless sensor node’s identity information, it can be used to cheat the other nodes.

c. Eavesdropping—For the open features of wireless channel used by sensor networks, any opponent can intercept radio communications between the wireless nodes freely and easily. Data stolen may be used for malicious acts.

d. Replaying—The attacker can eavesdrop a piece of valid information and resend it to original receiver after a while to achieve same purpose in different case.

Furthermore the attackers and hence the threats may be both internal and external. External attackers are not part of the system hence they are hard to deter. The primary purpose of these attacks is to steal valuable personal data. Since wireless media is always vulnerable than wired media, attackers find it easier. Once they are aware of the value of the personal health data, they may try to steal it by using both internal and external attacks.

Countering the attacks and measures

Any security issues must be resolved while designing the healthcare applications for sensor networks, or else they may give rise to serious social problems as discussed earlier. Authors in [28] argued that in the light of modern concepts of security, the safety should accompany the availability, scalability, efficiency and the quality parameters of inter-node communication. Hence countering the

<table>
<thead>
<tr>
<th>Attack assumptions</th>
<th>The risks to WBAN</th>
<th>Security requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computational capabilities</td>
<td>Data modification</td>
<td>Data integrity</td>
</tr>
<tr>
<td></td>
<td>Impersonation</td>
<td>Authentication</td>
</tr>
<tr>
<td>Listening capabilities</td>
<td>Eavesdropping</td>
<td>Encryption</td>
</tr>
<tr>
<td>Broadcast capabilities</td>
<td>Replaying</td>
<td>Freshness protection</td>
</tr>
</tbody>
</table>

Table 2 Security risks to WBAN and corresponding security requirements
system and information security threats should include all aspects of the network and its applications.

a) System Security

In WBAN scenario, where a person wears various devices, centralized control device can be used for data transmission from in and out of the network. This control device can also act as the gateway between internal network and outside world communication. Security measures such as authentication, firewalls and similar checks can be applied at the controller level to monitor the traffic as shown in Fig. 4.

Security in sensor networks applications in healthcare cannot be compromised. Outmost measures are necessary in this regards. We feel that security safeguard measures should be applied in three levels—Administrative, Physical and Technical. These also come under the domain of network management with emphasis on security.

i) Administrative Level Security

Effective administrative control is necessary to manage the system. Security measures should be applied to check the security breaches by the staff or people responsible for overall system operation. A well defined user hierarchy along with strong authentication measures may prevent security breaches at this level. The security measures must include kind of access mechanisms so that only authorized users can access the data. Similarly, it may be also a case where data forwarding may be only to the place or people which are previously authorized.

A start topology with all devices connected to a centralized system can help minimize overheads in network management. These will also help preventing attacks such as DoS and eavesdropping.

ii) Physical Level Security

At this level, measures may include controlling access to physical devices and data in the system for supposed stealing or tempering. The devices are vulnerable from people with malicious intent and from natural causes such as wear and tears. In case of natural disasters, the system may malfunction and may pose serious problems to the overall system operation. Hence, careful designing of devices to make them temper proof is necessary. But it is also understood that avoiding physical tempering of devices is hard to achieve. Another preventive measure can be that only authorized people should be allowed to physically handle the devices while in operation. Users must be strongly advised regarding this type of security measures.

iii) Technical Level Security

Technical level security checks are necessary mostly on hardware such as servers, disks, and other such devices. If the network is such that data is sent to central servers, server based security measures should be used at the server side and client based security at the end-user side. This is particularly necessary for safe propagation of information. This may again increase load on sensors at user side and thereby increase the overall cost. Hence some trade-offs between these issues will be necessary. It is also likely that more powerful motes will need to be designed in order to support the increasing requirements for computation and communication [16]. Securing the routing of data can also be applied as security measure. Wireless networks are very much susceptible to intrusion. Intrusion detection and prevention techniques are a must in these networks. Due to sensitive nature of healthcare applications, extra measures such as encryption of data, and constant monitoring of the network is necessary. While constant monitoring may not be a cost effective measure, encryption and creation of secure user groups can be effective as well as cost saving. Routing is another area where technical level security is required. If the data is sent to some remote host (e.g. doctors or some other hospital computers), routing is necessary. Attackers may cause routing inconsistencies resulting in wrong destination and receivers of data. Hence proper routing protocol and management is necessary to prevent such attacks.

b) Information security

WBAN applications contain not only medical but also personal information. Security and privacy are key concerns of all parts in WBAN. However, placing increasing amounts of valuable and confidential data on WBAN devices puts the data at serious risk to theft, sabotage, exploitation and manipulation. As shown in Table 2, there are several security risks surrounding any healthcare system using sensor networks.

The security mechanism of the system is responsible for providing the following security services on specified biomedical data when requested to do so by the applications.

a. Data Encryption—The data is encrypted so that it is not disclosed whilst in transit. Data encryption service provides confidentiality against eavesdropping attacks.

b. Data Integrity—Data integrity service consists of data integrity and data origin authentication. With data integrity the recipient can be sure that the data has not been altered or changed. Data origin authentication proves to the recipient that the stated sender has originated the data. It is an efficient method against data modification attack.

c. Authentication—Authentication service consists of association process among nodes. It is an efficient method against impersonation attacks.
d. Freshness Protection—This security service prevents the attacker from replaying the old frames that it eavesdropped by using nonce or time token.

To counter the major threats to information, two broad level security measures can be applied—encryption and authentication mechanisms. Any communication of personal health information and data over the networks must be encrypted. Authors in [29] proposed an ID-based cryptography and propose a novel secure architecture to enable secure communications in large-scale multi-domain wireless mesh networks. This can be extended easily to the healthcare applications of sensor networks which shows that clients can conveniently gain services securely even when roaming. Furthermore as mentioned by authors in [5], preventing unauthorized modifications of data while at the same time ensuring that only legitimate devices can create and inject data to the network prevents many of the previously discussed attacks. Authentication mechanisms can be used to ensure the data is coming from the person/entity it is claiming to be from [14]. We feel that even if the network is unattended for longer time, security measures should be always in highest priority mode. Authors in [30] discuss unattended network security in detail. It is well argued that adversarial models and defense techniques in prior WSN literature about security are unsuitable for the unattended WSN setting. This can be very helpful for healthcare applications where a person is not subject to constant monitoring or for monitoring elderly patients and illiterate people in remote/underdeveloped areas.

It should be noted that end to end security is must to make the wireless sensor networks in healthcare applications usable and acceptable by the common people. Threats such as tempering with data, denial of service (DoS), physical tampering and eavesdropping need far more special attention than any other common networks. Unless it happens acceptance of wireless sensor networks applications in healthcare will not be easy.

Privacy issues

Privacy is also among major concerns in wireless sensor networks with regard to healthcare applications. The health related data are always private in nature. Privacy issues arise from many reasons. It may be personal belief, social and cultural environment and other general public/private causes. Sending data out from a patient through wireless media can pose serious threats to the privacy of an individual. Concerns regarding privacy have been raised by some authors such as [3]. They have emphasized that if the issues associated with privacy are not honestly debated in a reasoned and open ways there is a risk that there will be a public backlash which will result in mistrust and consequently the technology will not be used for the many valuable applications where it can provide significant benefits. Whether the data are obtained with the consent of the person or without it due to the need by the system (for example emergency data from a patient), misuse or privacy concerns may restrict people from taking advantage of the full benefits from the system.

There are major questions raised by people from time to time. For example, authors in [8] have raised questions regarding guarding the privacy of an individual such as, where should the health data be stored, and who can view a patient’s medical record. There are also questions such as to whom should this information be disclosed to without the patient’s consent and who will be responsible for maintaining these data in case any problem arise, who will be held accountable. These are among several important issues that should be resolved in order to protect privacy as well as to some extent the security of the information.

In normal circumstances there are only few users of the data: the physicians, nurses and some other clinical/technical staffs. This limits the number of users in the system. Well defined regulations and firm guidelines regarding use of data for these users may limit the concerns
for privacy. But it should also be noted that in some cases such as emergency, disasters or remote patient monitoring may necessitate disclosure of information to other people in order to serve the patient in need. So the system must be flexible enough and users should be made to accept or compromise to some extent. Still procedures must be placed to make the users of the private healthcare data accountable for their actions or else these people may not care about the privacy concerns of an individual which may lead to bad implications on the social life of the person concerned. Authors in [24] have argued that without appropriate privacy safeguards the information may go into the public domain straight away, which is potentially undesirable for a number of reasons. People may not want some personal data be available in public domain. For example, early stage pregnancy, the details of certain medical conditions, may be made freely available to close family members and friends, but may not be appropriate for the general public. It is also important that these data should not fall into the hands of people with malicious intent and hence managing these types of data is very important in order to maintain the privacy of the person.

Privacy measures

Besides those mentioned above, some other measures may include:

a) All communications over wireless networks and Internet are required to be encrypted to protect the user’s privacy. Some countries have added this type of clause in their existing legal acts or enacted new laws. For example, the US Federal law HIPAA 1996 has this provision in it [22].

b) It is also necessary that, specific users should not be identified unless there is a need.

c) Another important measure is to create awareness in general public. It can be extremely beneficial if people are educated regarding security and privacy issues and their implications from now on. It is mentioned by authors in [3] that common people do not understand the technology and therefore may not be in a position to make balanced judgments concerning the extent to which it may have a negative impact on their own standards of privacy. Therefore educating the common people will greatly help in this regard.

The role of wireless infrastructure in healthcare applications is expected to become more prominent with an increasingly mobile society and with the deployment of mobile and wireless networks [13]. Hence it is always a better idea to be ready for such situations before the time comes for it. Educating people about the future ahead can make them more relaxed as well.

Conclusion

Sensor networks applications in healthcare being research and deployed all over the world. With the rise of these applications, implications will arise too. In this paper we tried to raise the concerns of major social implications like privacy and security. We have tried to analyze the cause and effects of these two issues. We feel that without taking care of these issues, the necessary growth and development will face major obstacles in coming future. Proper coordination between different government agencies, research institutes and manufactures is necessary to overcome these obstacles and have smooth implementation. General public should also be made aware of the benefits and implications so that they are better prepared. Rules and regulations like that of cyber laws and existing health regulations should be formalized and implemented.

Acknowledgment This research was supported by the MKE (Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the IITA (Institute of Information Technology Assessment) (IITA-2008-C1090-0801-0019).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

8. Meingast, M., Roosta, T., and Sastry, S., Security and privacy issues with health care information technology. 28th Annual International Conference of the IEEE Engineering in Medicine...
Personal health records: a scoping review

N Archer,1 U Fervier-Thomas,1 C Lokker,2 K A McKibbon,2 S E Straus3

ABSTRACT
Electronic personal health record systems (PHRs) support patient centered healthcare by making medical records and other relevant information accessible to patients, thus assisting patients in health self-management. We reviewed the literature on PHRs including design, functionality, implementation, applications, outcomes, and benefits. We found that, because primary care physicians play a key role in patient health, PHRs are likely to be linked to physician electronic medical record systems, so PHR adoption is dependent on growth in electronic medical record adoption. Many PHR systems are physician-oriented, and do not include patient-oriented functionalities. These must be provided to support self-management and disease prevention if improvements in health outcomes are to be expected. Differences in patient motivation to use PHRs exist, but an overall low adoption rate is to be expected, except for the disabled, chronically ill, or caregivers for the elderly. Finally, trials of PHR effectiveness and sustainability for patient self-management are needed.

INTRODUCTION
Widespread internet use and the availability of medical information on the web have made patients much more aware of symptoms, diagnostic tests, diseases, and treatment options. Much of the North American population relies on information from the internet for healthcare education and to make and reinforce decisions about medications, treatments, and lifestyle choices for themselves and others.1,2 Forty-two percent of the US population keeps health records for themselves and their families, with 87% of these in paper format.3 Many trials of electronic personal health record systems (PHRs) have shown that they supplement and improve patient and family access to knowledge for self-management of health and wellness issues. Although one study found 91% of PHR commercial products in use by firms, institutions, or governments,4 only 7% of American adults use PHRs. Growth in PHR use parallels the adoption of electronic medical record systems (EMRs) by primary care physicians (PCPs). This growth is because the EMRs are often used as a source of data for PHRs, and PCPs play a predominant role in advising and supporting patients in education and health self-management.5 PHRs have the potential to change and possibly to improve patient–provider relationships, enhance patient–physician shared decision making, and enable the healthcare system to evolve toward a more personalized medical model.5

There are a number of different fundamental designs for PHRs. We use the term PHR to refer to the records themselves and to the information systems used to support them. Electronic versions can include internet-based portals or computer-based applications. PHRs can be ‘tethered’, where subsets of information are provided by organizations that maintain patient data electronically, such as physician EMRs, health plan providers, hospitals, or employers.7 Untethered PHRs can be installed on isolated personal computers or internet-based portal services where only the user enters and maintains personal health data. Paper-based PHRs may also be used by patients to monitor their illnesses, and these can be carried physically by patients to doctor appointments, for example. Note that ‘patient’ is used interchangeably with ‘consumer’ in this paper, because all patients are consumers, and consumers almost always become patients.

The objective of this paper is to review the literature on PHRs and to describe the design, functionality, implementations, applications, outcomes, and perceived and real benefits of PHRs, with an emphasis on experience in the USA and Canada.

METHODS
Medline, Embase, CINAHL, Business Source Complete, IEEE and ACM Digital were searched from 1985 to March 2010. Included articles focused on studies assessing PHRs. For the purpose of this study, PHRs were defined as electronic or paper-based collections of health or wellness data arising from multiple sources about an individual’s health, that are managed, controlled, or shared by that individual or designate.

Search terms relating to PHRs included phr, ephr, patient internet portal, patient portal, patient-shared/held/carried record, patient accessible records, personal medical record. Two reviewers independently screened titles and abstracts to determine inclusion status. A second screen of article full-text, again by two independent reviewers, ensured that the studies described PHRs (see figure 1). Qualitative and quantitative articles of any study type were included.

Screening identified 2836 papers, of which 150 met the criteria. Article data were extracted manually into a database. Meta-analysis of the findings was not attempted because most of the papers lacked a common set of attributes that could be combined.

RESULTS
Ninety-two of the included studies reported research or evaluation data or referenced relevant descriptive supporting papers, and are categorized (based partially on Tang et al8) in table 1 as reporting on:
System attributes (record content, architecture, privacy and security, functionality, cost and financing)

Purpose (patient-provider communications, education and lifestyle changes, health self-management)

Adoption and acceptance (adoption and use, acceptance and satisfaction, usability)

Barriers to PHR adoption and use

Clinical outcomes and process changes.

In table 1, the total number of papers exceeds 92, because some contributed in more than one category. Functionality, patient-provider communication, and adoption and use were the categories most frequently addressed. Generally, most studies were oriented toward the care provider point of view, and only a few focused on self-management of chronic conditions. Some findings from the papers are discussed in more detail below.

System attributes

System attributes relate to the characteristics of PHR systems.

Content

Little consensus exists on what information to include in PHRs. Information from practitioner sources should use easy-to-understand language for laypersons.11 Information entered by patients may not be as complete, accurate, and organized as data exchanged between healthcare providers.17 Suggestions for data to be included in PHRs appear in table 2, based primarily on recommendations of the American Medical Informatics Association’s College of Medical Informatics,8 with additions from other researchers.91020 Certain chronic illnesses may require additional information.

Content must be important, understandable, and credible to patients and their caregivers, and appropriate for web access by patient-authorized individuals.18 Physician experience has shown that patient problem lists, clinical notes, medication and allergy data, and laboratory and diagnostic test results can be shared with patients.19 An attempt should be made (particularly in the case of bad news) to adjust office workflows so physicians can discuss results with patients before they appear in online records.

Table 1 Categorization of PHR papers and study type

<table>
<thead>
<tr>
<th>Categories</th>
<th>No.</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>System attributes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record content</td>
<td>13</td>
<td>Cohort studies3 10 Qualitative4 11–13 Observational14 Survey15–17 Descriptive18–20</td>
</tr>
<tr>
<td>Architecture</td>
<td>14</td>
<td>Cohort studies10 Qualitative8 Survey21 22 Descriptive18 23–31</td>
</tr>
<tr>
<td>Privacy, security</td>
<td>14</td>
<td>Cohort32–34 Survey15 21 25 35–37 Descriptive30 38–41</td>
</tr>
<tr>
<td>Functionality</td>
<td>27</td>
<td>RCTs42–44 Cohort3 23 34 45 46 Qualitative4 47 Survey15 16 48–50 Descriptive18 19 24 29 51–56 Systematic review57</td>
</tr>
<tr>
<td>Cost, financing</td>
<td>6</td>
<td>Cohort58 Survey15 59 60 Descriptive61 Systematic review62</td>
</tr>
<tr>
<td>Purposes of PHRs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient—provider communication</td>
<td>22</td>
<td>RCTs43 63–66 Cohort10 46 67 68 Qualitative4 8 12 47 69 Observational14 Survey36 37 48 70 Descriptive39 55 71 Systematic review57</td>
</tr>
<tr>
<td>Education, lifestyle changes</td>
<td>10</td>
<td>Cohort23 73 Qualitative4 12 74 75 Descriptive76 77</td>
</tr>
<tr>
<td>Health self-management</td>
<td>6</td>
<td>Qualitative4 78–80 Descriptive76 Systematic review8</td>
</tr>
<tr>
<td>Adoption, acceptance, and usability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adoption and use</td>
<td>39</td>
<td>RCTs42 43 65 81–83 Cohort23 33 37 45 46 67 68 84 Qualitative4 11–13 47 85–87 Observational14 Survey15 16 22 25 59 70 88 89 Descriptive18 24 25 52 54 55 80–92</td>
</tr>
<tr>
<td>Acceptance and satisfaction</td>
<td>6</td>
<td>Survey36 48 70</td>
</tr>
<tr>
<td>Usability</td>
<td>9</td>
<td>Qualitative4 1 12 Observational95 96 Survey21 36 Descriptive62 53 90 Systematic review57 94</td>
</tr>
<tr>
<td>Barriers to adoption and use</td>
<td>6</td>
<td>Survey15 21 Descriptive8 81 91 Systematic review97</td>
</tr>
<tr>
<td>Clinical outcomes and process changes</td>
<td>10</td>
<td>RCTs10 42 43 63–65 72 73 82 83</td>
</tr>
</tbody>
</table>

PHR, personal health record system; RCT, randomized controlled trial.
Architecture

Allowing patients to enter or view their own health data in their healthcare provider’s EMR can convey much more to the patient than stand-alone PHRs, enabling patients to gather their entire fragmented medical history in one place. Information that patients may keep for their personal use may also be valued by healthcare providers. Shared medical records are almost universally endorsed across ethnic and socio-economic groups. Linkages between PHRs and PCP EMRs thus appear to be critical to the successful use of PHRs.

A personally controlled PHR, integrated with a primary care EMR, can manage communications for prescriptions and appointments at reasonable cost. System interoperability is critical to giving consumers access to health records in hospital, physician, and laboratory systems, but this relies on the adoption of interoperability standards such as HL7 that support record sharing between systems. Other related approaches involve centralization of all patient records at a regional level, with access through online portals. Portability may also be important to consumers, although Lafsky and Horan found that the top preference of patients was to have information shared only within their circle of care.

Privacy and security

Two-thirds of adult consumers are concerned about the privacy and security of their health information, but most of those using PHRs are not worried about privacy implications. Those who are concerned about privacy may change their attitudes with appropriate framing of arguments favoring record use. The chronically and acutely ill and those who frequently use healthcare services tend to be less concerned about privacy than are health professionals. Current security protection mechanisms need to be enhanced for record protection, but to maintain privacy, security levels must not become so tight that health records are unusable. Work is underway (eg, Weider and Chan) to address security issues regarding wireless transmission of patient medical data, including the privacy, integrity, and confidentiality of the data, and the authentication and authorization of users.

Functionality

PHR functionalities can be classified as: (1) information collection, (2) information sharing and exchange, and (3) information self-management. Retrospective analyses of PHR implementations can provide data on functionalities. Functionalities include sending and receiving electronic messages to and from doctors’ offices; completing prescription renewal forms, appointments, and referral authorizations; viewing lists of current medications and allergies; and accessing health and practice information. Decision support can also assist patients in managing chronic illnesses, based on monitoring data. The nature of the patient’s illness affects preference for functionalities. For example, a ranked priority list of patients with diabetes included (priority percentages): a personal log to record and compare blood glucose levels (86%); a daily blood glucose log calculator to estimate diabetes control for the past 3 months (86%); links to educational websites (71%); an electronic newsletter for answers to questions, diabetes-friendly recipes, and information on community events (67%); online scheduling for routine appointments (67%); and e-reminders for appointments (67%).

Cost and financing

Public agencies are unlikely to fund PHRs unless they will reduce healthcare costs or substantially improve efficiency. Evaluating PHRs to determine if they improve health and reduce costs can be expensive. A systematic review of electronic health records in hospital settings showed evidence for cost reduction but little improvement in treatment quality. Similar studies are needed for PHRs. Consumers with serious chronic diseases, disabilities, or multiple health problems may experience direct and immediate benefits from PHRs, motivating continued use. Some data indicate that consumers might be willing to pay US$1.80 to US$4.50 per month for PHR support. Consumers with serious chronic illnesses may be willing to pay more.

Purpose of PHRs

Patient—provider communication

The benefits and satisfaction with PHRs have included easy access to test results and better communication with healthcare practitioners. A paper-based PHR was used successfully by patients to keep up-to-date records for encounters with different healthcare providers, reducing the need for inter-provider communications to access updated medical information. Patients prefer email communication for some interactions (eg,
requesting prescription renewals, obtaining general information, and in-person communication for others (eg, treatment instructions). Physicians generally prefer telephone or face-to-face communication. One EMR web portal designed to assist the self-management of ambulatory patients with diabetes included secure e-communication with the physician’s office, preventive healthcare reminders, and disease-specific tools and information. Over 2.5 years patient emails increased steadily. Participants reported that the system enhanced communication with the office, and that the reminder system was helpful.

Education and lifestyle changes

In addition to personal data and data from the provider EMR and monitoring devices (eg, weight, blood glucose), a PHR could store other data on, for instance, social status, family history, or living and work environment. It could also include information on healthy lifestyles (diet, exercise, smoking, weight loss, and working habits). In one study patients could access education and automated advice programs and add their own information to hospital systems. In this case, patients primarily reviewed laboratory results; patients and physicians reported enhanced communications and patient understanding. Jerden and Weinhard reported lifestyle changes associated with a paper PHR. After 6 months, 25% of patients reported changes in their health situation (exercise, diet, and habits related to stress). Patients may also benefit from sharing information on their conditions with others having similar problems using online patient communities, e-forums, private messaging, and comments. An example of such an online resource for cancer patients is the Caring Voices site at the Princess Margaret Hospital in Toronto, Canada.

Health self-management

Patient health self-management can be supported by PHRs that allow patients to record, edit, and retrieve their healthcare data, including blood glucose and blood pressure measurements, weight and activity logs, and stress scales. Frequent monitoring can lead to early detection of critical situations and timely intervention. Self-care monitoring tools are becoming more mobile and reliable, particularly in ‘smart home’ applications. Encouraging innovations are appearing in mobile monitoring and decision support tools for active individuals.

Adoption, acceptance, and usability

In the well-known Delone and McLean model of information systems success, there are several inter-related measures of success, including system quality, use, user satisfaction, individual impacts, and organizational impacts. In this section we review findings on related PHR characteristics, such as adoption, use, acceptance, satisfaction, and usability. A sustainable PHR implementation depends on positive results from all these characteristics as well as favorable individual and organizational impacts. While it is not the purpose of this study to review the Delone and McLean model of success in the PHR context, it is highly relevant in this environment.

Adoption and use

People with disabilities and chronic conditions, frequent users of healthcare services, and people caring for elderly parents tend to have the most interest in PHRs. Whether they actually adopt and use them is another matter. Simply providing online access to medical records is not useful unless the technology is integrated into the patient’s existing health and psychosocial support infrastructure. Participants tend to want unfettered access to health records and expect to use technologies to communicate with clinicians. Consumers tend to keep at least some paper records on test results and medications. Patients with chronic disease are more likely to keep summaries of their health histories, medications, and physician names. Others keep diaries, appointment notes, and questions for subsequent visits.

Low provider awareness and preparedness, and high patient expectations for personally controlled PHRs can reduce their chances of successful adoption and subsequent use. Other factors influencing adoption and use include educational and technical support for users and providers to handle access issues, especially among older consumers. Adoption issues among healthcare providers include new workflow demands and resistance to change, inadequate technology literacy, responsibility for ensuring the accuracy and integrity of health information across multiple interconnected data systems, and confidentiality and privacy risks. Ralston et al found monthly user rates per 1000 enrollees in an online health maintenance organization (HMO) system were: medical test results reviews (46), medication refills (37), patient-provider clinical messaging (27), after-visit summary review (27), medical condition review (20), appointment requests (10), immunization review (10), and allergy review (6). The adoption and use of PHRs reflect intense interest in patient health self-management, because PHRs can empower patients. In a study of 210 individuals, Lefkay and Horan found that one of the most important motivations for PHR users with disabilities is being able to use the system for support in emergencies, but this usually requires portable solutions such as paper records or smart cards. Users of PHRs tethered to healthcare providers have indicated a high value for PHRs in emergencies.

Physicians are less likely than patients to anticipate benefits, and more likely to anticipate problems from patient PHR use. They also have concerns about physician use of patient-entered information and whether PHR adoption will create unreimbursed work. Physicians are generally receptive to patient access to most laboratory and other EMR information, but with restricted access to physician notes. In some cases providers seemed to view PHRs as a source of medical information when the patient’s record is unavailable. Education of physicians thus appears to be needed on how PHRs can also support patient empowerment, disease prevention and control, and health self-management.

Acceptance and satisfaction

Denton found three compelling reasons motivating patients to maintain PHRs: serious chronic illnesses, unexpected health events, and the availability of inexpensive and secure computers. In a study of healthy, chronically ill, mentally ill, and pregnant patients, Fisher et al found that patient access to online medical records fitted three classifications: participation in care, quality of care, and self-care strategies. Patients felt that access helped reinforce trust and confidence in doctors and made them feel more like partners in healthcare. A measure of adoption success is sustainability—the degree to which an innovation continues to be used after initial efforts to secure adoption are completed. Sustainability was rarely if ever mentioned in any of the papers reviewed, although satisfaction, a related term, was often reported. For example, a survey of 4200 patients with access to certain healthcare provider EMR information reported that they found the system easy to use and their medical record information was complete, accurate, and understandable. Patient attitudes to the system were mostly positive,
but some patients were mildly concerned about confidentiality and privacy, and of learning about abnormal test results electronically. Clinicians were less positive and preferred telephone rather than email for routine interactions. Most studies suggest at least modest benefits ranging from enhanced doctor–patient communication to patient access to medical records. A systematic review assessed the effectiveness of paper PHRs for cancer patients and found that most patients welcomed them.

### Usability

Usability (user interface and support) is key to the adoption and use of PHRs. Kim et al. developed a web-based patient-centered PHR in a housing authority for low-income families, the elderly, and the disabled. Patients were assisted weekly by graduate nursing assistants. Patient use of the system was low, but users were satisfied with the system in general, paid more attention to their health conditions and care, and shared records with primary care providers, specialists, and nurse practitioners. They agreed that the system had improved their overall healthcare quality. This suggests that usability contributes through both the interface and user support for the elderly, for people who are uncomfortable with technology, and for the disabled who have technology difficulties.

Atreja et al. found that the perceptions of multiple sclerosis patients were similar to those of non-disabled patients (need for better health information, belief in the trustworthiness of online health content, excessive reliance on search engine results for health information). Use barriers were similar to those faced by the elderly (low contrast, inappropriate font size, poor navigational design, etc), while additional barriers were more specific to the disease (problems with flashing or moving objects, crowded or cluttered screens, and difficulty seeing red). These findings may be generalizable to the design of self-support systems for patients with specific diseases.

Usability also relates to non-user interface issues, such as communications management, where prompt response in high priority situations is important. This may involve re-engineering office workflows to increase patient satisfaction. Negative experiences may result from misunderstanding medical terms in doctors’ notes or laboratory tests. Solutions could include aids to assess the significance of laboratory and other diagnostic tests, hyperlinks that define technical terms, annotation of records that patients find in error, and electronic notifications when new or changed information appears. Hassol et al. found that electronic messaging assists patients communicating with providers about routine details (eg, renewing prescriptions), but is less appropriate for discussing complex and sensitive issues such as informing a patient about a new chronic condition or about disease management.

### Barriers to PHR adoption and use

PHR adoption has many perceived and real barriers. As with any new technology, failure can often be linked to little consumer involvement during planning, design, and implementation. Lack of trust in the provider is another barrier, as are poor computer and internet skills, fear of technology, inadequate access, low health literacy, and limited physical and cognitive abilities.

The ideal PHR appears to be one that provides access to all or most of the patient’s clinical information. This requires information from the consumer to be integrated through interoperable networks that gather information from facilities that have treated the patient, possibly assembled in the patient’s primary care provider’s EMR. Such PHRs are ‘integrated’ with the healthcare system. There are a number of technical and non-technical barriers to successfully implementing such ideal PHRs (summarized in table 3).

### Clinical outcomes and process changes

Ten outcome evaluation studies were found involving randomized controlled trials (RCTs) of PHRs. Although statistical results and their significance were reported in each study, there

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Barriers to PHR implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Barrier</strong></td>
<td><strong>Related issues</strong></td>
</tr>
<tr>
<td>Healthcare system and incentives</td>
<td>- Balancing between physician and patient autonomy</td>
</tr>
<tr>
<td></td>
<td>- Lack of technology training, interest, or ability of physicians</td>
</tr>
<tr>
<td></td>
<td>- Resistance to change</td>
</tr>
<tr>
<td></td>
<td>- Scope of work and responsibilities of healthcare providers</td>
</tr>
<tr>
<td></td>
<td>- Physician compensation and incentives</td>
</tr>
<tr>
<td></td>
<td>- Provider concerns about liability risks</td>
</tr>
<tr>
<td>Consumer confidence and trust</td>
<td>- Safeguarding the privacy of consumer medical information</td>
</tr>
<tr>
<td>Technical standards for system interoperability</td>
<td>- Data interchange standards</td>
</tr>
<tr>
<td></td>
<td>- Minimum data set standards in specific provider specializations</td>
</tr>
<tr>
<td></td>
<td>- Security and privacy standards</td>
</tr>
<tr>
<td></td>
<td>- Certification of health information technology products</td>
</tr>
<tr>
<td>Lack of EMR adoption by practitioners, institutions</td>
<td>- A major problem in the USA and Canada</td>
</tr>
<tr>
<td>Lack of health information technology infrastructure</td>
<td>- Lack of resources supporting system integration</td>
</tr>
<tr>
<td></td>
<td>- Range of existing non-compatible systems</td>
</tr>
<tr>
<td></td>
<td>- Need for mediating networks, organizational structures to support integration</td>
</tr>
<tr>
<td></td>
<td>- Limited online services at healthcare providers and institutions</td>
</tr>
<tr>
<td>Digital divide</td>
<td>- Considerations of racial, education, and socio-economic status</td>
</tr>
<tr>
<td></td>
<td>- Health literacy</td>
</tr>
<tr>
<td></td>
<td>- Special needs: visual, cognitive, or physical limitations</td>
</tr>
<tr>
<td></td>
<td>- Financial resources</td>
</tr>
<tr>
<td>Empirical justification</td>
<td>- Improved cost efficiency</td>
</tr>
<tr>
<td></td>
<td>- Healthcare effectiveness</td>
</tr>
<tr>
<td>Uncertainty in market demand</td>
<td>- Lack of success by many small vendors</td>
</tr>
<tr>
<td></td>
<td>- Increased certainty needed to drive investment in development of higher quality PHRs</td>
</tr>
</tbody>
</table>

EMR, electronic medical record system; PHR, personal health record system.
was little similarity among the conditions measured and meta-analysis was therefore not feasible. Of the 10 RCTs found, only three involved electronic PHRs. Two of these three focused on whether health promotion or coaching interventions changed prevention, detection, or management. Only Bourgeois et al. found improved recognition of signs of myocardial infarction or stroke. In a similar study involving paper records, knowledge, attitudes, and behavior concerning health promotion improved if patients were given computer-generated summaries and booklets. The third electronic PHR study found significant effects of patient access to vital health information on provider—patient communication through documentation of immunizations, allergies, medications, and surgical interventions.

In the six other PHR (paper-based) studies, one found no behavior change from preventive measures associated with written educational materials. In three of these studies, communications between providers and patients were examined; no improvements were reported in patient satisfaction with communications or with information provided to patients. One study of the effectiveness of a patient-held paper record for schizophrenic patients found no effect on satisfaction or the use of secondary care services. Finally, one study evaluated a comprehensive strategy of regular examinations and documentation to help improve diabetes care for participants. They found small improvements in HbA1c and diastolic blood pressure levels, but could not ascertain if this was due to effects of keeping medical records or to more regular examinations of the participants.

In general, the results of the few reported PHR RCTs are disappointing in that few significant effects were seen in the clinical endpoints. Although most users reported value in having access to more information, little if any improvement in actual health outcomes was reported.

DISCUSSION

Our scoping review has found that a significant amount of research is being done on PHR adoption, use, and satisfaction for various groups of users, with the main focus on providers. There is some evidence for the inclusion of certain functionalities in PHR systems, especially from the patient perspective, as gleaned from the utilities they use most. However, the clinical effectiveness and cost effectiveness of PHR interventions has not been adequately confirmed. From the limited and heterogeneous literature that was synthesized, the following themes emerged:

1. **Primary care physicians play a key role in the management of their patients’ health. Based on our review, we believe that sharing some proportion of their EMR records with patient PHRs can provide patients with useful information that allows them to be positively engaged in health self-management. A key to PHR adoption in North America is therefore rapid and continuing growth in physician adoption of EMRs from its current relatively low rate. In 2008, EMR adoption by primary care physicians was in the range of 24–28% in the USA, and 20–25% in Canada.** Primary care EMR adoption is likely to have grown considerably since then in both countries, due to provincial subsidies for EMR adoption in Canada, and the implementation of meaningful use requirements and significant allocations to healthcare information technology in the American Recovery and Reinvestment Act of 2009 in the USA.

2. **Although a number of good quality studies of PHRs have produced interesting results, many of these studies have been physician-oriented. Patients in the studies had access to their information through their doctors’ or hospital EMRs (tethered PHRs). EMRs are designed to provide doctors with the functionality and information they need, and their use for patients does not necessarily meet patient needs. Some studies included certain considerations of patient-oriented support such as the ability to join communities of interest, general information from high quality internet sites, information from healthcare professionals and internet sites on treatment programs for lifestyle, weight management, support for self-monitoring programs for chronic conditions, etc. But many did not. Until such integrated support is made available to patients, PHRs are not likely to demonstrate their full potential for supporting tangible or intangible improvements in patient health outcomes.**

3. **People with serious chronic conditions, individuals with disabilities, parents with small children, people with a strong interest in maintaining healthy lifestyles, and the elderly or their caregivers are more likely to adopt PHRs. Therefore, although a low overall PHR sustainable adoption rate can be expected, steps need to be taken by developers to improve the performance of PHRs and their long-term benefits for the people most likely to use them. This includes involving potential user groups with specific health self-management needs directly in requirement specification, design, and testing, to ensure that the PHRs match the cognitive abilities of their intended users and thereby support health self-management and disease prevention.**

4. **In a recent review of consumer health informatics, Gibbons et al. report that applications that provided individual tailoring, personalization, and behavioral feedback had the most significant impact on patient health outcomes. However, research is needed to develop a more detailed understanding of what motivates people to not only adopt but to continue using PHRs. Long-term sustainability of PHR use by patients was an issue that was not mentioned in any of the literature we examined. Sustainability involves not just positive results from factors such as adoption, use, acceptance, satisfaction, and usability, but favorable individual and organizational impacts. This is extremely important if healthcare systems are to avoid the specter of financing apparently successful PHR innovations that are abandoned or under-utilized by patients after an initial flurry of use.**

5. **RCTs are needed to test assumptions about the comparative effectiveness of PHRs on outcomes for various patient populations, using systems designed specifically for patient health self-management and disease prevention.**

Limitations

One of the limitations to this study is that new papers are being published quite regularly on PHRs, and some may have been missed in this study. Second, although we have tried to discuss some of the most important findings in the literature, it is impossible in a limited space to detail all the aspects we found that affect PHR attributes, purposes, benefits, usage, user satisfaction, and barriers to adoption and use.

CONCLUSIONS

The objective of this study was to describe existing electronic and paper-based PHR research and to determine whether PHRs can provide benefits to consumers/patients. We found many relevant papers, indicating a generally growing interest in PHR use, but there is much more to be done in tailoring PHRs for patient health self-management and sustainability. Although there is a large amount of survey, observational, cohort/panel,
and anecdotal evidence of PHR benefits and satisfaction for patients, more research is needed that gathers evidence to evaluate the results of PHR implementations in the context of works such as the Delone and McLean model of information systems success. At this point there is little solid evidence from RCTs or other studies of proven effectiveness in improved patient health outcomes through the use of PHRs. More research is also needed that addresses the current lack of understanding of optimal functionality and usability of these systems, and how they can play a beneficial role in supporting self-managed healthcare.

Acknowledgments The authors acknowledge the very helpful comments of the reviewers in improving the quality of this paper.

Funding This research was funded by a grant from the Canadian Institutes for Health Research. The funding agency did not have any role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; and preparation, review, or approval of the manuscript. Any opinions expressed are only those of the authors and do not necessarily represent the views of any of their affiliated institutions.

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

3. Taylor H. Two in five adults keep personal or family health records and almost everybody thinks this is a good idea. Health Care News 2004.
Qualitative research methods are enjoying unprecedented popularity. Although checklists have undoubtedly contributed to the wider acceptance of such methods, these can be counterproductive if used prescriptively. The uncritical adoption of a range of “technical fixes” (such as purposive sampling, grounded theory, multiple coding, triangulation, and respondent validation) does not, in itself, confer rigour.

In this article I discuss the limitations of these procedures and argue that there is no substitute for systematic and thorough application of the principles of qualitative research. Technical fixes will achieve little unless they are embedded in a broader understanding of the rationale and assumptions behind qualitative research.

Checklists in quantitative research

In medical research the question is no longer whether qualitative methods are valuable but how rigour can be ensured or enhanced. Checklists have played an important role in conferring respectability on qualitative research and in convincing potential sceptics of its thoroughness. They have equipped those unfamiliar with this approach to evaluate or review qualitative work (by providing guidance on crucial questions that need to be asked) and in reminding qualitative researchers of the need for a systematic approach (by providing an aide-mémoire of the various stages involved in research design and data analysis).

Qualitative researchers stress the importance of context but sometimes forget that research itself is carried out against an ever-changing backdrop. Now that it has secured a place in the methodological mainstream, qualitative research is increasingly being influenced by funding and editorial policies. Despite disclaimers by authors that their checklists should be viewed as being “reflective rather than constitutive of good research,” there is evidence that checklists are sometimes being used prescriptively.

Over the past two years, several researchers have informed me that they must comply with various procedures (such as respondent validation, multiple coding, etc) in order to satisfy the requirements of specific journals where they hope to publish their work. (I am not concerned here with the accuracy of such claims, although my own experience suggests these are exaggerated.) While we all attempt to tailor our writing to match the style and format of the journal in question, the strategic adoption of such technical fixes has wider repercussions. The complex dilemmas in research design that qualitative researchers face with regard to sampling, choice of methods, and approaches to analysis cannot be solved by formulaic responses. If we succumb to the lure of “one size fits all” solutions we risk being in a situation where the tail (the checklist) is wagging the dog (the qualitative research).

From reading recent journals and my experience of reviewing journal articles and grant submissions, I find that the five technical fixes currently enjoying the greatest popularity are purposive sampling, grounded theory, multiple coding, triangulation, and respondent validation (table). The rest of this article outlines their limitations and provides a more realistic appraisal of their potential.

Summary points

Checklists can be useful improving qualitative research methods, but overzealous and uncritical use can be counterproductive

Reducing qualitative research to a list of technical procedures (such as purposive sampling, grounded theory, multiple coding, triangulation, and respondent validation) is overly prescriptive and results in “the tail wagging the dog.”

None of these “technical fixes” in itself confers rigour; they can strengthen the rigour of qualitative research only if embedded in a broader understanding of qualitative research design and data analysis.

Otherwise we risk compromising the unique contribution that systematic qualitative research can make to health services research.

Purposive sampling

Rather than aspiring to statistical generalisability or representativeness, qualitative research usually aims to reflect the diversity within a given population. In the past qualitative research often relied on convenience samples, particularly when the group of interest was difficult to access. Purposive (or theoretical) sampling, however, offers researchers a degree of control rather than a pre-determined role.
<table>
<thead>
<tr>
<th>Technical fix</th>
<th>Concerns addressed</th>
<th>Realistic potential</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purposive sampling</td>
<td>Bias</td>
<td>Enhancing sample coverage and providing a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>framework for analysis</td>
</tr>
<tr>
<td>Grounded theory</td>
<td>Original theorising</td>
<td>Developing existing theory or, occasionally,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>new theories</td>
</tr>
<tr>
<td>Multiple coding</td>
<td>Inter-rater reliability</td>
<td>Refining interpretations or coding frameworks</td>
</tr>
<tr>
<td>Triangulation</td>
<td>Confirmation or refutation of internal validity</td>
<td>Corroborating or, more often, finding refluences</td>
</tr>
<tr>
<td>Respondent validation</td>
<td>Confirmation or refutation of interpretations</td>
<td>Corroborating or, more often, refining findings</td>
</tr>
</tbody>
</table>

than being at the mercy of any selection bias inherent in pre-existing groups (such as clinic populations). With purposive sampling, researchers deliberately seek to include “outliers” conventionally discounted in quantitative approaches. It allows for such deviant cases to illuminate, by juxtaposition, those processes and relations that routinely come into play, thereby enabling “the exception to prove the rule.”

Some strategies claimed as examples of purposive sampling in effect involve hybrids, which retain elements of random or convenience sampling and which are unlikely to yield the spread of respondents required. When they are provided at all, details of sampling are often dealt with in the methods section of papers and are disregarded in the analysis section, which often consists of little more than a description of undifferentiated themes that emerged during data analysis. For example, we are likely to be told in the methods section that a third of the sample were men, but the analysis section does not discuss how their perspectives differed from those of female respondents.

Such approaches do not use qualitative datasets to full advantage. That would involve applying the constant comparative method to continuously compare the views and experiences of respondents who have been selected precisely—indeed, purposively—in order to illuminate subtle but potentially important differences. In other words, samples may have been selected purposively, but they are not being used purposefully to interrogate the data collected.

**Grounded theory**

In its purest form the grounded theory approach to data analysis alleges that all explanations or theories are derived from the dataset itself rather than from a researcher’s prior theoretical viewpoint. In practice, however, you are unlikely to obtain research funding without having carried out a thorough literature review or having formulated some idea of the content of the data you are likely to collect.

According to many researchers who invoke the concept of grounded theory, coding categories reflect the content of data collected rather than the questions on the interview schedule or focus group topic guide and often use concepts or vocabulary borrowed from respondents. However, few published papers yield the surprises likely to be a feature of analyses driven entirely by respondents’ concerns, and the terminology and theories to which papers appeal generally bear an uncanny resemblance to current disciplinary concerns and debates.

Bryman and Burgess have criticised the use of grounded theory as “an approving bumper sticker” invoked to confer academic respectability rather than as a helpful description of the strategy used in analysis. Melia claims that most researchers use a pragmatic variant, whereby they can achieve added value by identifying new themes from the data alongside those that could have been anticipated from the outset. All too often, however, the tension between these two different sorts of insight—and its potential to illuminate the topic being studied—is not explored in the presentation of findings.

In the absence of an attempt to systematically analyse the commonalities and contradictions reflected in the data, many researchers produce an artificially neat and tidy account that is descriptive rather than analytical and which militates against formulating in-depth analyses. Uncritical adoption of grounded theory can result in explanations tinged with the “near mysticism” that Melia derides in the original text on grounded theory.

A slight of hand produces a list of “themes,” and we are invited to take it on trust that theory somehow emerges from the data without being offered a step by step explanation of how theoretical insights have been built up.

**Multiple coding**

Multiple coding concerns the same issue as the quantitative equivalent “inter-rater reliability” and is a response to the charge of subjectivity sometimes levelled at the process of qualitative data analysis. Although multiple coding does not usually demand complete replication of results, it does involve the cross checking of coding strategies and interpretation of data by independent researchers. While I would caution against multiple coding of entire datasets (on the grounds of economy in both cost and effort), some element of multiple coding can be a valuable strategy. It can be useful to have another person cast an eye over segments of data or emergent coding frameworks, and this is a core activity of supervision sessions and research team meetings.

Although six experienced researchers who independently coded one focus group transcript showed substantial agreement, Armstrong et al found considerable variation in the ways that they packaged coding frameworks (including the language used). This is not surprising, given the complexity of qualitative data and the range of disciplinary backgrounds and interests of qualitative researchers. Indeed, Mauthner et al have shown how researchers’ original interpretations may shift when they revisit previously collected data.

However, the degree of concordance between researchers is not really important; what is ultimately of value is the content of disagreements and the insights that discussion can provide for refining coding frames. The greatest potential of multiple coding lies in its capacity to furnish alternative interpretations and thereby to act as the “devil’s advocate” implied in many of the checklists in alerting researchers to all potentially competing explanations. Such exercises encourage thoroughness, both in interrogating the data at hand and in providing an account of how an analysis was developed. Whether this is carried out by a conscientious lone researcher, by a team, or by involving independent experts is immaterial: what matters is that a systematic process is followed and that this is rendered transparent in the written research project.
Triangulation

The current heavy reliance on triangulation in grant applications testifies both to the respect accorded to this concept and to its perceived value in demonstrating rigour. Triangulation addresses the issue of internal validity by using more than one method of data collection to answer a research question. In principle, it sounds eminently feasible to combine, say, observational fieldwork and interviews or focus groups in order to get a broader view. However, triangulation is difficult to perform properly: data collected using different methods come in different forms and defy direct comparison. This is true for different types of qualitative data, such as interview and focus group transcripts, as well as for the more obvious differences between qualitative and quantitative data.

The production of similar findings from different methods merely provides corroboration or reassurance; the absence of similar findings does not, however, provide grounds for refutation. This is because different methods used in qualitative research furnish parallel datasets, each affording only a partial view of the whole picture.

Triangulation relies on the notion of a fixed point, or superior explanation, against which other interpretations can be measured. Qualitative research, however, is usually carried out from a relativist perspective, which acknowledges the existence of multiple views of equal validity. Therefore, it does not readily lend itself to the production or observance of such a hierarchy of evidence. Richardson suggests that it is more helpful to conceive of complementary rather than competing perspectives and offers the term “crystallisation” as an alternative to triangulation. Qualitative research, with its distinctive approach to harnessing the analytical potential of exceptions, allows a research question to be examined from various angles. As Mays and Pope conclude, comprehensiveness may be a more realistic goal for qualitative research than is internal validity. According to this approach, apparent contradictions (or exceptions) do not pose a threat to researchers’ explanations; they merely provide further scope for refining theories.

Respondent validation

Given the current focus on consumerism, respondent validation, which involves cross checking interim research findings with respondents, has a ready appeal. Respondents’ reactions to emerging findings can certainly help refine explanations—as can key informants—but several commentators have questioned whether it is always appropriate. As Mays and Pope point out, researchers seek to provide an overview whereas respondents have individual concerns, and this can result in apparently discrepant accounts. Sometimes researchers choose to disregard their own interpretations and to accept those of respondents at face value. This can be costly but may lead to conclusion: Atkinson has warned of the dangers of “romanticising” respondents’ accounts. Respondent validation exercises, such as reading of drafts, make considerable demands on participants’ time and, depending on the research topic and content of transcripts, can even be exploitative or distressing.

Respondent validation can be particularly valuable in action research projects, where researchers work with participants on an ongoing basis to facilitate change. Most health services research, however, involves a one-off data collection exercise, in which respondent validation may be more trouble than it is worth.

Conclusion

Although some of the technical fixes discussed here may seem appealing in the face of the dual imperatives of securing grant funding and publication, each has limitations. Reducing qualitative research to a list of technical procedures, however extensive, is overly prescriptive and results in “the tail wagging the dog.” None of these technical fixes, in itself, confers rigour. They can strengthen the rigour of qualitative research only if they are embedded in a broad understanding of qualitative research design and data analysis. Otherwise we run the risk of compromising the unique contribution that systematic and thoughtfully carried out qualitative research can make to health services research.

This article is based on a presentation to the British Sociological Association’s Regional Medical Sociology Group in London in March, 2000. I am grateful to those who attended for their constructive feedback; also to Helen Richards and Graham Watt for helpful comments on an earlier draft. Competing interests: None declared.

1 Pyper J, Williams G, Rogers A. Rationale and standards for the systematic review of qualitative literature in health services research. Qual Health Res 1998;8:341-51.
17 Mauthner NS, Parry O, Backen-Milburn K. The data are out there, or are they? Implications for archiving and revisiting qualitative data. Sociology 1998;32:733-45.
22 Atkinson P. Narrative turn or blind alley? Qual Health Res 1997;7:325-44.
Short Communication


Till Bärnighausen,1 Frank Tanser,1 and Marie-Louise Newell1,2

Abstract

To understand the dynamics of the HIV epidemic and to plan HIV treatment and prevention programs, it is critical to know how HIV incidence in a population evolves over time. We used data from a large population-based longitudinal HIV surveillance in a rural community in South Africa to test whether HIV incidence in this population has changed in the period from 2003 through 2007. We observed 563 seroconversions in 8095 individuals over 16,256 person-years at risk, yielding an overall HIV incidence of 3.4 per 100 person-years (95% confidence interval 3.1–3.7). We included time-dependent period dummy variables (in half-yearly increments) in age-stratified Cox regressions in order to test for trends in HIV incidence. We first did regression analyses separately for women and men. In both regressions, the coefficients of all period dummy variables were individually insignificant (all $p > 0.338$) and jointly insignificant ($p = 0.764$ and $p = 0.111$, respectively). We then did regression analysis using the pooled data on women and men, controlling for sex and interactions between sex and age. Again, the coefficients of the eight period dummy variables were individually insignificant (all $p > 0.387$) and jointly insignificant ($p = 0.701$). We show for the first time that high levels of HIV incidence have been maintained without any sign of decline over the past 5 years in both women and men in a rural South African community with high HIV prevalence. It is unlikely that the HIV epidemic in rural South Africa can be reversed without new or intensified efforts to prevent HIV infection.

Estimates of the development over time of HIV incidence in the general population are crucial for understanding the dynamics of the HIV epidemic, assessing the population impact of HIV prevention strategies, and predicting antiretroviral treatment need. Several countries in sub-Saharan Africa have in recent years experienced declines in HIV prevalence among pregnant women attending antenatal clinics, including Botswana,1 Ethiopia,2 Kenya,3 Malawi,4 Zambia,5 and Zimbabwe.6 In South Africa, the country with the largest absolute number of HIV-positive people worldwide, HIV prevalence among women attending antenatal clinics has been reported to be leveling off after nearly two decades of steady increase.7,8 Trends in antenatal HIV prevalence are sometimes taken to suggest similar trends in HIV incidence in the general population.1,7 However, such inference may be incorrect. Findings based on surveys of women in antenatal care may not generalize to the population as a whole.8,9 Moreover, changes in HIV prevalence are a function of both the rate of new infections and mortality in HIV-positive people, so that trends in HIV prevalence do not imply specific trends in incidence. For instance, Wawer and colleagues find that “HIV-related mortality contributed most to the [HIV] prevalence decline” observed in Uganda from 1990 to 1992, while HIV incidence did not change significantly over the period.10

Population trends in HIV incidence can be directly measured through repeated HIV testing of the same individuals in longitudinal HIV surveillance. However, longitudinal surveillances are rare, because they are difficult to establish and costly to maintain.11 Only a few past studies in Africa have thus analyzed population trends in HIV incidence using longitudinal data.10,12–15 We use data from a large population-based longitudinal HIV surveillance conducted by the Africa Centre for Health and Population Studies, University of

1Africa Centre for Health and Population Studies, University of KwaZulu-Natal, Mtubatuba, South Africa.
2Centre for Paediatric Epidemiology and Biostatistics, Institute of Child Health, University College of London, London, UK.
KwaZulu-Natal, in a rural community in the Umkhananyakude district of KwaZulu-Natal, South Africa, to test whether HIV incidence in this population has changed in the period from 2003 through 2007. Previous studies in this community found high levels of crude HIV prevalence in adults (21.4% in 2003/4)\(^{16}\) and crude HIV incidence (3.2 per 100 person-years from 2003 through 2005).\(^{17}\) Over the past years, the community in which the surveillance took place has been exposed to HIV prevention campaigns organized both at the national level (such as the loveLife campaign)\(^{18}\) and the local level (such as the Star for Life program).\(^{19}\) and has had access to voluntary counseling and testing (VCT) in the public and private sector.\(^{20,21}\) A public-sector antiretroviral treatment program was established in the study area in October 2004; by July 2008, more than 5200 patients were receiving antiretroviral medication through the program.

The data for this study were collected between July 2003 (the start of the HIV surveillance) and December 2007, covering the first four rounds of the surveillance. The study area, data collection, and laboratory methods of the surveillance have been described elsewhere.\(^{22}\) All participants in the surveillance provided written informed consent for the analysis of their information, including HIV status. Residents in the surveillance area had to meet age eligibility criteria (women between 15 and 49 and men between 15 and 54 years of age) in order to participate in the HIV surveillance. Of 36,813 eligible residents ever contacted in the first four rounds of the surveillance 21,709 consented at least once to provide a blood sample for HIV testing (i.e., the participation proportion over the first four surveillance rounds was 59%). We enrolled all individuals into the HIV incidence cohort for this study who have been described elsewhere.\(^{22}\) All participants in the surveillance were randomized to the HIV incidence analysis contributed 16,256 person-years at risk and 563 seroconversions to the analysis (Table 2), yielding an overall HIV incidence of 3.4 per 100 person-years [95% confidence interval (95% CI) 3.1–3.7]. Figure 1 shows the point estimates and 95% CIs of crude HIV incidence by half-year calendar period from June 2003 through December 2007. The 95% CIs of crude HIV incidence by half-year calendar period overlap with the 95% CIs of the incidence estimates in all other periods. To test whether HIV incidence varied significantly by time period, we included calendar period at risk as a time-dependent covariate in half-yearly increments (from the period July 2003–December 2003

<table>
<thead>
<tr>
<th>Eligible</th>
<th>Included in analysis</th>
<th>Follow-up proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Women (n)</td>
<td>Men (n)</td>
</tr>
<tr>
<td>15–19</td>
<td>2963</td>
<td>3076</td>
</tr>
<tr>
<td>20–24</td>
<td>1080</td>
<td>1017</td>
</tr>
<tr>
<td>25–29</td>
<td>481</td>
<td>356</td>
</tr>
<tr>
<td>30–34</td>
<td>442</td>
<td>238</td>
</tr>
<tr>
<td>35–39</td>
<td>578</td>
<td>218</td>
</tr>
<tr>
<td>40–44</td>
<td>697</td>
<td>249</td>
</tr>
<tr>
<td>45–49</td>
<td>607</td>
<td>279</td>
</tr>
<tr>
<td>50–54</td>
<td>213</td>
<td>133</td>
</tr>
<tr>
<td>Total</td>
<td>6848</td>
<td>5646</td>
</tr>
</tbody>
</table>

\(^{a}\)CI, confidence interval.

Table 2. Cox Regression Analysis of HIV Acquisitiona,b

<table>
<thead>
<tr>
<th>Regression</th>
<th>(1) Women</th>
<th>(2) Men</th>
<th>(3) Women and men</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004/I</td>
<td>1.07 (0.36)</td>
<td>0.76 (0.53)</td>
<td>0.99 (0.30)</td>
</tr>
<tr>
<td>2004/II</td>
<td>1.11 (0.37)</td>
<td>1.12 (0.72)</td>
<td>1.09 (0.32)</td>
</tr>
<tr>
<td>2005/I</td>
<td>0.87 (0.30)</td>
<td>1.33 (0.86)</td>
<td>0.99 (0.30)</td>
</tr>
<tr>
<td>2005/II</td>
<td>0.681 (0.30)</td>
<td>0.656 (0.86)</td>
<td>0.964 (0.30)</td>
</tr>
<tr>
<td>2006/I</td>
<td>0.94 (0.32)</td>
<td>0.90 (0.60)</td>
<td>0.92 (0.28)</td>
</tr>
<tr>
<td>2006/II</td>
<td>0.85 (0.30)</td>
<td>0.56 (0.39)</td>
<td>0.76 (0.24)</td>
</tr>
<tr>
<td>2007/I</td>
<td>1.09 (0.38)</td>
<td>0.65 (0.46)</td>
<td>0.97 (0.30)</td>
</tr>
<tr>
<td>2007/II</td>
<td>1.10 (0.40)</td>
<td>0.85 (0.62)</td>
<td>1.04 (0.34)</td>
</tr>
<tr>
<td>2008/I</td>
<td>0.81 (0.37)</td>
<td>2.08 (1.58)</td>
<td>1.07 (0.41)</td>
</tr>
<tr>
<td>2008/II</td>
<td>0.643 (0.37)</td>
<td>0.338 (0.86)</td>
<td>0.859 (0.32)</td>
</tr>
</tbody>
</table>

Female sex

- - - 5.89 (1.17) <0.0001

Sex*Age 20–24

- - - 0.34 (0.09) <0.0001

Sex*Age 25–29

- - - 0.16 (0.05) <0.0001

Sex*Age 30–34

- - - 0.16 (0.07) <0.0001

Sex*Age 35–39

- - - 0.14 (0.06) <0.0001

Sex*Age 40–44

- - - 0.14 (0.07) <0.0001

Sex*Age 45–49

- - - 0.10 (0.05) <0.0001

N 4733 3362 8095
Seroconversions 427 136 563
Person-years at risk 9776 6480 16,256
Log likelihood 4.94 13.03 139.53
LR χ2 0.764 0.111 <0.00001

The numbers in the cells are the adjusted hazard ratio (aHR), standard error of the aHR, and p value. LR = likelihood ratio, I = first half of a year, II = second half of a year, *indicates interaction.

The Cox baseline hazard is stratified by 5-year age category. The reference category in regression (1) and (2) is 2003. The reference categories in regression (3) are 2003/II, male sex, and Sex*Age 15–19.

FIG. 1. Crude HIV incidence by half-year period.

To the period July 2007–December 2007 in Cox regression analysis of HIV acquisition. When we included age (at the time of the first negative HIV test) in a number of different functional forms on the right-hand side of the Cox regressions equations (linear age term, third-order polynomial of the centered age variable, and 5-year age categories), the Cox proportional hazards assumption was consistently violated, as diagnosed by the Grambsch–Thernau test. We thus stratified the Cox baseline hazard by age category (in 5-year increments, starting at age 15).

We know from previous studies in this population that the age pattern of HIV incidence differs by sex.25 We first did regression analyses separately for women and men [Table 2, regressions (1) and (2)]. In both regressions, the coefficients of the eight time-dependent period dummy variables were individually insignificant (all p ≥ 0.338) and jointly insignificant [likelihood ratio test of joint significance, χ2 4.94 (p = 0.764) in regression (1) and χ2 13.03 (p = 0.111) in regression (2)]. To maximize the statistical power to detect significant differences in the hazard of HIV acquisition by time period, we then did regression analysis using the pooled data on women and men. In this regression, we controlled for sex and interactions between sex and the 5-year age strata. No interaction term between sex and the oldest age stratum (50–54 years of age) was included in the regression because women older than 49 years of age were not eligible to participate in the surveillance. The results of the regression using the pooled data confirmed the results of the separate regressions for women and men. The coefficients of the eight period dummy variables were individually insignificant (all p ≥ 0.387) and jointly insignificant [likelihood ratio test of joint significance, χ2 5.52 (p = 0.701)] [Table 2, regression (3)]. In none of the three final regressions was the proportional hazards assumption violated [for any of the individual independent variables (all p ≥ 0.104) or globally (all p ≥ 0.517)]. All analyses were undertaken using STATA 10.1.

A previous longitudinal population-based study in another rural community in South Africa (in Limpopo province) between 2001 and 2004 found HIV incidence levels similar to those in this community (4.9 and 2.2 per 100 person-years in women and men, respectively).25 Our study demonstrates for the first time that such high levels of HIV incidence in women and men have been maintained in a rural South African community without any sign of decline over the past 5 years. This finding supports the conclusion in the 2008 UNAIDS Report that “there is no evidence yet of major changes in HIV-related behaviour” in South Africa.26
would have been without past and present HIV prevention efforts, it is clear from our findings that the prevention strategies used are not sufficient to reduce the spread of the epidemic, primarily because they do not reach sufficient numbers of people, they do not reach the right target groups, or they are ineffective. It is unlikely that the HIV epidemic in rural South Africa can be reversed without new approaches (such as male circumcision programs and behavioral interventions targeting HIV-positive people) or intensified effort to prevent HIV infection using established methods (such school-based sexual health interventions).

Acknowledgments

We thank Phumzile Dlamini, Thobeka Mgomezulu, Zanomsa Gqwede, Claudia Wallrauch, Kubus Herbst, and the field staff at the Africa Centre for Health and Population Studies at the University of KwaZulu-Natal, South Africa, for their work in collecting the data used in this study and the communities in the Africa Centre demographic surveillance area for their support and participation in this study. Core funding for the Africa Centre’s Demographic Surveillance Information System (GR065377/Z/01/H) and Population-based HIV Survey (GR065377/Z/01/B) was received from the Wellcome Trust, UK. Till Barnighausen and Frank Tanser are supported by Grant 1R01-HD058482-01 from the National Institute of Child Health and Human Development (NICHD). The funding organizations had no role in the design and conduct of the study, in the collection, analysis, and interpretation of the data, or in the preparation, review, or approval of the manuscript.

Disclosures

No competing financial interests exist.

References


Address reprint request to:
Till Bärnighausen
Africa Centre for Health and Population Studies
University of KwaZulu-Natal
PO Box 198
Mtubatuba 3935, South Africa

E-mail: tbarnighausen@africacentre.ac.za
This article has been cited by:

1. Samuelina Arthur, Martin Bangha, Osman Sankoh. 2013. Review of contributions from HDSSs to research in sexual and reproductive health in low- and middle-income countries. *Tropical Medicine & International Health* 18:12, 1463-1487. [CrossRef]


5. John Imrie, Graeme Hodkinott, Sebastian Fuller, Stephen Oliver, Marie-Louise Newell. 2013. Why MSM in Rural South African Communities Should be an HIV Prevention Research Priority. *AIDS and Behavior* 17:51, 70-76. [CrossRef]

6. Carol S. Camlin, Rachel C. Snow, Victoria Hosegood. 2013. Gendered Patterns of Migration in Rural South Africa. *Population, Space and Place* n/a-n/a. [CrossRef]


The Impact of eHealth on the Quality and Safety of Health Care: A Systematic Overview

Ashly D. Black1, Josip Car1, Claudia Pagliari2, Chantelle Anandan2, Kathrin Cresswell2, Tomislav Bokun1, Brian McKinstry2, Rob Proctor3, Azeem Majeed4, Aziz Sheikh2*

1 eHealth Unit, Department of Primary Care and Public Health, Imperial College London, London, United Kingdom, 2 eHealth Research Group, Centre for Population Health Sciences, The University of Edinburgh, Edinburgh, United Kingdom, 3 National Centre for e-Social Science, University of Manchester, Manchester, United Kingdom, 4 Department of Primary Care and Public Health, Imperial College London, London, United Kingdom

Abstract

Background: There is considerable international interest in exploiting the potential of digital solutions to enhance the quality and safety of health care. Implementations of transformative eHealth technologies are underway globally, often at very considerable cost. In order to assess the impact of eHealth solutions on the quality and safety of health care, and to inform policy decisions on eHealth deployments, we undertook a systematic review of systematic reviews assessing the effectiveness and consequences of various eHealth technologies on the quality and safety of care.

Methods and Findings: We developed novel search strategies, conceptual maps of health care quality, safety, and eHealth interventions, and then systematically identified, scrutinised, and synthesised the systematic review literature. Major biomedical databases were searched to identify systematic reviews published between 1997 and 2010. Related theoretical, methodological, and technical material was also reviewed. We identified 53 systematic reviews that focused on assessing the impact of eHealth interventions on the quality and/or safety of health care and 55 supplementary systematic reviews providing relevant supportive information. This systematic review literature was found to be generally of substandard quality with regards to methodology, reporting, and utility. We thematically categorised eHealth technologies into three main areas: (1) storing, managing, and transmission of data; (2) clinical decision support; and (3) facilitating care from a distance. We found that despite support from policymakers, there was relatively little empirical evidence to substantiate many of the claims made in relation to these technologies. Whether the success of those relatively few solutions identified to improve quality and safety would continue if these were deployed beyond the contexts in which they were originally developed, has yet to be established. Importantly, best practice guidelines in effective development and deployment strategies are lacking.

Conclusions: There is a large gap between the postulated and empirically demonstrated benefits of eHealth technologies. In addition, there is a lack of robust research on the risks of implementing these technologies and their cost-effectiveness has yet to be demonstrated, despite being frequently promoted by policymakers and “techno-enthusiasts” as if this was a given. In the light of the paucity of evidence in relation to improvements in patient outcomes, as well as the lack of evidence on their cost-effectiveness, it is vital that future eHealth technologies are evaluated against a comprehensive set of measures, ideally throughout all stages of the technology’s life cycle. Such evaluation should be characterised by careful attention to socio-technical factors to maximise the likelihood of successful implementation and adoption.

Please see later in the article for the Editors’ Summary.


Academic Editor: Benjamin Djulbegovic, University of South Florida, United States of America

Received November 12, 2009; Accepted November 19, 2010; Published January 18, 2011

Copyright: © 2011 Black et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was funded by the NHS Connecting for Health Evaluation Programme (NHS CFHEP 001), http://www.haps.bham.ac.uk/publichealth/cfhep/. ADP is supported by an ESRC PhD studentship in eHealth. BM is supported by a CSO Career Scientist Award. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: BM has received research grants from Intel, Tunstall, Chief Scientist Office, and NHS Lothian to explore the use of Telehealthcare in long-term conditions. AD is on the PLOS Medicine Editorial Board.

Abbreviations: CDSS, computerised decision support system; CPOE, computerised provider (or physician) order entry; EHR, electronic health record; ePrescribing, electronic prescribing; NHS, National Health Service; PACS, picture archiving and communication system

* E-mail: aziz.sheikh@ed.ac.uk
Introduction

Implementations of potentially transformative eHealth technologies are currently underway internationally, often with significant impact on national expenditure. England has, for example, invested at least £12.8 billion in a National Programme for Information Technology (NPfIT) for the National Health Service, and the Obama administration in the United States (US) has similarly committed to a US$33 billion eHealth investment in health care [1]. Such large-scale expenditure has been justified on the grounds that electronic health records (EHRs), picture archiving and communication systems (PACS), electronic prescribing (ePrescribing) and associated computerised provider (or physician) order entry systems (CPOE), and computerised decision support systems (CDSSs) will help address the problems of variable quality and safety in modern health care. However, the scientific basis of such claims—which are repeatedly made and seemingly uncritically accepted—remains to be established [2–7].

Moving this agenda forward thus requires a scientifically informed perspective. However, there remains a disparity between the evidence-based principles that underpin health care generally and the political, pragmatic, and commercial drivers of decision making in the commissioning of eHealth tools and services. Obtaining an evidence-informed perspective on the current situation may serve to ground unrealistic expectations that might hinder longer-term progress within the field, help to suggest priorities by identifying areas with greatest potential for benefit, and also inform ongoing deliberations on eHealth implementations that are being considered internationally.

To inform these global deliberations, we systematically reviewed the preexisting systematic review literature on eHealth technologies and their impact on the quality and safety of health care delivery. We synthesised and contextualised our findings with the broader theoretical and methodological literature with a view to producing a comprehensive and accessible overview of the field. We present here a synopsis and updated version of a much larger recently published report covering the period 1997–2010 [8].

Methods

Overview of Methods

Systematic reviews of reviews have been particularly advocated to inform policy, clinical, and research deliberations by providing an evidence-based summary of inter-related technologies [9]. Our approach involved drawing on established systematic review methodology (i.e., those developed by The Cochrane Collaboration) to ensure rigour by minimising the risk of bias [10]; we also drew on more novel methods of evidence synthesis (i.e., those developed by the UK National Health Service [NHS] Service Delivery and Organisation Programme) with the aim of producing an overview that we hoped would prove useful to decision makers [11]. We present here a summary of the methods used.

Developmental Work

Inherent difficulties associated with systematic reviews of health care organisation and delivery intervention include the considerable effort required at the outset to facilitate their conduct [9]. Accordingly, we began with an in-depth exploration of the fields of health care quality and safety, as well as eHealth functionalities used in health care delivery. This exploration entailed conceptually mapping the fields to understand various processes involved as well as how these relate to each other.

For quality and safety considerations, we identified existing taxonomies and frameworks to facilitate this conceptual mapping exercise, which helped to delineate the scope of our work. For the field of eHealth, we drew from existing team members’ conceptual and empirical work to aid our construction of a conceptual map for eHealth technologies [12,13]. This exercise allowed us to categorise interventions with regards to over-arching similarities. We characterised eHealth technologies as having three main overlapping functions: (1) to enable the storage, retrieval, and transmission of data; (2) to support clinical decision making; and (3) to facilitate remote care. Given the strategic focus of the English National Programme for Information Technology (NPfIT) (and other similar large-scale programmes) on electronic record and professional decision support systems [1], the first two functions were prioritised in this initial phase of our work. The current reported work thus concerns the related areas of EHRs, PACS, CPOEs, ePrescribing, and computerised systems for supporting clinical decision making. Remote care and consumer health informatics are the subjects of a subsequent 3-y research enquiry, which is currently in progress.

Selection and Critical Appraisal of Systematic Reviews

On the basis of the areas identified for prioritisation, we developed a detailed list of interventions that were to be included/excluded [Text S2]. End users of applicable interventions were limited to health care professionals; any findings relating to patient-focused interventions were therefore excluded. Of interest were systematic reviews that focused on the assessment of patient, practitioner, or organisational outcomes. We detailed the following methodological criteria for the identification of systematic reviews: (1) reference to the study as being a systematic review by the authors within the title, abstract, or text; and/or (2) evidence from the description of the methods that systematic review principles had been utilised in searching and appraising the evidence.

All systematic reviews having been identified as potentially suitable were assessed for inclusion by two independent reviewers, with arbitration by a third reviewer if necessary. Data from systematic reviews meeting the above criteria, henceforth referred to as “reviews,” were independently critically reviewed by two reviewers, and relevant data were abstracted. Systematic reviews not primarily concerned with assessing impact on patients, professionals, or the organisation, but nonetheless intervention focused, were drawn on to provide additional contextual information. These supplementary systematic reviews (henceforth referred to as “supplementary reviews”) were not subjected to formal critical appraisal.
Critical appraisal was undertaken using an adapted version of the Critical Appraisal Skills Programme (CASP) tool for systematic reviews [14]. These modifications were informed by the growing literature regarding both the methodological and reporting issues with primary research in health informatics (Table S1). The details of this process and the tool’s associated properties will be the subject of a separate publication in due course.

Data Synthesis
A standard approach was taken for each of the eHealth technologies of interest. Definitions were first clarified and then the individual use and broader scope for deployment conceptualised. Juxtaposing this with the aforementioned conceptual maps of the fields of eHealth, quality and safety provided a literature-based framework for delineating the principal theorised benefits and risks associated with each intervention. We used this framework to guide synthesis of the empirically demonstrated benefits and risks of implementing eHealth technologies.

The body of literature identified was too diverse to allow quantitative synthesis of empirical evidence and we therefore undertook a narrative synthesis. This synthesis involved initially describing the technologies and outcomes studies using the above-described framework for each of the included reviews, which was followed by developing a summary of our assessment of and the key findings from each review (Table S2). We then employed a modified version of the World Health Organization’s Health Evidence Network system for appraising public health evidence, which classifies evidence into three main categories, i.e., strong, moderate or weak; this assessment being based on a combination of the overall consistency, quality, and volume of evidence uncovered. These review-derived data were then thematically synthesised in relation to each of the technologies under consideration, drawing on key findings from the additional reviews, as appropriate [8].

Results
Our searches retrieved a total of 46,349 references from which we selected a total of 108 reviews for inclusion (Figure 1). Our final selection of 53 reviews provided the main empirical evidence base in relation to assessing the impact of the selected eHealth technologies (see Table 1 for our critical appraisal of these studies) [15–67], full details of which can be found in Table S2. An additional 55 supplementary reviews provided context to the findings [68–122], aiding in their interpretation [123]. In the case of systematic review updates, only the most recent review in a

![Figure 1. PRISMA flow diagram.](https://doi.org/10.1371/journal.pmed.1000387.g001)
Impact of eHealth on Quality and Safety

Table 1. Critical appraisal of ‘‘reviews’’ (see legend for description of quality assessment criteria).

Lead Author and Year

Q1

Q2

Q3

Q4

Q5

Q6

Q7

Q8

Q9

Q10

Q11

Q12

Q13

Q14

Q15

Totala

Ammenwerth 2008

2

2

2

1

2

2

2

1

2

2

1

2

1

1

1

24

Anderson 1997

2

2

2

1

2

1

1

2

1

2

1

2

1

1

2

23

Balas 2004

1

1

2

1

2

1

2

1

2

2

1

2

1

2

1

22

Bennett 2003

1

1

1

2

2

1

2

1

1

1

1

1

1

1

1

18

Bryan 2008

1

2

1

2

2

2

1

1

1

1

1

1

1

1

1

19

Charvet-Protat 1998

1

2

1

0

0

0

1

1

0

1

1

1

1

1

1

12

Chatellier 1998

2

2

2

0

1

1

2

2

2

2

1

1

1

1

1

21

Chaudhry 2006

1

1

2

2

0

1

1

1

2

1

0

2

0

2

2

18

Clamp 2005

1

1

2

0

1

1

1

1

2

1

1

1

1

2

1

17

Delpierre 2004

1

1

0

1

0

2

1

0

2

1

0

1

1

1

1

13

Dexheimer 2008

2

1

1

1

2

1

2

1

2

2

1

2

1

0

0

19

Durieux 2008

2

2

1

2

2

1

2

1

2

2

2

1

1

1

1

23

Eslami 2007

2

2

1

1

0

1

2

2

1

1

0

1

1

2

1

18

Eslami 2008

2

2

1

1

0

1

2

2

1

1

0

1

1

2

1

18

Eslami 2009

2

2

1

0

1

2

2

0

1

0

0

2

1

1

1

16

Fitzmaurice 1998

2

1

0

2

2

0

1

1

1

1

0

1

0

0

0

12

Garg 2005

1

1

2

2

2

1

1

1

2

2

1

1

1

1

1

20

Georgiou 2007

2

2

1

2

1

1

1

2

2

1

0

2

1

2

1

21

Hayward 2009

2

0

0

1

0

1

2

1

2

1

0

2

1

1

1

15

Hender 2000

1

1

2

1

2

1

1

1

1

0

0

0

0

0

0

11

Heselmans 2009

2

2

2

2

2

1

2

1

2

1

1

1

1

1

1

22

Hider 2002

1

1

2

2

1

2

1

1

2

2

1

2

1

1

1

21

Irani 2009

2

1

1

1

1

2

2

1

1

1

1

1

1

1

1

18

Jamal 2009

1

1

2

2

2

1

1

1

2

1

0

1

1

0

0

16

Jerant 2000

1

2

0

0

2

1

1

1

2

1

0

2

1

1

1

16

Kaushal 2003

2

2

2

1

1

1

1

2

0

2

1

1

0

0

0

16

Mador 2009

2

2

1

2

2

2

2

2

1

0

0

1

1

1

1

20

Mitchell 2001

1

1

1

2

2

1

1

1

0

1

1

1

1

2

1

17

Montgomery 1998

2

2

2

1

1

1

1

1

1

1

0

0

1

0

0

14

Niazkhani 2009

2

2

2

1

0

1

1

2

2

1

1

2

1

2

2

22

Oren 2003

1

2

1

0

0

1

1

1

0

1

0

1

1

2

2

14

Pearson 2009

2

2

2

1

2

2

2

2

2

1

1

1

1

1

1

23

Poissant 2005

2

1

2

2

2

1

2

1

2

2

2

2

1

2

1

25

Randell 2007

1

1

1

2

2

1

2

1

1

1

0

1

1

1

1

17

Reckmann 2009

2

2

1

2

0

2

1

1

1

1

0

2

1

2

2

20

Rothschild 2004

1

1

1

1

0

1

2

1

2

1

0

2

1

1

1

16

Schedlbauer 2009

1

2

2

1

2

2

2

1

2

1

1

2

1

1

1

22

Shachak 2009

2

1

1

0

0

0

2

1

1

1

1

2

1

1

1

15

Shamliyan 2008

1

2

1

1

1

2

0

1

1

1

1

1

0

0

0

13

Shebl 2007

1

1

2

1

1

1

1

1

2

1

1

1

1

1

1

17

Shekelle 2006

1

1

2

1

1

2

1

1

2

1

1

2

1

2

1

20

Shekelle 2009

1

1

2

1

0

2

1

1

2

1

1

2

2

2

2

21

Shiffman 1999

1

1

2

0

0

2

1

1

1

1

0

1

1

1

1

14

Shojania 2009

1

2

2

1

2

2

2

1

2

2

2

1

1

1

0

22

Sintchenko 2007

1

1

1

0

1

1

2

1

2

2

1

1

1

1

1

17

Smith 2007

1

2

0

1

1

2

1

1

1

1

0

2

1

2

1

17

Tan 2005

2

2

1

2

2

2

2

2

2

1

0

1

0

0

0

19

Thompson 2009

2

1

1

1

0

0

1

2

0

1

0

0

0

0

0

9

Uslu 2008

2

1

0

0

2

1

2

2

1

1

1

1

0

1

1

16

van Rosse 2009

2

2

1

1

2

2

2

0

1

2

2

2

1

2

2

24

PLoS Medicine | www.plosmedicine.org

4

January 2011 | Volume 8 | Issue 1 | e1000387


series of updates was selected. In the case of full and summary publications, we drew on the more substantive reports. Three related reviews – an update, a fuller report, and its more concise counterpart – were an exception due to the complementary nature of the reports rather than these being duplicative [22,55,56].

Data Storage, Management, and Retrieval Systems

Electronic health records. The EHR is a complex construct encompassing digitised health care records and the information systems into which these are embedded [8]. Whilst there are a number of operational definitions, the US’ Institute of Standards and Technology defines an EHR as “a longitudinal collection of patient-centric health care information available across providers, care settings, and time. It is a central component of an integrated health information system” [124]. EHRs can be used for the digital input, storage, display, retrieval, printing, and sharing of information contained in a patient’s health record [8]. We found that these systems vary on multiple dimensions, including levels of sophistication, detail, data source, timeframe (single service encounter to complete health record), and extent of integration (across intra- and interservice boundaries). In addition to patient histories and details of recent care, these records may also incorporate digital images and scanned documents. More detailed EHRs further often include nonclinical data relevant to health care administration and/or planning such as, for example, bed management and commissioning data. EHRs can therefore be used by a variety of end users such as clinicians, administrators, and patients themselves. EHRs can also have varying degrees of added clinical functionality including the ability to interface with a digital PACS, enter orders electronically (i.e., CPOE), prescribing (ePrescribing), and access to CDSSs.

The theoretical benefits and risks associated with EHRs are largely related to data storage and management functionality. These functions include increased accessibility, legibility, “searchability,” manipulation, transportation, sharing, and preservation of electronic data. Consequently, improved organisational efficiency and secondary uses of data are typically amongst the most commonly expected benefits. However, digitising health records can also introduce new risks. Paper persistence can result in threats to patient safety, unsecured networks can lead to illegitimate access, and increased time needed to document and retrieve patient data can result in organisational inefficiency. Moreover, the dynamic of the patient-provider interaction could become less personal with the intrusion by the computer as a “third person” in the consultation. If anticipated benefits are not realised, this may therefore mean that ultimately the EHR may be rendered cost-ineffective.

Although a number of reviews purporting to assess the impact of EHRs were found, many of these in fact investigated auxiliary systems such as CDSS, CPOE, and ePrescribing. As a result, most of the impacts assessed were more relevant to these other systems. We found only anecdotal evidence of the fundamental expected benefits and risks relating to the organisational efficiency resulting from the storage and management facilities within the EHR and thus the potential for secondary uses (Table 2). We did find, however, a small amount of secondary research relating to time efficiency for some health care professionals and administrators and data quality (in particular legibility, completeness, and comprehensiveness), which demonstrated weak evidence of benefit for both. Risks largely went ignored apart from anecdotal evidence of time-costs associated with recording of data due to both end-user skill and the inflexibility of structured data, increased costs of EHRs, and a decrease in patient-centeredness within the consultation (Table 3).

Picture archiving and communication systems. PACS are clinical information systems used for the acquisition, archival, and post-processing distribution of digital images. An image must either be directly acquired using digital radiography or be digitised from a paper-based format. It can be stored using an electronic, magnetic, or optical storage device. PACS can be integrated or interface with EHRs and CDSSs, or be stand-alone systems.

Much like the digitisation of health records, certain benefits – i.e., accessibility, image (rather than data) quality, searchability, transportation, sharing, and preservation – can be expected from the digitisation of medical images, which were previously film based. Again, certain improvements to organisational efficiency should in theory follow on from this digitisation, including time-savings, continuity of care, and ability to remotely view images. Conversely, digitising medical images can lead to decreased organisational efficiency if increased time is needed for retrieval owing to the difficulties associated with navigating a new or cumbersome system or in the event of system downtime. If the potential benefits of a PACS implementation are not realised, high expenditure might render the application cost-inefficient.

Although only three reviews on PACS were located, in contrast to the reviews on EHRs the impacts assessed in reviews of PACS were more congruent with the theoretically derived benefits (Table 4). This assessment involved a focus on improved organisational efficiency through time savings resulting from increased productivity of radiology services, reduced transit time, and improved access to new, recently stored, and archived images, as well as reducing physical space requirements for images; there was also an interest in the assessments of costs relating to purchasing and processing film. Worth noting however was the transient negative impact of implementation as well as issues with access due to system “loss” and downtime; access was sometimes impeded by the new workflows, which could result in a decrease in opportunistic interactions between clinicians and radiologists (Table 5). Overall, despite some promising findings, the weak evidence for the beneficial impact of digitising medical images is largely due to a low volume of research and somewhat inconsistent

Table 1. Cont.

<table>
<thead>
<tr>
<th>Lead Author and Year</th>
<th>Q1</th>
<th>Q2</th>
<th>Q3</th>
<th>Q4</th>
<th>Q5</th>
<th>Q6</th>
<th>Q7</th>
<th>Q8</th>
<th>Q9</th>
<th>Q10</th>
<th>Q11</th>
<th>Q12</th>
<th>Q13</th>
<th>Q14</th>
<th>Q15</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wolfscheid 2008</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Wong 2010</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>13</td>
</tr>
<tr>
<td>Yourman 2008</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>Average scores</td>
<td>1.49</td>
<td>1.47</td>
<td>1.3</td>
<td>1.09</td>
<td>1.19</td>
<td>1.28</td>
<td>1.45</td>
<td>1.19</td>
<td>1.43</td>
<td>1.21</td>
<td>0.7</td>
<td>1.32</td>
<td>0.87</td>
<td>1.08</td>
<td>0.91</td>
<td>18</td>
</tr>
</tbody>
</table>

Maximum total score of 30, each question (or Q) having a maximum of 2 points, refer to Table S1 for the critical appraisal form for additional details. doi:10.1371/journal.pmed.1000387.t001
findings across studies. For example, the overall cost-effectiveness of systems could not be determined, as the findings from economic analyses were often contradictory and of poor quality.

Supporting Clinical Decision Making

Computerised provider (or physician) order entry. CPOE systems are typically used by clinicians to enter, modify, review, and communicate orders; and return results for laboratory tests, radiological images, and referrals (for pharmacy see ePrescribing) [8]. These systems can be integrated within EHRs and/or integrate or interface with CDSSs. They not only integrate orders (similar to EHRs) with patient data and PACS images, but they also have the explicit purpose of electronic transfer of orders and the return of results. The electronic request of orders and return of results is expected to result in organisational efficiency gains and time savings. However, potential risks of these systems include increased time spent on computer-related activity and increased infrastructure costs, thereby decreasing overall organisational efficiency.

We found relatively few reviews on CPOE that were not focused primarily on the ordering of medications, rather than the ordering of laboratory tests and medical images. Within the reviews, we found that what had been empirically evaluated generally mirrored the theorised impacts (Tables 6 and 7). The findings from these reviews indicated weak evidence of an impact on organisational efficiency. Individual efficiency and workload both increased and decreased between providers. Additionally, while the speed at which orders were received led to better preparation and a modest effect on time taken to process and deliver results, it did not affect when the patient or their specimen was made available or when their results were acted upon. Findings supported moderate evidence of an impact on practitioner performance. The provision of relevant information at the time of ordering had a moderate impact on increasing cost-conscious ordering and subsequently on decreasing those orders deemed inappropriate; and following system-generated suggestions led to increased ordering of routine care as well as withdrawal of potentially injurious care. There was however evidence that the

### Table 2. Evidence of benefits associated with EHRs.

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Review ID</th>
<th>Data Security</th>
<th>Legibility</th>
<th>Accessibility</th>
<th>Completeness</th>
<th>Comprehensiveness</th>
<th>Efficiency</th>
<th>Secondary Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clamp 2005</td>
<td>N/A</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Irani 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Jamal 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Mador 2009</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Mitchell 2001</td>
<td>+/-</td>
<td>+</td>
<td>N/A</td>
<td>+</td>
<td>+</td>
<td>+/+</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Poissant 2005</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Shachak 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Shekelle 2006</td>
<td>N/A</td>
<td>+</td>
<td>N/A</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Shekelle 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Thompson 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Uslu 2008</td>
<td>N/A</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td></td>
</tr>
</tbody>
</table>

Evidence of benefits: N/A, not assessed; +/-, none; +, weak; +/-+, weak to moderate; ++, moderate.

doi:10.1371/journal.pmed.1000387.t002

### Table 3. Evidence of risks associated with EHRs.

<table>
<thead>
<tr>
<th>Risks</th>
<th>Review ID</th>
<th>Paper Persistence</th>
<th>Patient Disengagement</th>
<th>Insecure Data</th>
<th>Increased Time</th>
<th>Increased Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Clamp 2005</td>
<td>N/A</td>
<td>-</td>
<td>N/A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Irani 2009</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Jamal 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Mador 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Mitchell 2001</td>
<td>N/A</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Poissant 2005</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Shachak 2009</td>
<td>-</td>
<td>-/ -</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Shekelle 2006</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Shekelle 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Thompson 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Uslu 2008</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Evidence of risks: N/A, not assessed; +/-, none; -, weak; -/-, weak to moderate; --, moderate.

doi:10.1371/journal.pmed.1000387.t003
use of CPOE had a negative impact on practitioners because of the increased time needed to complete orders by having to enter them into the computer system, or incompatibility between professional routines and those imposed by the new system. Changes in workflows also posed an opportunity cost for collaboration, and the potential exclusion of certain providers from processes. Additionally, workload could either decrease or increase as a result of changes in workflow, which when unaccounted for were dealt with on an ad hoc basis and allowed for the redesignation of responsibilities.

**ePrescribing.** ePrescribing refers to clinical information systems that are used by clinicians to enter, modify, review, and output or communicate medication prescriptions. This term thus includes stand-alone CDSSs for prescribing purposes [8]. ePrescribing systems can integrate or interface with EHRs or be an element of a broader CPOE system. Like systems for computerised order entry, those for prescribing also have the explicit purpose of electronic transfer between the prescriber and the pharmacy and are rarely mentioned without decision support functionality [125]. ePrescribing systems should result in similar benefits as CPOE systems, including improvements in organisational efficiency and practitioner performance in relation to prescribing. Furthermore, the direct relationship between the therapeutic nature of prescribing of medications and patient outcomes suggests that better prescribing should lead to improved patient outcomes. Finally, as the prescribing of medications is a potentially larger contributor to risks to patient safety than the ordering of laboratory tests or radiology images, there is greater scope for improvements in patient safety by reducing errors in the prescribing process. On the contrary, a flawed or cumbersome system design (e.g., suboptimal specificity and/or sensitivity) and deployment strategies (e.g., insufficient training) may contribute to errors in prescribing and lead to workarounds, putting patients at risk and resulting in clinician dissatisfaction. Prescribers can also become over-reliant on decision support or overestimate its functionality, resulting in decreased practitioner performance.

ePrescribing was the most commonly studied intervention amongst the included reviews. Consequently, we found multiple papers covering most of the theorised impacts (Tables 8 and 9).

### Table 4. Evidence of benefits associated with PACS.

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Review ID</th>
<th>Data Integrity</th>
<th>Image Resolution</th>
<th>Image Access</th>
<th>Cost Savings</th>
<th>Time Savings</th>
<th>Diagnostic Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anderson 1997</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Charvet-Protat 1998</td>
<td>+</td>
<td>N/A</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Clamp 2005</td>
<td>+</td>
<td>N/A</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Evidence of benefits: N/A, not assessed; +/-, none; +, weak; ++, weak to moderate; ++, moderate.
doi:10.1371/journal.pmed.1000387.t004

### Table 5. Evidence of benefits associated with PACS.

<table>
<thead>
<tr>
<th>Risks</th>
<th>Review ID</th>
<th>Film Persistence</th>
<th>Record Loss</th>
<th>Increased Time</th>
<th>Increased Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Anderson 1997</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Charvet-Protat 1998</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Clamp 2005</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>-</td>
</tr>
</tbody>
</table>

Evidence of risks: N/A, not assessed; +/-, none; -, weak; +/-, weak to moderate; --, moderate.
doi:10.1371/journal.pmed.1000387.t005
systems should also be able to help address disparities in care by facilitating standardisation, especially when part of an EHR, PACS, CPOE, or ePrescribing system. Improved practitioner performance should result in a variety of beneficial impacts depending on the care activity targeted (e.g., increased immunisation rates, reduced resource utilisation, more timely diagnosis) or better disease control. In addition, if practitioner’s performance is directly related to patient outcomes, then these too should improve. The main theorised risks relating to the use of CDSSs include a potential decline in practitioner performance due to deskilling or flawed system design, and related threats to patient safety.

Actual improved practitioner performance rather than just behaviour change in general was supported by only weak evidence (Tables 10 and 11). While most findings were able to demonstrate some degree of behaviour change it did not always translate into the provision of higher quality care. While some subgroups seemed to fare better than others, the evidence was still only modest at best. The most notable findings were hallmarked by relative consistency across findings and thusly provided moderate evidence. These included increased provision of preventive care measures, disease-specific examinations or measurements, corollary orders to monitor side effects, and the decreased use of unnecessary or redundant care. Efforts at influencing practitioners to change practice patterns to adhere to a certain model of care were however less successful. No evidence was indicated for an impact on patient outcomes outside prescribing; while surrogate outcomes were modestly improved in some cases there was inconsistency across studies.

### Discussion

Our systematic review of systematic reviews on the impact of eHealth has demonstrated that many of the clinical claims made about the most commonly deployed eHealth technologies cannot be substantiated by the empirical evidence. Overall, the evidence base in support of these technologies is weak and inconsistent, which highlights the need for more considered claims, particularly in relation to the patient-level benefits, associated with these technologies. Also of note is that we found virtually no evidence in support of the cost-effectiveness claims (Tables 2–11) that are frequently being made by policy makers when constructing business cases to raise funding for the large-scale eHealth deployments that are now taking place in many parts of the world [1].

This work is characterised by a number of strengths and limitations, which need to be considered when interpreting this work. Strengths include the multifaceted approach to the identification of systematic reviews and the synthesis of this body of evidence. Juxtaposing the conceptual maps of the fields of quality, safety, and eHealth permitted us to produce a comprehensive framework for assessing the impact of these technologies in

---

### Table 6. Evidence of benefits associated with CPOE.

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Review ID</th>
<th>Resource Utilisation</th>
<th>Indicated Care</th>
<th>Patient Outcomes</th>
<th>Cost Savings</th>
<th>Time Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chaudry 2006</td>
<td>+/-</td>
<td>++</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Garg 2005</td>
<td>+/-</td>
<td>+++</td>
<td>+/-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Georgiou 2007</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Jamal 2009</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Niyazkhani 2009</td>
<td>N/A</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>+/- +</td>
</tr>
<tr>
<td></td>
<td>Poissant 2005</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/- +</td>
</tr>
<tr>
<td></td>
<td>Rothschild 2004</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Shellel 2006</td>
<td>+/-</td>
<td>+++</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Evidence of benefits: N/A, not assessed; +/-, none; +, weak; +/-, weak to moderate; ++, moderate.

doi:10.1371/journal.pmed.1000387.t006

### Table 7. Evidence of risks associated with CPOE.

<table>
<thead>
<tr>
<th>Risks</th>
<th>Review ID</th>
<th>Increased Time</th>
<th>Interruptions</th>
<th>Increased Costs</th>
<th>Workarounds</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Chaudry 2006</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Garg 2005</td>
<td>N/A</td>
<td>-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Georgiou 2007</td>
<td>-</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Jamal 2009</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Niyazkhani 2009</td>
<td>-</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Poissant 2005</td>
<td>-</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Rothschild 2004</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Shellel 2006</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Evidence of risks: N/A, not assessed; +/-, none; -, weak; +/-, weak to moderate; --, moderate.

doi:10.1371/journal.pmed.1000387.t007
an otherwise poorly ordered discipline. In addition, reflecting on methodological considerations and socio-technical factors enabled us to produce an overview that is sensitive to the intricacies of the methodological considerations and socio-technical factors enabled an otherwise poorly ordered discipline. In addition, reflecting on methodological considerations and socio-technical factors enabled us to produce an overview that is sensitive to the intricacies of the discipline.

Given the poor indexing of this literature and the fact that our searches were centred on English-language databases, there is the possibility that we may have missed some systematic reviews. Our use of a novel, multimethod approach may be criticised as being less rigorous than a conventional systematic review in that we were not in a position to appraise individual primary studies. These more novel methods of synthesis are less well developed and employed, and therefore less evaluated [126]. The fact that we needed to adapt the instrument used for critical appraisal is another potential limitation. Further, our assumptions about the theoretical benefits expected presumes that the eHealth technologies considered are capable of delivering these and are used in a manner that allows them to do so. Likewise, it could be argued that some of the expected benefits outlined in this overview are assured and perhaps do not therefore require formal evaluation. It is our view, based on the prevailing climate surrounding EHRs and large-scale implementations underway globally, that the claims made about these technologies are subjected to critical review in the light of the empirical evidence. The overlap in reviews and inconsistent use of terminology subjected to critical review in the light of the empirical evidence. The overlap in reviews and inconsistent use of terminology

<table>
<thead>
<tr>
<th>Reference ID</th>
<th>Surrogate Outcomes</th>
<th>Guideline Adherence</th>
<th>Safer Prescribing</th>
<th>Communication</th>
<th>Patient Outcomes</th>
<th>Resource/ Cost Savings</th>
<th>Time Savings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammenwerth 2008</td>
<td>Pot. ADEs +/- +</td>
<td>N/A</td>
<td>MEs +/+</td>
<td>N/A</td>
<td>ADEs +</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Bryan 2008</td>
<td>+/+</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+</td>
<td>N/A</td>
</tr>
<tr>
<td>Chattellier 1998</td>
<td>+/+/+</td>
<td>N/A</td>
<td>++/+</td>
<td>N/A</td>
<td>Death +/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Clamp 2005</td>
<td>++</td>
<td>+</td>
<td>MEs ++</td>
<td>+</td>
<td>ADEs +</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Delpiere 2004</td>
<td>+/-</td>
<td>+</td>
<td>MEs+</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Durieux 2008</td>
<td>+/-</td>
<td>+</td>
<td>++/+</td>
<td>N/A</td>
<td>Death +/-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td>Eslami 2007</td>
<td>+</td>
<td>+/-</td>
<td>+/+</td>
<td>N/A</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Eslami 2008</td>
<td>+</td>
<td>+/-</td>
<td>+/+</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Eslami 2009</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Fitzmaurice 1998</td>
<td>+</td>
<td>N/A</td>
<td>+</td>
<td>N/A</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Garg 2005</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td>Hider 2002</td>
<td>+/-</td>
<td>+</td>
<td>++/+</td>
<td>+</td>
<td>+/-</td>
<td>-</td>
<td>N/A</td>
</tr>
<tr>
<td>Jamal 2009</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>N/A</td>
<td>+/-</td>
<td>-</td>
<td>N/A</td>
</tr>
<tr>
<td>Mitchell 2001</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Molland 2009</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td>Niyazkhani 2009</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td>Poissant 2005</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td>Rothschild 2004</td>
<td>+/-</td>
<td>+</td>
<td>MEs ++</td>
<td>N/A</td>
<td>ADEs +</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Schedlbauer 2009</td>
<td>+/-</td>
<td>+</td>
<td>MEs ++</td>
<td>N/A</td>
<td>Renal ADEs+ Falls+</td>
<td>+</td>
<td>N/A</td>
</tr>
<tr>
<td>Shamiyani 2008</td>
<td>+</td>
<td>N/A</td>
<td>MEs ++</td>
<td>N/A</td>
<td>ADEs +</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Shekelle 2006</td>
<td>+/-</td>
<td>N/A</td>
<td>MEs +</td>
<td>N/A</td>
<td>ADEs +</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td>Shiffman 1999</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Shojania 2009</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Sintchenko 2007</td>
<td>++</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>Death +/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Tan 2005</td>
<td>+</td>
<td>N/A</td>
<td>MEs +</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td>Van Rosse 2009</td>
<td>+/-</td>
<td>N/A</td>
<td>++</td>
<td>+</td>
<td>Death +/-</td>
<td>ADEs +/-</td>
<td>N/A</td>
</tr>
<tr>
<td>Wolfstadt 2008</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>ADEs +/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Youman 2008</td>
<td>+</td>
<td>+</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Evidence of benefits: N/A, not assessed; +/-, none; +, weak; +/-+, weak to moderate; ++, moderate.

doi:10.1371/journal.pmed.1000387.t008
those oriented towards professionals. We are currently engaged in follow-on work, which broadens our field of enquiry along these lines [127–131]. Finally, our synthesis was limited by critical deficits within the literature, which undermined our efforts to generate a fully reproducible quantitative summary of findings [132].

At the most elementary level, the literature that constitutes the evidence base is poorly referenced within bibliographic databases reflecting the nonstandard usage of terminology and lack of consensus on a taxonomy relating to eHealth technologies [133–135]. There were, furthermore, varying degrees of overlap between individual reviews and contradictory findings even amongst reviews of the same primary studies. In addition, we found considerable heterogeneity in the ways in which findings and other aspects relating to the fundamental features of reviews (motivation, objectives, methods, presentation of findings, etc.) from individual papers were presented. This imprecision and nonstandard usage of terminology, as well as the poor quality of reviews, posed additional challenges, both with respect to interpretation of findings from individual reviews and in relation to synthesising the overall body of evidence.

Our greatest cause for concern was the weakness of the evidence base itself. A strong evidence base is characterised by quantity, quality, and consistency. Unfortunately, we found that the eHealth evidence base falls short in all of these respects. In addition, relative to the number of eHealth implementations that have taken place, the number of evaluations is comparatively small. Apart from several barriers and challenges that impede the evaluation of eHealth interventions per se [136–141], a number of factors might contribute to evaluative findings going unpublished [142]. Conflict of interests can, in particular, make it difficult to publish negative findings [142], which means that the potential for publication bias should not be underestimated in this discipline [102,143]. Moreover, published primary research has been repeatedly found to be of poor quality – particularly with regards to outcome measurement and analysis [73,74,80,86,118]. The highly heterogeneous and complex nature of these interventions makes consistency of findings, even across very similar scenarios, difficult to detect. Our critical appraisal exercise found the same to be true for secondary research. How the included reviews fared with regards to our critical appraisal, merits further comment and will be the subject of a further publication.

### Table 9. Evidence of risks associated with ePrescribing.

<table>
<thead>
<tr>
<th>Risks</th>
<th>Reference ID</th>
<th>Patient Harm</th>
<th>Increased Time</th>
<th>Increased Costs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ammenwerth 2008</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Bryan 2008</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Chatellier 1998</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Clamp 2005</td>
<td>+/-</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Delpierre 2004</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Durieux 2008</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Eslami 2007</td>
<td>+/-</td>
<td>-/-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Eslami 2008</td>
<td>+/-</td>
<td>-/-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Eslami 2009</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Fitzmaurice 1998</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Garg 2005</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Hider 2002</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Jamal 2009</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Mitchell 2001</td>
<td>+/-</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Mollon 2009</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Niyazkhani 2009</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Poissant 2005</td>
<td>N/A</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Rothschild 2004</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Schedlbauer 2009</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Shamliyan 2008</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Shekelle 2006</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Shiffman 1999</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Shojania 2009</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Sintchenko 2007</td>
<td>+/-</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Tan 2005</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Van Rosse 2009</td>
<td>+/-</td>
<td>+/-</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Wolfstadt 2008</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Yourman 2008</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Evidence of risks: N/A, not assessed; +/-, none; -, weak; +/- -, weak to moderate; --, moderate. doi:10.1371/journal.pmed.1000387.t009
Another commonly criticised element of the existing evidence base is its utility [144]. Evaluations have to date largely favoured simplistic approaches, which have provided little insight into why a particular outcome has occurred [145]. Understanding the underlying mechanisms, typically by studying the particular context of the evaluation, is critical for drawing conclusions in relation to causal pathways and effectiveness of eHealth interventions [146]. In addition, evaluations have tended to focus on the benefits with little attention to the risks and costs, which are rarely assessed or rigorously appraised [73,74,80,86,118]. Consequently, the existing evidence base is often of little utility to decision making in relation to the strategic direction of implementation efforts [144].

A handful of high-profile primary studies demonstrating the greatest evidence of benefit often serve as exemplars of the transformative power of clinical information systems [22]. These often include advanced multifunctional clinical information systems incorporating storage, retrieval, management, decision support, order and results communication, and viewing functionality. Evidence of the beneficial impact of such systems is limited, however, to a few academic clinical centres of excellence where the systems were developed in house, undergoing extensive evaluation with continual improvement, supported by a strong sense of local ownership by their clinical users [31,56]. The contrast between the success of these systems and the relative failure of much of the wider body of evidence is striking. Clearly, there are important lessons to be learned from these centres of excellence, but the extent to which the results of these primary studies can be generalised beyond their local environment to those institutions procuring “off-the-shelf” systems is questionable. It is encouraging, however, to see evaluations of commercial systems increasingly taking place [55].

A range of factors tend to contribute to the lack of successful implementations of these off-the-shelf systems. In particular, these commercial systems typically have assumptions about work

---

**Table 10. Evidence of benefits associated with CDSS.**

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Reference ID</th>
<th>Indicated Care</th>
<th>Guideline Adherence</th>
<th>Surrogate Outcomes</th>
<th>Patient Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Balas 2004</td>
<td>+</td>
<td>++</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Bryan 2008</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Chaudhry 2006</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Delpierre 2004</td>
<td>++</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Dexheimer, 2008</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Garg 2005</td>
<td>+/+ +</td>
<td>++</td>
<td>+/+ +</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Hayward 2009</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Heselmans 2009</td>
<td>N/A</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Jamal 2009</td>
<td>+/-</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Jerant, 2000</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Montgomery 1998</td>
<td>+</td>
<td>N/A</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Randell 2007</td>
<td>+/-</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Shekelle 2006</td>
<td>++</td>
<td>++</td>
<td>+/+ +</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Shiffman 1999</td>
<td>+</td>
<td>+/-</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Shojania 2009</td>
<td>+/+ +</td>
<td>++</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Sintchenko 2007</td>
<td>+</td>
<td>+</td>
<td>+/+ +</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Smith 2007</td>
<td>N/A</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Tan 2009</td>
<td>+/-</td>
<td>N/A</td>
<td>+/-</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Evidence of benefits: N/A, not assessed; +/-, none; +, weak; +/+, weak to moderate; ++, moderate.
doi:10.1371/journal.pmed.1000387.t010

**Table 11. Evidence of risks associated with CDSS.**

<table>
<thead>
<tr>
<th>Risks</th>
<th>Reference ID</th>
<th>Practitioner performance</th>
<th>Patient outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Balas 2004</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Bryan 2008</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Chaudhry 2006</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Delpierre 2004</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Dexheimer, 2008</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Garg 2005</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Hayward 2009</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Heselmans 2009</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>Jamal 2009</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Jerant, 2000</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Montgomery 1998</td>
<td>+/-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Randell 2007</td>
<td>-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Shekelle 2006</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Shiffman 1999</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Shojania 2009</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Sintchenko 2007</td>
<td>+/-</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Smith 2007</td>
<td>N/A</td>
<td>+/-</td>
</tr>
<tr>
<td></td>
<td>Tan 2009</td>
<td>+/-</td>
<td>+/-</td>
</tr>
</tbody>
</table>

Evidence of risks: N/A, not assessed; +/-, none; -, weak; +/-, weak to moderate; -, moderate.
doi:10.1371/journal.pmed.1000387.t011
practices embedded within them, which are often not easily transferable to different contexts of use. Additionally, it is not unusual for insufficient time and effort to be devoted to the all-important customisation process [147]. NHS Connecting for Health’s difficulties with the implementation of EHRs into hospitals in England is a prime example of the challenges that can ensue if such socio-technical factors are given insufficient attention [145].

Keeping in mind the above, the maturation of evaluation is vital to the success of eHealth [149,150]. There is some indication that the quality of evaluations is beginning to improve with regards to methodological rigour [74], but there is clearly still considerable scope for improvement [118]. Most of the reviews we included in our work made calls for more rigorous research to establish impact with some calling for more randomised controlled trials (RCTs) in particular [61,151]. A growing number of authors have however argued for trials of eHealth interventions to employ guidance specifically for complex interventions [152]. However, there are a number of challenges to conducting RCTs of eHealth [153], and many calls have also been made for using other complementary methodologies [24,146]. Strategies for improving the quality of research should include building the capacity and competency of researchers. In the shorter term, developing resources, tool-kits, frameworks, and the like for researchers and consumers of research should be prioritised [154–156]. Such developments are pivotal to furthering the science of evaluation in eHealth and the use of evidence-based principles in health informatics [157].

Another important development that is needed is the collaboration of different disciplines in evaluation [158,159].

We found an important literature pertaining to the design and deployment aspects of eHealth technologies. This literature is central to understanding why some interventions succeed and others fail (or being judged as such). At the individual level, “human factors” play an important role in the design of an intervention, determining usability and ultimately adoption [160]. At the aggregate level, “organisational issues” are critical in strategising deployment that ultimately influences adoption [160]. Although both enablers and barriers to success are being elicited retrospectively from the literature for design, development, and deployment, the findings for both of these concepts, inter-related as they are, have largely gone untested prospectively. Although there is greater attention being paid to the socio-technical aspects in formal evaluations than ever before, there is still much that needs to be understood [161].

Conclusions

It is clear that there is now a large volume of work studying the impact of eHealth on the quality and safety of health care. This might be seen as setting a firm foundation for realising the potential benefits of eHealth. However, although seminal reports on quality and safety of health care invariably point to eHealth as one of the main vehicles for driving forwards sweeping improvements [2–7], our work indicates that realising these benefits is not guaranteed and if it is to be achieved, this will require substantial research resources and effort.

Our major finding from reviewing the literature is that empirical evidence for the beneficial impact of most eHealth technologies is often absent or, at best, only modest. While absence of evidence does not equate with evidence of ineffectiveness, reports of negative consequences indicate that evaluation of risks—anticipated or otherwise—is essential. Clinical informatics should be no less concerned with safety and efficacy than the pharmaceutical industry. Given this, there is a pressing need for further evaluations before substantial sums of money are committed to large-scale national deployments under the auspices of improving health care quality and/or safety.

Promising technologies, unless properly evaluated with results fed back into development, might not “mature” to the extent that is needed to realise their potential when deployed in everyday clinical settings. The paradox is that while the number of eHealth technologies in health care is growing, we still have insufficient understanding of how and why such interventions do or do not work [123]. To resolve this, it is essential to not only devote more effort to evaluation, but to ensure that the methodology adopted is multidisciplinary and thus capable of untangling the often complex web of factors that may influence the results. Moreover, a fuller description of the rationale for the choice of methodological approach employed to evaluate eHealth technologies in health care would facilitate synthesis and comparison.

Finally, it is equally important that deployments already commissioned are subject to rigorous, multidisciplinary, and independent evaluations. In particular, we should take every opportunity to learn from the largest eHealth commissioning and deployment project in health care in the world—the £12.8 billion NPfIT and the at least equally ambitious national programme that has recently begun in the US [162–166]. These and similar initiatives being pursued in other parts of the world offer an unparalleled opportunity not just for improving health care systems, but also for learning how to (or how not to) implement eHealth systems and for refining these further once introduced.

Supporting Information

Table S1 Critical appraisal form. Found at: doi:10.1371/journal.pmed.1000387.s001 (0.05 MB DOC)

Table S2 Characteristics and main findings of “reviews.” Found at: doi:10.1371/journal.pmed.1000387.s002 (0.42 MB DOC)

Text S1 Search strategy (databases, string, and filters). Found at: doi:10.1371/journal.pmed.1000387.s003 (0.05 MB DOC)

Text S2 Intervention inclusion and exclusion criteria. Found at: doi:10.1371/journal.pmed.1000387.s004 (0.03 MB DOC)

Acknowledgments

We are grateful to the Independent Project Steering Committee comprising Denis Protti (chair), David Bates, Richard Lilford, Maureen Baker, Antony Clutter, and Jo Foster for their valuable guidance and support. Our many thanks to Uğurbek Nurmutav for his work in quality assessment as well as to Ann Hansen for her work in running the searches. This work draws on a report published by the NHS Connecting for Health Evaluation programme, the full text of which is available from: http://www.pcpoh.bham.ac.uk/publichealth/cfhep/documents/NHS_CFHEP_001_Final_Report.pdf

Author Contributions

ICMJE criteria for authorship read and met: ADB JC CP CA KC TB BM AM AS. Designed the experiments/the study: JC CP CA KC TB BM AM AS. Analyzed the data: ADB JC CA BM. Collected data/did experiments for the study: ADB JC CP CA TB. Wrote the first draft of the paper: ADB. Contributed to the writing of the paper: JC CP CA KC TB BM AM AS. Coapplicant on the grant that enabled this research to proceed; authored or coauthored key sections of the report (e.g., developed the conceptual maps, building on a previous analysis for NHS SDO; CP.


84. Gurses AP, Xiao Y (2006) A systematic review of the literature on


59. Sintchenko V, Magrabi F, Tipper S (2007) Are we measuring the right end-


56. Sintchenko V, Magrabi F, Tipper S (2007) Are we measuring the right end-


Impact of eHealth on Quality and Safety
**Editors’ Summary**

**Background.** There is considerable international interest in exploiting the potential of digital health care solutions, often referred to as eHealth—the use of information and communication technologies—to enhance the quality and safety of health care. Often accompanied by large costs, any large-scale expenditure on eHealth—such as electronic health records, picture archiving and communication systems, ePrescribing, associated computerized provider order entry systems, and computerized decision support systems—has tended to be justified on the grounds that these are efficient and cost-effective means for improving health care. In 2005, the World Health Assembly passed an eHealth resolution (WHA 58.28) that acknowledged, “eHealth is the cost-effective and secure use of information and communications technologies in support of health and health-related fields, including health-care services, health surveillance, health literature, and health education, knowledge and research,” and urged member states to develop and implement eHealth technologies. Since then, implementing eHealth technologies has become a main priority for many countries. For example, England has invested at least £12.8 billion in a National Programme for Information Technology for the National Health Service, and the Obama administration in the United States has committed to a US$38 billion eHealth investment in health care.

**Why Was This Study Done?** Despite the wide endorsement of and support for eHealth, the scientific basis of its benefits—which are repeatedly made and often uncritically accepted—remains to be firmly established. A robust evidence-based perspective on the advantages on eHealth could help to suggest priority areas that have the greatest potential for benefit to patients and also to inform international eHealth deliberations on costs. Therefore, in order to better inform the international community, the authors systematically reviewed the published systematic review literature on eHealth technologies and evaluated the impact of these technologies on the quality and safety of health care delivery.

**What Did the Researchers Do and Find?** The researchers divided eHealth technologies into three main categories: (1) storing, managing, and transmitting data; (2) clinical decision support; and (3) facilitating care from a distance. Then, implementing methods based on those developed by the Cochrane Collaboration and the NHS Service Delivery and Organisation Programme, the researchers used detailed search strategies and maps of health care quality, safety, and eHealth interventions to identify relevant systematic reviews (and related theoretical, methodological, and technical material) published between 1997 and 2010. Using these techniques, the researchers retrieved a total of 46,349 references from which they identified 108 reviews. The 53 reviews that the researchers finally selected (and critically reviewed) provided the main evidence base for assessing the impact of eHealth technologies in the three categories selected.

In their systematic review of systematic reviews, the researchers included electronic health records and picture archiving communications systems in their evaluation of category 1, computerized provider (or physician) order entry and e-prescribing in category 2, and all clinical information systems that, when used in the context of eHealth technologies, integrate clinical and demographic patient information to support clinician decision making in category 3. The researchers found that many of the clinical claims made about the most commonly used eHealth technologies were not substantiated by empirical evidence. The evidence base in support of eHealth technologies was weak and inconsistent and importantly, there was insufficient evidence to support the cost-effectiveness of these technologies. For example, the researchers only found limited evidence that some of the many presumed benefits could be realized; importantly, they also found some evidence that introducing these new technologies may on occasions also generate new risks such as prescribers becoming overly-reliant on clinical decision support for e-prescribing, or overestimate its functionality, resulting in decreased practitioner performance.

**What Do These Findings Mean?** The researchers found that despite the wide support for eHealth technologies and the frequently made claims by policy makers when constructing business cases to raise funds for large-scale eHealth projects, there is as yet relatively little empirical evidence to substantiate many of the claims made about eHealth technologies. In addition, even for the eHealth technology tools that have proven to be successful, there is little evidence to show that such tools would continue to be successful beyond the contexts in which they were originally developed. Therefore, in light of the lack of evidence in relation to improvements in patient outcomes, as well as the lack of evidence on their cost-effectiveness, the authors say that future eHealth technologies should be evaluated against a comprehensive set of measures, ideally throughout all stages of the technology’s life cycle, and include socio-technical factors to maximize the likelihood of successful implementation and adoption in a given context. Furthermore, it is equally important that eHealth projects that have already been commissioned are subject to rigorous, multidisciplinary, and independent evaluation.

**Additional Information.** Please access these websites via the online version of this summary at http://dx.doi.org/10.1371/journal.pmed.1000387.

- More information is available on the World Health Assembly eHealth resolution
- The World Health Organization provides information at the Global Observatory on eHealth, as well as a global insight into eHealth developments
- The European Commission provides Information on eHealth in Europe and some examples of good eHealth practice
- More information is provided on NHS Connecting for Health
“Clinical features of women with gout arthritis.”
A systematic review

KJM Jansen Dirken-Heukensfeldt · TAM Teunissen ·
EH van de Lisdonk · ALM Lagro-Janssen

Abstract Clinically, gout is generally considered as a preferential male disease. However, it definitely does not occur exclusively in males. Our aim was to assess differences in the clinical features of gout arthritis between female and male patients. Five electronic databases were searched to identify relevant original studies published between 1977 and 2007. The included studies had to focus on adult patients with primary gout arthritis and on sex differences in clinical features. Two reviewers independently assessed eligibility and quality of the studies. Out of 355 articles, 14 were selected. Nine fulfilled the quality and score criteria. We identified the following sex differences in the clinical features of gout in women compared to men: the onset of gout occurs at a higher age, more comorbidity with hypertension or renal insufficiency, more often use of diuretics, less likely to drink alcohol, less often podagra but more often involvement of other joints, less frequent recurrent attacks. We found interesting sex differences regarding the clinical features of patients with gout arthritis. To diagnose gout in women, knowledge of these differences is essential, and more research is needed to understand and explain the differences, especially in the general population.

Keywords Female · Gender · Gout · Sex characteristics · Sex differences

Introduction

Gout is a frequent form of arthritis, presenting as a severe and painful inflammatory arthritis, mostly of the first metatarsophalangeal joint (podagra). It occurs suddenly, and in most patients it disappears completely within 5 to 14 days [1].

Gout affects around 1% of adult men in Western countries, mostly aged over 45 years. The estimated incidence of gout in these countries is 0.6 to 2.1 per 1,000 per year, with a prevalence of 3 to 7.5 per 1,000 per year [2–4]. Accumulating data support an increase in the prevalence of gout that is potentially attributable to recent shifts in diet, lifestyle, medical care, and increased longevity. The increased longevity of the population in industrialized countries may contribute to a higher prevalence of gout through the disorder’s association with age-related diseases such as hypertension and cardiovascular diseases. The increasing prevalence of gout worldwide indicates a need for improved effort to identify these patients early in the disease process [5].

Clinically, gout is often considered a preferential male disease. The condition is more common in men than in women. Hippocrates was the first who stated that it was a sex-related disease. Among patients younger than 65 years, men have a fourfold greater prevalence than women [6]. However, the incidence of gout in the elderly has a more equal sex distribution. The impact of female hormones cannot be neglected because estradiol can lower serum urate in females, and the serum urate rises after the menopause [7]. In patients older than 65 years, the sex gap narrows to one woman to every three men with gout and/or hyperuricemia [5]. Research about gout arthritis in the general population is rare; even less is
known about gout in female patients in general population. The aim of this review is to evaluate the available studies on the sex differences in clinical features of gout arthritis. Therefore, our research question was: What is known in medical literature about the clinical features of gout in females compared to males?

Methods

The present review was conducted in cooperation with a trained librarian (EP), according to the methodology of the Effective Practice and Organization of Care module from the Cochrane library [8]. Electronic searches were undertaken in MEDLINE, EMBASE, the Cochrane Clinical Trials Register, and Web of Science. The search strategy aimed to identify original relevant research papers published between 1977 and 2007 on gout and sex differences. It consisted of the AND combination of two main concepts: gout and the gender-specific filter. Meaning: Gout AND sex characteristics OR sex distribution OR sex factors OR sex differences OR gender identity OR gender OR sex distribution OR women’s health OR menopause OR pregnancy OR breast feeding OR menstruation OR gonadal hormones OR men’s health—in Mesh headings and text words [9].

The strategy used to identify relevant articles was adapted to the specific search criteria required for each database. Hyperuricemia can be present in a patient who does not have acute gout arthritis, and a patient, however seldom, can have gout without hyperuricemia. For this reason, we left the term hyperuricemia out of the search strategy. No age limit was used. Language was limited to Dutch, English, German, and French. We started the selection in 1977 because we used the Wallace gout criteria defined in 1977, later called the American College of Rheumatology (ACR) criteria of the American Rheumatology Association [10]. Additional studies were identified by searching the reference lists of relevant trials and reviews. Details of the database searches can be obtained from the author.

Selection

To select studies for further assessment, two observers (KD and TT) have independently reviewed the titles, abstracts, and keywords of every record retrieved. Articles were selected if the information from title, abstract, or keywords included gout not caused by another underlying disease and if the article reported an explicit comparison of male and female clinical features and a clear description about how they made the diagnosis of gout. If the information of title, abstract section, or keywords was unclear, the full article was retrieved for clarification. If no consensus was reached about whether to include a study, a third observer (ALJ) was asked to make a definite decision. Interobserver agreement for study selection was measured using the kappa statistic.

The two reviewers assessed the quality of each retrieved trial independently, using the “Quality criteria list on cohort studies” (CBO 2003: http://www.cbo.nl/product/richtlijnen/handleiding_ebro/article20050427141202/articleCBObasic_view). In this way, we checked whether the study met the international guideline criteria. We checked the full description of the stated aim of the study, study population, patient characteristics such as ethnicity, socioeconomic status, and setting (inpatient or outpatient, general practice), definition of the study inclusion and exclusion criteria, observation period, outcome measurements, statistic analysis, and data presentation. If any of this information was missing, the article was excluded.

Quality assessment of selected articles

The included full articles were graded and given a quality score based on the following information: method of diagnosing gout, female versus male comparisons, the absolute number and the age of the female patients, and the type of population and adequate outcome measures (Table 1). The diagnosis of gout was considered accurate if there was documentation on urate crystals aspirated from an affected joint or, failing this, if the clinical characteristics recorded were sufficient to fulfill the ACR diagnostic criteria of gout [10]. We also included studies in which the diagnosed gout was based on a dramatic clinical improvement of the arthritis as response to a treatment with colchicine.

Results

From MEDLINE, we retrieved 178 original articles. The two observers (KD, TT) selected 69 articles on title and abstract (kappa 0.57). From EMBASE and the Cochrane Library, we retrieved another 20 articles. In the Web of Science, we searched for articles quoted by the authors of our top ten articles. No new articles were found. Out of the 89 full articles, we selected 14 original studies with specific information on sex differences of gout and a clear description about how they made the diagnosis of gout. These articles were scored according to the quality criteria mentioned in Table 1. The maximum score was 5. The minimum score for inclusion was set on 2.5. Nine articles were included, and five were excluded (Table 2) [11, 15].

Two of the nine articles were conducted in a general population and seven in a hospital. The age of the male patients ranged from 53 to 70 years and of the female
patients from 49 to 61 years. The ethnicity of the study population differed in the studies. In two studies [16, 17], the diagnosis of gout was made by synovial tap, in one by using both synovial tap and the ACR criteria [18], in four by the ACR criteria [19–22], and in two by ACR criteria or clinical improvement on treatment with colchicines [23]. The best way to classify a patient as having acute gout is to demonstrate characteristic sodium urate monohydrate crystals in the joint fluid. If the diagnosis is based on colchicine or ACR criteria, other causes of arthritis were also included in the studies. Because of this, we reviewed the studies with crystal-proven gout separately from the studies using the ACR criteria and the colchicine definition.

The diagnosis of gout based on synovial fluid analysis

Two studies are based on a crystal-proven gout [16, 17]. In the study of Puig et al., in 89% of the study population, the gout was crystal-proven (n=40) [18]. Because of the high percentage, we include this study also in the group “crystal-proven gout” (Table 3).

Patient characteristics

The age of onset of gout as described in the three studies was on average 8.1 years later in women, and the duration of the gout at study entry was significantly higher in men. Gallerani et al. only studied the influence of the season on gout attacks and showed a significant peak of gout attacks in April in male patients, 36% of all the attacks [17]. Most of the women were postmenopausal at the onset of gout (86% and 92%).

Considering the use of diuretics, female gout patients received diuretics significantly more often than male patients, 57% of the women vs 14% of the men [18] and 83% of the women vs 47% of the men [16]. In the same studies, men with gout were more likely to drink alcohol than women, 14% of the women vs 55% of the men [18] and 10% of the women vs 45% of the men [16]. Puig et al. also analyzed sex differences in obesity in gout patients and found no differences between men and women.

Joint location

Puig et al. and Lally et al also described the location of the arthritis [16, 18]. Men seemed to have a higher prevalence of podagra at their first attack, 52% of the women vs 57% of the men [16]. Puig et al. found significantly more tophi in female gout patients compared to male, 27% in women vs 10% in men [18]. Lally et al. found a higher prevalence of polyarticular gout in male patients, 56% of the women vs 80% of the men [16]. They found no sex differences in

Table 1  Score of the quality criteria

<table>
<thead>
<tr>
<th>Prescription</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Method of proven gout</td>
<td>Synovial puncture</td>
</tr>
<tr>
<td>ACR criteria</td>
<td>0.5</td>
</tr>
<tr>
<td>Diagnosis on specifically defined clinical grounds</td>
<td>0.0</td>
</tr>
<tr>
<td>Male/female comparison</td>
<td>Specific numbers</td>
</tr>
<tr>
<td>Global, prevalence, incidence</td>
<td>0.5</td>
</tr>
<tr>
<td>Number of female patients</td>
<td>&gt;26</td>
</tr>
<tr>
<td>11–25</td>
<td>0.5</td>
</tr>
<tr>
<td>0–10</td>
<td>0.0</td>
</tr>
<tr>
<td>Study population</td>
<td>General population</td>
</tr>
<tr>
<td>Primary care</td>
<td>0.5</td>
</tr>
<tr>
<td>Hospital</td>
<td>0</td>
</tr>
<tr>
<td>Statistics</td>
<td>Odds ratio</td>
</tr>
<tr>
<td>Only prevalence/incidence</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Table 2  Selected studies with a score of less than 2.5 points

<table>
<thead>
<tr>
<th>Year</th>
<th>Score</th>
<th>Female</th>
<th>Male</th>
<th>Age of onset</th>
<th>Population</th>
<th>Diagnosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ter Borg et al. [12]</td>
<td>1987</td>
<td>2.0</td>
<td>22</td>
<td>18</td>
<td>F+3.0</td>
<td>Hosp</td>
</tr>
<tr>
<td>Cassim et al. [13]</td>
<td>1994</td>
<td>2.0</td>
<td>14</td>
<td>93</td>
<td>F+7.6</td>
<td>Hosp</td>
</tr>
<tr>
<td>Kowalec et al. [14]</td>
<td>1978</td>
<td>1.5</td>
<td>9</td>
<td>33</td>
<td>F+1.0</td>
<td>Hosp</td>
</tr>
<tr>
<td>Macfarlane et al. [15]</td>
<td>1985</td>
<td>1.0</td>
<td>9</td>
<td>51</td>
<td>F+21.0</td>
<td>Hosp</td>
</tr>
</tbody>
</table>
tophaceous gout, but in their study the upper limb seemed to be more involved in male patients, 44% in women vs 47% in men [16].

**Comorbidities**

Finally, Puig et al. and Lally et al. reported the incidence of “gout-associated comorbidities” as diabetes, hypertension, and renal insufficiency [16, 18]. In both studies, an association was found between renal insufficiency and postmenopausal women, 54% of the women vs 11% of the men [18] and 30% of the women vs 12% in men [16]. Puig et al. considered that hypertension was more common in female than in male gout patients, 78% of the women vs 14% of the men [18]. The sex distribution on diabetes and gout was equal [18].

Gout according to the ACR criteria or clinical improvement on colchicine

Six studies did not base the diagnosis on a crystal-proven gout (Table 4).
Patient characteristics

The age of onset of gout was on average 9.2 years later in women (mean age 66 years) compared to men (mean age 54 years), and the duration of the gout at study entry was significantly higher in men. Most of the women were postmenopausal at the onset of the gout (66–95%). Three studies described the use of diuretics [19–21]. In two studies, female gout patients received diuretics significantly more often than male patients, 77% in women vs 40% in men [19] and 72% in women vs 48% in men [20]. Three studies reported the use of alcohol in gout patients [20–22].

### Table 4 Sex differences according to the six selected articles with diagnose gout made by ACR criteria

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Female %</td>
<td>5</td>
<td>2.5</td>
<td>3.5</td>
<td>2.5</td>
<td>2.5</td>
<td>3.5</td>
</tr>
<tr>
<td>Male (n)</td>
<td>1,185</td>
<td>92</td>
<td>27</td>
<td>27</td>
<td>22</td>
<td>21</td>
</tr>
<tr>
<td>Average age of women (years)</td>
<td>70*</td>
<td>61</td>
<td>64</td>
<td>63</td>
<td>59</td>
<td>55</td>
</tr>
<tr>
<td>Average age of men (years)</td>
<td>58</td>
<td>51</td>
<td>61</td>
<td>49</td>
<td>52</td>
<td>54</td>
</tr>
<tr>
<td>Age of onset</td>
<td>F+12**</td>
<td>F+10</td>
<td>F+8*</td>
<td>F+9*</td>
<td>F+8*</td>
<td>F+4</td>
</tr>
<tr>
<td>Duration</td>
<td>–</td>
<td>–</td>
<td>M</td>
<td>–</td>
<td>–</td>
<td>M</td>
</tr>
</tbody>
</table>

### Patients characteristics (% women–% men)

- **Diuretics**: F** (77–40) M* (72–48) F – –
- **Alcohol**: – M* (2–20) M (7–16) – M* (55–82)
- **Obesity**: – – F (40–26) – –
- **Familiar gout**: – – F (40–26) – –
- **Spring peak**: – – – – –
- **Recurrence**: M** – – – –

### Location on onset (% women–% men)

- **Podagra**: – – M – M* (23–45) –
- **Ankle**: – – – – F* (45–23) –
- **Upper limb**: – – F** – – –
- **Fingers**: – – – – E F*
- **Monoarthritides**: – M** (27–57) – – M –
- **Polyarthritides**: – F (63–39) – F (41–24) –

### Comorbidity (% women–% men)

- **Dyslipemia**: F** (42–38) E F – – –
- **Hypercholesterol.**: E F F – –
- **Chr. heart disease**: F** (25–19) M (26–57) F (25–16) – M (0–11) –
- **Peripheral arterial disease**: F** (7–4) – – –
- **Diabetes**: F** (30–17) E M – F E

*Syn. tap* diagnosed on a synovial tap, *ACR* diagnosed according to the ACR criteria of gout, *Postmenopausal* number of postmenopausal women, *Duration* the length of the period since first gout diagnosis, *GP* general practice, *F* more prevalent in female, *M* more prevalent in men, *E* equal between men and women

*p<0.01 (significant); **p<0.001 (significant)*
They found that men with gout were more likely to drink alcohol than women, 2% of the women vs 20% of the men [20], 7% of the women vs 16% of the men [21], and 55% of the women vs 82% of the men. Tikley et al. found sex differences in the relation between obesity and gout. In women, they did not find any relation between the body mass index (BMI) and gout, but they found a higher prevalence of gout in men with a BMI > 25 (odds ratio 7.8) [22]. Harrold et al. studied the recurrence of gout attacks and reported a significantly higher number in men [19].

Gout in the family history was equally distributed between the sexes [21].

**Joint location**

The studies of De Souza et al. and Deesomchok et al. described the location of the arthritis [21, 23]. Men had a higher prevalence of podagra at their first attack, 23% of the women vs 45% of the men [23]. During the gout disease period, the sex differences in the location of the recurrent attack widened, with an increasingly higher prevalence of podagra in men [21]. Female patient more often had other joints involved, such as ankle, fingers, and upper limb [23]. Polycrystalline gout seemed to be more related to female gout patients, 63% of the women vs 39% of the men [20] and 41% of the women vs 24% of the men [23]. The presence of tophi differed in the various studies, 39% of the women vs 26% of the men [20], 30% of the women vs 48% of the men [21], 18% of the women vs 31% of the men [24], and 34% of the women vs 18% of the men [23].

**Comorbidities**

Five studies reported the incidence of several “gout-associated comorbidities” [19–23]. In four of them, hypertension was more common in female than in male gout patients, 81% of the women vs 57% of the men [19], 81% of the women vs 77% of the men [21], 45% of the women vs 39% of the men [23], and 65% of the women vs 59% of the men [22]. In two studies, this difference was significant [19, 22]. Five studies considered diabetes [19–23]. In one study, diabetes was significantly more prevalent among female than among male gout patients, 30% of the women vs 17% of the men [19]. The other four were contradicting.

Four studies analyzed the relationship between gout and cardiovascular heart disease. These results were also inconsistent, 25% of the women vs 19% of the men [19], 26% of the women vs 57% of the men [20], 25% of the women vs 16% of the men [21], and 0% of the women vs 11% of the men [23]. Dyslipemia was more common in women (42% vs 38%) [18–21] and Deesomchochok et al. did not find a sex difference in cerebral vascular accident in gout patients [23]. In this study, a significantly higher prevalence of hematologic malignancies was found in female gout patients, 22% of the women compared to 3% of the men [23].

Renal insufficiency was studied in five studies, and in all an association was found with gouty arthritis in postmenopausal women (in three of them, the association was significant), especially in those with preexisting joint disease, 18% of the women vs 10% of the men [19], 25% of the women vs 15% of the men [20], and 22% of the women vs 12% of the men [24].

We studied the conclusions of the five excluded studies and compared these with the nine included papers. No significant difference in outcome variables was found between the two groups.

**Discussion**

We selected nine articles which varied largely in the characteristics of the study population such as a hospital population, outpatient department or general population, the number of postmenopausal women, age, and the ratio of women and men. Moreover, the method of classifying gout varied from using the ACR criteria to detecting urate crystals. These factors influence the clinical features and make it difficult to compare the results. Therefore, we divided the studies in one group diagnosed on a synovial fluid analysis and another group diagnosed on the ACR/collchicine criteria. Nevertheless, there are great similarities in results in both groups. Our conclusions are that women are almost a decade older at the onset of gout arthritis, have more associated comorbidities such as hypertension and renal insufficiency, and use less alcohol. Also, the typical location in the first toe is less frequent in women, and women are more likely to be taking diuretics. In the crystal-proven gout group, polycrystalline gout tends to be more prevalent in male while in the ACR group polyarthritis seems to be related to women.

The difference in age between men and women at the onset of gout is remarkable. One of the reasons is possibly the menopause. After the menopause, the incidence of gout is high compared to the reproductive age [16, 18, 23, 24]. Kim et al. suggested a possible role of 17-beta-estradiol in the regulation of purine biosynthesis and uric acid metabolism and lowered serum urate [7]. Also, Hak et al. found in a recent large prospective study that menopause increases the risk of gout [25].

Women with gout have more often renal insufficiency and hypertension, and they more often use diuretics. Is it because of the older age of female gout patients that they had a higher percentage of renal dysfunction? Lally et al. and Bero et al. controlled their study results for age at the
gout onset, and after this correction they still found a higher prevalence of renal insufficiency and diuretic use in female gout patients [4, 8, 16]. Renal dysfunction reduces urate excretion, and hyperuricemia has been found to increase tubular reabsorption of sodium and may thus be predisposing for hypertension and diuretic use. Conversely, hypertension can induce renal dysfunction [26]. Women in general use diuretics much more frequently than men: a causal relationship between gout and diuretics is therefore unclear.

Alcohol increased the risk of gout but plays a less important role in the development of gout in women. Alcohol causes hyperuricemia by reducing renal excretion of urate as a result of increased adine nucleotide turnover.

Female gout patients differ in the location of the gout arthritis. In women, not only podagra is involved but also other joints such as fingers and ankle. Therefore, gout should be considered seriously in the differential diagnosis of elderly women with an acute (oligo)arthritis, especially of the ankle. These atypical locations may cause a delay in the diagnosis because of unfamiliarity of physicians with the atypical presentation of gout in women. Especially since a function of monosodium urate crystals in joint fluid is seldom used in primary care in the diagnostic procedure.

Another reason for a delay in the diagnosis in women may be the severity of coexisting diseases because of their older age compared to men, which overshadows gout. Studies in which the arthritis location is described in detail are rare. It seems that gout in female patients has fewer recurrences.

Harrold et al. performed the first population-based study with a high number of female patients [19]. In this study, female gout was more associated with hypertension, dyslipidemia, chronic heart disease, peripheral arterial disease, diabetes, and renal insufficiency compared to male patients with gout. The other studies examining sex differences in clinical characteristics of gout were very small studies, and only the study of Chang et al. was population-based [24].

Strengths and limitations of this study

This is, as far as we know, the first systematic review about clinical features of women with gout arthritis.

This study has some limitations. Firstly, most studies are based on physician diagnosis by using the ACR criteria or the response to colchicine, a definition that is subject to misclassification. The unreliable diagnosis of gout is a problem. If we had strictly accepted the gout diagnosis on the basis of urate crystals, only two studies would have been included [16, 17]. Malik et al. found a positive predictive value of 66% for ACR criteria compared to the golden standard crystal identification by synovial tap [27]. So, up to a third of the cases included in these studies might have false positives. Despite this problem, the results in the crystal-proven gout studies were quite equal to the studies using the ACR criteria, and in daily practice the diagnosis of gouty arthritis is based on clinical grounds without the use of crystal identification in de synovial fluid.

Another problem is the small number of women in most studies. Only the study of Harrold et al. performed a population-based study with a high number of female patients [19]. In all other included studies, the total number of included patients, especially women, is much too low to explore sex differences in gout-related comorbidities. Besides, most studies are performed in a clinical setting and depend therefore on presented morbidity. Only the studies of Harrold et al. and Chang et al. were population-based [15, 18].

Conclusion

This review is a unique study and a start for further research, in order to realize evidence-based diagnostic procedures and treatments for female patients with gout arthritis in general population.

Our systematic review about sex differences in gout arthritis draws attention to some of the less-known characteristics of gout. Conclusions must be drawn with great caution because of the small number of available studies on female gout, which however were carried out in different populations, with different criteria and mostly small numbers of women.

We found differences in the clinical features of gout arthritis between men and women. The onset of gout occurs at a later age in women; they are more likely to have comorbidities such as hypertension or renal insufficiency, and they use diuretics more often. Men with gout are more likely to drink alcohol, and they have a higher prevalence of typical podagra than female patients. Women may more often have other joints involved than just one toe, and the gout recurs less often.

More research in general population is necessary to identify sex differences in clinical features, in order to avoid a possible delay in diagnosing gout and recognize patients who are at risk for developing gout. We need population-based studies with a large number of female patients. In new research, more attention also have to be given to the difference in first recurrent episodes of gout. Do clinical features differ in these two groups?

Because gout is a rare disease in women and it can have an unusual way of manifesting itself, it is very important to recognize the symptoms. Physicians should consider gout arthritis especially in elderly female patients with hypertension, diuretic use, and renal insufficiency and arthritis in one or more joints. It is easy to overlook or pay little attention to the chronic joint problems because the female patients are older and frequently have chronic diseases. The
recognition of gout even with an atypical course is important because there are therapeutic strategies to reduce recurring unnecessary pain and complications like tophi deformities, nephropathy, and joint destruction.

Acknowledgement Elmie Peters, Librarian UMCN, St. Radboud.

Disclosures None.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Online Medical Professionalism: Patient and Public Relationships: Policy Statement From the American College of Physicians and the Federation of State Medical Boards

Jeanne M. Farnan, MD, MHPE; Lois Snyder Sulmasy, JD; Brooke K. Worster, MD; Humayun J. Chaudhry, DO, MS, SM; Janelle A. Rhyne, MD, MA; and Vineet M. Arora, MD, MAPP, for the American College of Physicians Ethics, Professionalism and Human Rights Committee; the American College of Physicians Council of Associates; and the Federation of State Medical Boards Special Committee on Ethics and Professionalism*

User-created content and communications on Web-based applications, such as networking sites, media sharing sites, or blog platforms, have dramatically increased in popularity over the past several years, but there has been little policy or guidance on the best practices to inform standards for the professional conduct of physicians in the digital environment. Areas of specific concern include the use of such media for nonclinical purposes, implications for confidentiality, the use of social media in patient education, and how all of this affects the public’s trust in physicians as patient–physician interactions extend into the digital environment. Opportunities afforded by online applications represent a new frontier in medicine as physicians and patients become more connected. This position paper from the American College of Physicians and the Federation of State Medical Boards examines and provides recommendations about the influence of social media on the patient–physician relationship, the role of these media in public perception of physician behaviors, and strategies for physician–physician communication that preserve confidentiality while best using these technologies.


* This paper, written by Jeanne M. Farnan, MD, MHPE; Lois Snyder Sulmasy, JD; Brooke K. Worster, MD; Humayun J. Chaudhry, DO, MS, SM; Janelle A. Rhyne, MD, MA; and Vineet M. Arora, MD, MAPP, was developed by the American College of Physicians Ethics, Professionalism and Human Rights Committee; the American College of Physicians Council of Associates; and the Federation of State Medical Boards Special Committee on Ethics and Professionalism: Janelle A. Rhyne, MD, MA (Chair); Thomas E. Reznik, MD, MPH; and Michael N. Young, MD. Members of the 2012–2013 ACP Ethics, Professionalism and Human Rights Committee: David A. Fleming, MD, MA (Chair); Ana María López, MD, MPH (Vice-Chair); Jeffrey T. Berger, MD; Thomas A. Bledsoe, MD; Clarence H. Braddock III, MD, MPH; David L. Bronson, MD; Nitin S. Damle, MD, MS; Kathy Faber-Langendoen, MD; Phyllis A. Guze, MD; Nathaniel E. Lepp, MPH; Alejandro Moreno, MD, MPH, JD; Upasna (Mini) Swift, MBBS; Jon C. Tilburt, MD; and Michael N. Young, MD. Members of the 2012–2013 ACP Council of Associates: Jay D. Bhatt, DO, MPH, MPA (Chair); Ryan Clark Van Voorhok, MD; John Peter Biebelhausen, MD, MBA; Stephen F. Darrow, MD; Morganna L. Freeman-Keller, DO; Gaurav Jain, MBBS; Ali M. Khan, MD, MPH; Brent Wallace Lacey, MD; Arta Lahiji, MD, MPH; Julissa Lombardo, MD; Thomas E. Reznik, MD; Shruti Tandon, MD; Zoe Tseng, MD; and Michael N. Young, MD. Members of the 2011–2012 FSMB Special Committee on Ethics and Professionalism: Janelle A. Rhyne, MD, MA (Chair); Radheshyam M. Agrawal, MD; Constance G. Diamond, DA; Robert P. Fedor, DO; John P. Kopetski (deceased); M. Myron Leinwetter, DO; Lance A. Talmage, MD; and Bruce D. White, DO, JD. Approved by the FSMB Board of Directors on 21 October 2012 and the ACP Board of Regents on 17 November 2012.
Table. Online Physician Activities: Benefits, Pitfalls, and Recommended Safeguards

<table>
<thead>
<tr>
<th>Activity</th>
<th>Potential Benefits</th>
<th>Potential Pitfalls</th>
<th>Recommended Safeguards</th>
</tr>
</thead>
<tbody>
<tr>
<td>Communications with patients using e-mail, text, and instant messaging</td>
<td>Greater accessibility</td>
<td>Confidentiality concerns</td>
<td>Establish guidelines for types of issues appropriate for digital communication</td>
</tr>
<tr>
<td></td>
<td>Immediate answers to nonurgent issues</td>
<td>Replacement of face-to-face or telephone interaction</td>
<td>Reserve digital communication only for patients who maintain face-to-face follow-up</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ambiguity or misinterpretation of digital interactions</td>
<td></td>
</tr>
<tr>
<td>Use of social media sites to gather information about patients</td>
<td>Observe and counsel patients on risk-taking or health-averse behaviors</td>
<td>Sensitivity to source of information</td>
<td>Consider intent of search and application of findings</td>
</tr>
<tr>
<td></td>
<td>Intervene in an emergency</td>
<td>Threaten trust in patient-physician relationship</td>
<td>Consider implications for ongoing care</td>
</tr>
<tr>
<td>Use of online educational resources and related information with patients</td>
<td>Encourage patient empowerment through self-education</td>
<td>Non-peer-reviewed materials may provide inaccurate information</td>
<td>Vet information to ensure accuracy of content</td>
</tr>
<tr>
<td></td>
<td>Supplement resource-poor environments</td>
<td>Scam “patient” sites that misrepresented therapies and outcomes</td>
<td>Refer patients only to reputable sites and sources</td>
</tr>
<tr>
<td>Physician-produced blogs, microblogs, and physician posting of comments by others</td>
<td>Advocacy and public health enhancement</td>
<td>Negative online content, such as “venting” or ranting, that disparages patients and colleagues</td>
<td>“Pause before posting”</td>
</tr>
<tr>
<td></td>
<td>Introduction of physician “voice” into such conversations</td>
<td></td>
<td>Consider the content and the message it sends about a physician as an individual and the profession</td>
</tr>
<tr>
<td>Physician posting of physician personal information on public social media sites</td>
<td>Networking and communications</td>
<td>Blurring of professional and personal boundaries</td>
<td>Maintain separate personas, personal and professional, for online social behavior</td>
</tr>
<tr>
<td></td>
<td>Impact on representation of the individual and the profession</td>
<td></td>
<td>Scrutinize material available for public consumption</td>
</tr>
<tr>
<td>Physician use of digital venues (e.g., text and Web) for communicating with colleagues about patient care</td>
<td>Ease of communication with colleagues</td>
<td>Confidentiality concerns</td>
<td>Implement health information technology solutions for secure messaging and information sharing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unsecured networks and accessibility of protected health information</td>
<td>Follow institutional practice and policy for remote and mobile access of protected health information</td>
</tr>
</tbody>
</table>

Methods

This position statement was authored on behalf of the American College of Physicians (ACP) Ethics, Professionalism, and Human Rights Committee, the ACP Council of Associates, and the Federation of State Medical Boards (FSMB) Special Committee on Ethics and Professionalism. They and the authors developed the statement between May 2011 and October 2012. After literature reviews and an environmental assessment to determine the scope of issues, drafts were debated, and a consensus was reached on issues through facilitated discussion. A draft then underwent external peer review and review by the College and FSMB committees, councils, and leadership, as well as journal peer review. After revisions based on those comments, the position paper was reviewed and approved by the committees and the FSMB Board of Directors and ACP Board of Regents. The position paper is official ACP and FSMB policy.

Positions

Position 1: Use of online media can bring significant educational benefits to patients and physicians, but may also pose ethical challenges. Maintaining trust in the profession and in patient-physician relationships requires that physicians consistently apply ethical principles for preserving the relationship, confidentiality, privacy, and respect for persons to online settings and communications.

The Patient–Physician Relationship

Standards for professional interactions should be consistent across all forms of communication between the patient and physician, whether in person or online. Encounters between patients and physicians should only occur within the bounds of an established patient-physician relationship, which entails rights and obligations for both parties. As stated in the ACP Ethics Manual, physicians “must be careful to extend standards for maintaining professional relationships and confidentiality from the clinic to the online setting” (4). E-mail and other electronic means of communication can supplement, but not replace, face-to-face encounters.

Establishing positive patient-physician relationships and maintaining professional decorum are core elements of training that should be fostered from medical school through all stages of professional development. Online professionalism can pose challenges because of the ambiguity of written language without the context of body language or lack of awareness of the potential abuses of such media (5). The ease of use and immediacy of social media tools—especially if users do not engage in “pausing before posting”—can lead to unintended outcomes or messages.
Many state medical boards have received reports of violations of online professionalism (6).

The initial decision about whether to extend the patient–physician relationship to the online setting includes the following factors: the intended purpose of the exchange and the content of conversation; the immediacy of electronic media and expectations, including response time; how communication will take place (for example, through social networking sites, microblogging, or professional e-mail on a protected server) while maintaining confidentiality; and how emergency or urgent situations will be managed.

The Patient–Physician Relationship: To Friend (and Google) or Not to Friend (and Google)?

Patients will sometimes initiate online communication. One recent study suggested that many patients extend online “friend” requests to their physicians, although very few physicians reciprocate or respond (7). Organizational policy statements increasingly discourage personal communication between physicians and patients online (8). The FSMB specifically discourages physicians from “interacting with current or past patients on personal social networking sites such as Facebook” (9).

Information exchanged on the Web is at least a 2-way street because it may also be available to the general public. Just as patients may learn about the personal behavior of physicians, physicians may observe patients participating in risk-taking or health-averse behaviors. Information about a patient from online sources may be helpful in the care of that patient, but physicians should be sensitive to the source. They should use clinical judgment in determining whether and how to reveal it during their management of the patient.

This online practice, known as patient-targeted Googling, has been described in many settings, including an attempt to identify an unconscious patient in the emergency department. But often, it instead can be linked to “curiosity, voyeurism and habit” (10). Although anecdotal reports highlight some benefit (for example, intervening when a patient is blogging about suicide), real potential exists for blurring professional and personal boundaries. Digitally tracking the personal behaviors of patients, such as determining whether they have indeed quit smoking or are maintaining a healthy diet, may threaten the trust needed for a strong patient–physician relationship (11). Commentators encourage physicians to consider the intent of the search, whether it affects continuing therapy for the patient, and how to appropriately document findings with implications for ongoing care.

Patient and Physician Education

The Internet can be a powerful tool for education. Patients can share and discuss information using illness-specific social networking pages (10). The Pew Internet and American Life Project estimates that 8 in 10 Internet users go online for health information, making it the third most popular activity online among those in Pew Internet surveys (12).

Physicians should consider the quality of online resources they recommend and guide patients to peer-reviewed media and Web sites where the quality control of information can be checked. Using and sharing recommendations from state medical boards or the College may help direct physicians and patients to resources that are more accurate and objective.

Online learning opportunities can be used by patients and physicians. New care delivery models embrace social media, especially for sharing resources in resource-poor environments (13, 14). Online decision aids are growing in popularity among motivated patients seeking health information, and they warrant familiarity by physicians (15). Continuing medical education and faculty development activities are now on the Web, with online learning modules and social media platforms available for specialists and generalists to share experiences and network.

The Internet and social networking can also serve the public health (16). For example, text messaging on a public health level can bring health benefits. But online activities also bring ethical challenges for the profession and individual physicians. Digital media may help to increase physician–physician interaction and education via online discussion communities and similar means; however, it is the responsibility of physicians to ensure to the best of their ability that professional networks are secure and that only verified and registered users have access to shared information. Online postings can also be used to help advocate for public health issues and broadly educate groups of patients on specific conditions and treatment. Clinical vignettes, however, must have all personal identifying information removed, including any revealing references to a patient who serves as the basis for an illustrative narrative. Consent from the patient to use his or her personal story online should be obtained.

Just as with informal in-person discussions among colleagues, the airing of frustrations and “venting” may occur in online forums. The ACP and the FSMB recommend against this practice, even among close contacts, as it may be disrespectful and undermine professionalism. We also caution against this practice in other forums, specifically blog postings or microblog sites, such as Twitter, as the material may present the physician or physician-in-training in an inappropriate or unprofessional light (17). Physicians criticizing late-arriving patients or disparaging patients for not adhering to behavior changes (such as diet and weight loss) can undermine trust in the profession.

Confidentiality

Confidentiality respects patient rights and privacy, and this encourages patients to seek medical care and openly discuss issues. Confidentiality may be hard to maintain given electronic health records, electronic data processing, e-mail, the faxing of patient information, third-party pay-
ment for medical services, and the sharing of patient care and information among several health professionals and institutions; therefore, “Physicians must follow appropriate security protocols for storage and transfer of patient information to maintain confidentiality, adhering to best practices for electronic communication and use of decision making tools” (4). In addition, they should be aware of state and federal legal requirements, including the privacy rule from the Health Insurance Portability and Accountability Act of 1996 (HIPAA) and updates to the rule (18).

In digital environments, the sharing of patient information must always be held to a higher level of security than standard residential Internet connections. Encrypted or virtual proxy network connections in hospital-based information technology systems should be used for all patient information exchange and review to ensure a secure digital environment. Institutional-based policies on home access of the electronic health record should be reviewed before use, specifically maintaining the level of security required for use on personal devices. Many institutions use mobile device management systems for smartphones and tablet devices. This allows for remote monitoring of the hospital’s digital “perimeter” and remote disabling of devices that are lost or confiscated.

Because many physicians use mobile devices to help manage their professional careers, mobile solutions are required to ensure confidentiality, especially when such devices or tablet computers are used to access electronic medical records. Digital devices must be configured to protect patient information should the devices be misplaced or stolen; mobile management solutions can help provide such a safety net (19). In addition, the use of public, unsecured wireless networks and cellular device networks is discouraged given their inherent public accessibility and the potential for patient information to be compromised. The recent Imprivata study of text messaging in health care settings echoes these concerns, with 64% of physician respondents classified as very concerned over HIPAA compliance when sending patient health information by text. Nearly 72% believed that secure text messaging solutions would replace standard numerical pagers in current use within 3 years (20). The disposal of old devices with hospital-based connectivity or access to the electronic health record should be managed on the basis of institutional policy.

With respect to more specific use and sharing of digital media, cell phone photography, for example, is still considered a form of photography. Despite its ease of use and ubiquity, it requires obtaining formal written consent from the patient. In taking a patient photograph or radiographic image, the physician is accepting responsibility to protect this information just as for all health records. Deidentification of radiographic images in the context of educational lectures must be ensured (21).

### Medicine and Society

Professionalism is the foundation for the social contract between physicians and society (22). In exchange for the privilege of caring for patients, as well as the status, respect, and financial compensation that accompanies that privilege, society expects physicians to practice in a professional and empathetic manner (23) and to self-regulate (4).

The intimate nature of the relationship between physicians and patients results in the expectation of high ethical behavior by physicians (24). Societal expectations often extend beyond professional practice and into the daily activities of the physician. Poor judgment reflects not only on the individual physician but also on the profession. State medical boards have the authority to discipline physicians, including license restriction, suspension, or revocation, for inappropriate uses of social media, such as improper communication with patients (for example, sexual misconduct), unprofessional behavior, and misrepresentation of credentials.

The ACP Ethics Manual requires that “physicians’ conduct as professionals and as individuals should merit the respect of the community” (4). Explicit definitions and expectations of physician behaviors, both in and outside the presence of patients, have been defined by organizations, such as the United Kingdom’s General Medicine Council (25).

**Position 2: The boundaries between professional and social spheres can blur online. Physicians should keep the 2 spheres separate and comport themselves professionally in both.**

### Role and Representation

The ACP Ethics Manual stresses the importance of maintaining public trust in the medical profession and in patient–physician relationships. To maintain the respect of the community as individuals and as members of a profession, not only should the content of all online postings be considered but also the *role* of the individual posting the information. Are individuals posting material in their role as physicians, or are they merely stating opinions and also happen to practice medicine? Can this distinction be maintained?

The American Medical Association strongly suggests divorcing public and professional digital identities, specifically maintaining separate online sites or identities for the separate roles (16). This underscores the importance of education on the use of digital media and pertinent issues of confidentiality. The ACP Ethics Manual states, “Physicians who use online media, such as social networks, blogs, and video sites, should be aware of the potential to blur social and professional boundaries” (4). Problems occur when individuals post questionable material while identifying themselves as a physician or physician-in-training (26–28).

At times, physicians may be asked or may choose to write online about their professional experiences, or they...
may post comments on a Web site as a physician. When doing so, they must disclose their credentials and any conflicts of interest. They should consider the dangers of posting or responding to comments on the Web. Truly anonymous postings do not exist on the Web, and with the increased sophistication of searching and search engines, the ability to link posts or comments to the original contributor has expanded (29). Physicians should be aware that information posted on a social networking site may be disseminated (whether intended or not) to a larger audience, be taken out of context, and remain publicly available or retrievable online in perpetuity. Physicians should follow their institutional policy on digital media (30) and seek guidance from professional societies and state medical boards.

Maintaining Boundaries

The ACP and the FSMB advise against including patients in the physician’s personal and social interactions online. Professional distance and privacy are appropriate for both physician and patient. Physicians should not “friend” or contact patients through personal social media. Physicians should familiarize themselves with the privacy settings and terms of agreements for social media platforms to which they subscribe, and they should maintain strict privacy settings on personal accounts. Professional profiles should be constructed with an explicit purpose (such as networking and community outreach).

Public Consumption

Physicians-in-training, who at present are most apt to use social media platforms, agree on the responsibility to represent themselves professionally online and are aware that they, and the profession, are being assessed by their online behaviors (7). Although narrative work has described the psychological benefit of “collective venting toward the process of being doctored” (31), the public availability of online medical class skits, songs, shows, and other material previously intended for sharing in private, physician-only audiences has called into question these traditions.

Although we will not attempt to dissect the implications of such offerings, it is clear that these are experiences that are not generally intended for public consumption and, despite any value to the psyche of the trainees, should be examined more closely by medical educators and not shared online or in other mass media. It is prudent to consider the effect of publicly posting something that initially seems like harmless medical humor. Consideration should be given to how patients and the public would perceive the material and what effect this may have on the individuals involved as well as their institutions and the medical profession. Many institutional policy statements encourage a “pause-before-posting” moment where medical professionals are asked to reflect on how the general public may perceive the content.

Interprofessional Relationships

Another issue requiring consideration is online relationships between physicians of varying levels of training, specifically, attending physicians and their students and residents. Attending physicians frequently receive online “friend” requests from students and residents (32). These digital “relationships” can also blur professional and personal boundaries, especially when the faculty physician is in the role of evaluator. Faculty and trainees should examine the purpose of initiating an online relationship and decide whether it is for ongoing mentorship, research work, or career advice (32). Regardless of intent, the traditional boundaries encouraged in trainee–faculty relationships should apply when those parties interact through social media. These boundaries should also apply with staff, other clinicians, and allied health professionals.

Position 3: E-mail or other electronic communications should only be used by physicians in an established patient–physician relationship and with patient consent. Documentation about patient care communications should be included in the patient’s medical record.

Effective communication is a foundation of a strong patient–physician relationship. E-mail or other electronic communications can supplement face-to-face encounters if done under guidelines (4, 33). Using e-mail to provide therapeutic advice is not recommended when a patient–physician relationship has not been previously established. Some state laws (for example, those in Hawaii) do not require a preexisting relationship for e-mail or other electronic consultation between a physician and a patient (that is, the physician has not met or examined the patient) (34); however, the ACP and the FSMB do not support this practice.

Documentation of communications in an established patient–physician relationship, including those done electronically, should be maintained. “Medical records should contain accurate and complete information about all communications, including those done in-person and by telephone, letter or electronic means” (4).

Situations in which a physician is approached by electronic means for clinical advice in the absence of a patient–physician relationship should be handled with careful judgment; they should usually be addressed with encouragement that the individual schedule an office visit or, in the case of an urgent matter, go to the nearest emergency department.

E-Communication and Established Relationships

E-communication between patients and physicians with an existing relationship requires discussion and previous agreement before electronic exchange is initiated. Guidelines exist for interactions with patients via e-mail (33), including the appropriate type of information to share and the expectations about turnaround time. The nature of e-mail communication ensures a written copy of
the exchange, but patient confidentiality must be assured, such as through the use of a hospital-based server. A discussion of the protections in place to ensure patient privacy must also occur.

Documentation of the patient’s consent and awareness of the security and risks associated with the use of patient–physician e-mail should be included in the medical record (35). Physicians should not use personal e-mail accounts for these communications but rather encrypted messages over secure network connections. Web-based portals offer messaging through secure accounts on the portal. Physicians must maintain appropriate boundaries (36) and recognize that electronic communication merely supplements face-to-face encounters.

Electronic communication with patients, if done in a systematic and thoughtful way, can improve patient care and outcomes. Studies have demonstrated that in patients with chronic disease management needs, supplemental electronic communication served as a “booster” to physician advice and improved adherence to therapy (37, 38). It may also improve patient and physician satisfaction by increasing the actual or perceived time spent communicating and having questions answered (39). As other Web tools begin to show promise, this communication is often not limited to standard e-mail (40). Physicians and patients should be discouraged from communicating on health matters through social media tools that are publicly viewable, do not ensure patient confidentiality, and are not readily recordable or admissible to the medical record.

Physicians should be aware of legal requirements in their states about these communications and the risk for state medical board violations or other issues if the physician is not licensed in the state in which the electronic communications are received.

“The MD Will BRB [Be Right Back]”

Expectations for immediate access have led to non–Web-based forms of communication by means of multimedia messaging services and short or text messaging services (41). Several large pharmacies and insurers have piloted systems for prescription refills and appointment updates (42); however, these interactions are largely unidirectional (such as update or reminder texts) with several layers of encryption for security. Despite these advances, current technology does not provide adequate security to prevent third-party access to information. Also, text messaging is not analogous to e-mail because of its abbreviated format and the greater possibility of missed messages. Therefore, physicians should not use text messaging for medical interactions with even established patients except with extreme caution and with patient consent.

Position 4: Physicians should consider periodically “self-auditing” to assess the accuracy of information available about them on physician-ranking Web sites and other sources online.

Ranking, feedback, and other Web sites may offer patients insight into physician training and office practices. Physicians and patients should recognize that this information may not be complete or accurate. Physicians may have little recourse in deleting misrepresentations (43–45). Establishing a professional profile so that it “appears” first during a search, instead of a physician-ranking site, can provide some measure of control that the information read by patients before and after the initial encounter is accurate. Physicians should consider doing routine surveillance (46) of their online presence by searching for their names, and they should correct inaccurate information.

Position 5: The reach of the Internet and online communications is far and often permanent. Physicians, trainees, and medical students should be aware that online postings may have future implications for their professional lives.

How one is represented affects public, patient, and peer perceptions. Colleagues may often be superior or those in an evaluative capacity. The online behaviors an individual displays may harm employability and recruitment, may result in limitations in professional development and advancement, and may reflect poorly on the profession as a whole.

Many institutions have begun to harness the power of digital media to attract patients, new faculty, or trainees, especially in allied health professional education (47). These technologies can be used as recruitment or screening tools. Employers have turned away job applicants on the basis of questionable digital behavior, including provocative or inappropriate photographs or information, content that displays drinking or drug use, and evidence of poor communication skills (48). Anecdotal reports indicate that medical school admissions offices and residency training programs are increasingly using the Web to prescreen candidates. Many trainees may inadvertently harm their future careers by not responsibly posting material or not actively policing their online content. Educational programs stressing a proactive approach to digital image (online reputation) are good forums to introduce these potential repercussions.

The implications for professional life extend beyond being a prospective applicant to career advancement. A physician’s digital image can have positive or negative career repercussions. Several very public missteps have been documented, including physicians taking digital photographs during surgery (49), posing with weapons and alcohol (in some instances during humanitarian work) (50), and unprofessional microblog posts (for example, “tweets”) (51) that may ultimately harm both the individual and the profession. One’s digital image should be actively managed beyond training by maintaining the separation of professional and personal images and the clinical and nonclinical use of social media. Being proactive by controlling posted content, using privacy settings, and limiting access to per-
sonal information is in the best interest of both the profession and the individual physician.

CONCLUSION

Online technologies present both opportunities and challenges to professionalism. They offer innovative ways for physicians to interact with patients and positively affect the health of communities, but the tenets of professionalism and of the patient–physician relationship should govern these interactions. Institutions should have policies in place on the uses of digital media. Education about the ethical and professional use of these tools is critical to maintaining a respectful and safe environment for patients, the public, and physicians. As patients continue to turn to the Web for health care advice, physicians should maintain a professional presence and direct patients to reputable sources of information.

Digital media use for nonclinical purposes may affect societal perceptions of the profession, especially when questionable content is posted by physicians in their personal use of the Web. Maintaining separate personal and professional identities in Web postings may help to avoid blurring boundaries in interactions with patients and colleagues.

The ACP and the FSMB recognize that emerging technology and societal trends will continue to change the landscape of social media and social networking and how Web sites are used by patients and physicians will evolve over time. These guidelines are meant to be a starting point, and they will need to be modified and adapted as technology advances and best practices emerge. Physicians are encouraged to take a proactive approach to managing digital identity by routinely performing surveillance of publically available information. Physicians also need to familiarize themselves with these technologies to guide themselves, and their patients, as they navigate the online terrain.

From the University of Chicago, Chicago, Illinois; American College of Physicians and Thomas Jefferson University Hospitals, Philadelphia, Pennsylvania; Federation of State Medical Boards, Euless, Texas; and South East Area Health Education Center, Wilmington, North Carolina.

Acknowledgment: The American College of Physicians and the Federation of State Medical Boards thank reviewers of this position paper: Mitchell A. Adler, MD; Moises Auron, MD; Deborah L. Baruch-Bienen; Ma; Bradley H. Crotsy, MD; Robert A. Gluckman, MD; Jay A. Jacobson, MD; Terry Kind, MD, MPH; Arash Mostaghami, MD, MPA; Susan L. Rattner, MD; Thomas E. Reznik, MD; Michael C. Sha, MD; Earl Stewart Jr.; Thomas G. Tape, MD; Susan Thompson Hingle, MD; Alan H. Wynn, MD; and Annals of Internal Medicine reviewers.

Potential Conflicts of Interest: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M12-2111.

Requests for Single Reprints: Lois Snyder Sulmasy, JD, American College of Physicians, 190 N. Independence Mall West, Philadelphia, PA 19106-1572; e-mail, lsnyder@acponline.org.

Current author addresses and author contributions are available at www.annals.org.

References

Online Medical Professionalism

Position Paper


Current Author Addresses: Drs. Farnan and Arora: University of Chicago, 5841 South Maryland Avenue, M/C2007 AMB W216, Chicago, IL 60637.
Ms. Snyder Sulmasy: American College of Physicians, 190 N. Independence Mall West, Philadelphia, PA 19106-1572.
Dr. Worster: Thomas Jefferson University Hospitals, 111 South 11th Street, Philadelphia, PA 19107.
Dr. Chaudhry: Federation of State Medical Boards, 400 Fuller Wiser Road, Suite 300, Euless, TX 76039.
Dr. Rhyne: South East Area Health Education Center, 1601 Doctors Circle, Wilmington, NC, 28401.

Author Contributions: Conception and design: J.M. Farnan, L.S. Sulmasy, H.J. Chaudhry, V.M. Arora.
Critical revision of the article for important intellectual content: J.M. Farnan, L.S. Sulmasy, B. Worster, H.J. Chaudhry, J. Rhyne, V.M. Arora.
Administrative, technical, or logistic support: J.M. Farnan, L.S. Sulmasy, H.J. Chaudhry.
Background
Young people are disproportionately affected by sexually transmissible infections (STIs) [1]. Untreated STIs can have serious health consequences including infertility, ectopic pregnancy, and pelvic inflammatory disease in women [2]. As most STIs are asymptomatic, periodic screening for certain STIs, such as chlamydia, is recommended [2]; this requires adequate access to sexual health services.

Young people may face barriers to accessing sexual health services, including concerns over confidentiality and privacy, cost, limited transport, and too few medical providers [3-6]. Living in an isolated or remote region can also limit young people’s options because there may be no available sexual health specialist and no choice of male or female doctor. Many young women in Australia prefer speaking to a doctor of the same sex [4,5,7].
Adolescents have reported concerns about being identified entering a clinic and that medical staff might disclose to others the reason for their visit [3,5]. In both rural and urban areas, concerns over the implications of sexual activity and, in particular, the stigma surrounding STIs have been reported to limit willingness to seek medical care for sexual health [6]. These findings highlight the importance of access to confidential services.

One possible means to increase access is the use of telemedicine. Telemedicine is defined as “the delivery of health services when there is geographic separation between health-care provider and patients, or between health-care professionals” [8]. Telemedicine itself falls into a broad category incorporating a range of technologies such as telephone, facsimile, and webcam consultations over the computer (also referred to as video consultations or videoconferencing) [9,10]. For the purpose of this paper, telemedicine refers to communication between patients and medical professionals. Webcam consultations in Australia have been used successfully in fields such as psychiatry, emergency care, and paediatrics [9,11].

Reviews of telemedicine between patients and medical professionals have cited numerous advantages for patients including increased access to services and providers, lessened travel and waiting time to see a doctor, and reduced cost [9,12,13]. Despite these advantages, concern has been raised about the quality of doctor-patient communication during telemedicine consultations, as well as about privacy and security [8,13].

Webcam or telephone consultations between health care providers and patients would enable young people to consult a doctor from their home computer or smart phone, obviating the need for a clinic visit and increasing their options around medical providers. After a consultation, a home STI testing kit could be posted to patients. These kits have been found to be reliable and acceptable STI screening tools [14,15].

We initiated a literature review in July 2009 to examine what was known about telephone and webcam consultations (video consultation or videoconferencing) for STI care between patients and health care professionals. A comprehensive search of the published peer-reviewed literature via Scopus, MEDLINE, Web of Science, PsycINFO, PubMed, and Academic Search Complete yielded no articles about using webcam consultations for STI care between patients and providers. Only one article was retrieved that dealt (indirectly) with the use of telephone consultations for STI care [16]. This research study intended to fill this void in the literature.

The aim of this study was to examine young adults’ pre-use views on webcam and telephone consultations for sexual health in Australia.

Methods
Study Respondents
Young people aged 16-24, living in Australia, with Internet access were eligible to participate in the study.

Instrument
An online questionnaire was deemed most appropriate to examine the relationship between health care and the Internet. In the absence of a standardised questionnaire about pre-use views on telemedicine, a study-specific questionnaire was devised; where appropriate, questions were adapted from other published questionnaires [17-19]. The national, cross-sectional SHOUT (Sexual Health Online Using Telemedicine) questionnaire had five sections: information about the respondents, their access to health care, discussing their sexual health with a doctor, IT information, and sexual behaviour. Respondents were asked their general views on webcam, telephone, and in-person consultations. Five-point Likert scales (very willing → very unwilling) were used to assess people’s willingness to have a consultation by these different media. In addition to the fixed-answer questions, respondents could provide additional or explanatory comments in the free text response boxes. Next, respondents were asked to nominate their first preference for speaking to a doctor for an asymptomatic sexual health consultation if given the choice between an in-person, telephone, and webcam consultation. For this question, respondents were instructed to imagine two possible situations: living twenty minutes or two hours from a doctor. The questionnaire was accessible on the research study’s website [20]. After piloting the questionnaire with urban and rural Australian young people, it was available to complete anonymously online from September 2009 to May 2010.

Recruitment and Advertising
The survey used convenience sampling. Advertising was concurrent with the questionnaire’s availability (9 months). A variety of advertising approaches was used: through universities, Australian organisations targeting young people, Facebook, and radio. A total of 105 diverse youth organisations across Australia were contacted about placing survey information on their website and/or newsletter; 11 (10%) (predominantly government-affiliated and rural organisations) agreed to advertise. Advertisements were placed on the University of Melbourne’s online Student Portal Notice Board. Facebook groups targeting Australian youth were also contacted about posting information about the study on their Facebook page and a paid Facebook advertisement was placed online. Of the 77 Facebook groups contacted, 16 (21%) (mainly university groups or Facebook groups aimed at
people living in rural areas) agreed to advertise. Contacting Facebook groups also resulted in advertisements in related blogs and newsletters. Of the 1855 people who clicked on the paid Facebook advertisement, 24 (1%) completed the questionnaire.

Analysis
Descriptive statistics were used to describe the study sample. Variables based on Likert scales were collapsed into binary outcomes. Chi-square tests were used to assess associations of categorical variables, and t-tests or Mann-Whitney U tests were used to assess associations between binary and continuous variables. Multinomial logistic regression was used to explore the association between the three-level outcome variables (first preference in person, telephone or webcam, and demographic and behavioural variables); odds ratios and 95% CI were calculated using in-person consultation as the reference category. Free text responses were analysed thematically.

Ethics
The study was approved by the University of Melbourne Human Research Ethics Committee (#931507).

Results
Sample Characteristics
A total of 662 people completed the questionnaire. Forty four percent of respondents (n = 289) wrote comments in the free text sections. There were 2541 visits to the study’s website. Most respondents (66%) accessed the website from a referring website, 32% accessed the website through direct traffic, and 2% found the questionnaire through a search engine. The majority of the referrals came from the Facebook website. Most respondents (66%) reported hearing about the study through a website or an electronic newsletter/email.

Median age of respondents was 20 years; most were female (74%). Respondents reported residing in all Australian states and territories except the Northern Territory with most (83%) living in a major city. The majority (88%) were either currently in tertiary education or held a bachelor’s degree or higher. Overall, 76% reported having penetrative (vaginal or anal) sex in the previous twelve months. The median number of reported sexual partners in that time period, for both women and men, was 1 (range: 0-28 and 0-19 respectively). A higher proportion of males reported same-sex partners (23% vs. 10%; p =< 0.01). Respondents were more likely to be female and aged 20 to 24 years than the general population of same age [21], and males were more likely to report male-to-male sexual contact than similarly aged males [22] (Table 1).

Thirty four percent had had a past STI test, with 19% (n = 42) of this group reporting being diagnosed with an STI. Fifteen percent (n = 102) of respondents agreed with the statement “I feel I could be at risk for a sexually transmitted infection (STI)”. Women were more likely to have consulted a doctor in the last 12 months, with a median of 4 visits compared with 2 for men (p < 0.01). Twenty eight percent (n = 185) reported that they found it difficult to access a doctor with whom they would be willing to discuss a sexual health concern and 85% (n = 158) of these said that the main reason was not feeling comfortable talking to the local doctor about a sexual health concern. Respondents in their 20s (p =< 0.01), those born in Australia (p =< 0.01), and those with three or more sexual partners in the previous year reported finding it easier to access a doctor than respondents in their teens, those born outside Australia, and those with fewer than three sexual partners in the previous year.

Willingness to have a sexual health consultation by different media
Overall, 85% of the sample indicated they would be willing to have an in-person consultation with a doctor, 63% a telephone consultation, and 29% a webcam consultation (Table 2). Some differences were found in how willing respondents were to speak to a doctor by different media (Table 3). It is notable that respondents who had had an STI test in the past were more willing to have an in-person consultation than respondents who had never had an STI test (94% vs. 84%, p = 0.01). Men were more willing than women to have a webcam consultation (36% vs. 26%, p = 0.01), as were respondents who reported same-sex partners compared with those with no same-sex partners (45% vs. 27%, p =< 0.01). Additionally, respondents reporting three or more partners were more willing to have a webcam consultation than respondents reporting fewer partners (38% vs. 27%, p = 0.01).

Sixty eight percent (n = 453) of the sample reported having access to a webcam they could use for a sexual health consultation. Of those who did not own a webcam, only 13% (n = 26) reported being willing to purchase a webcam for this purpose. There was no association between owning a webcam and willingness to have a webcam consultation (p = 0.30).

Free text responses
In addition to the forced-choice answers to the questions in this section, respondents were invited to comment in the free text boxes. The free text responses provided further insight into young people’s views on telemedicine consultations. Three main advantages of telephone consultations were identified: 1) patients could remain anonymous; 2) the telephone was deemed a more convenient and less embarrassing medium for speaking to a doctor than in person; and 3) such...
consultations were assessed as saving time because no travel to a clinic was required (Table 4). Concerns about telephone consultations included difficulty verifying the doctor’s credentials and the potential for eavesdropping.

The main objections to webcam consultations were privacy and security concerns about the possibility of the webcam consultation being recorded, saved, and potentially searchable and retrievable online (Table 4). Others found webcam consultations unnecessary because telephone was adequate if the consultation could occur at a distance. Despite these objections, a few respondents reported viewing webcam consultations as advantageous either because they avoided travel or

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>N (%)</th>
<th>CI (95%)</th>
<th>Reference Population²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>16 to 19</td>
<td>214 (32%)</td>
<td>(29%-36%) 44%</td>
</tr>
<tr>
<td></td>
<td>20 to 24</td>
<td>448 (68%)</td>
<td>(64%-71%) 56%</td>
</tr>
<tr>
<td>Gender</td>
<td>Female</td>
<td>487 (74%)</td>
<td>(70%-77%) 49%</td>
</tr>
<tr>
<td></td>
<td>Male</td>
<td>173 (26%)</td>
<td>(23%-30%) 51%</td>
</tr>
<tr>
<td>Aboriginal or Torres Strait Islander</td>
<td>Yes</td>
<td>7 (1%)</td>
<td>(0%-2%) 2%</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>655 (99%)</td>
<td>(98%-100%) 98%</td>
</tr>
<tr>
<td>Remoteness</td>
<td>Major city</td>
<td>548 (83%)</td>
<td>(80%-86%) 84%</td>
</tr>
<tr>
<td></td>
<td>Non-major city</td>
<td>111 (17%)</td>
<td>(14%-20%) 16%</td>
</tr>
<tr>
<td>Country born</td>
<td>Australia</td>
<td>515 (78%)</td>
<td>(75%-81%) 78%</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>147 (22%)</td>
<td>(19%-25%) 22%</td>
</tr>
<tr>
<td>Education</td>
<td>Did not complete high school</td>
<td>6 (1%)</td>
<td>(0%-2%) —+</td>
</tr>
<tr>
<td></td>
<td>Still studying high school</td>
<td>23 (4%)</td>
<td>(2%-5%) 45%</td>
</tr>
<tr>
<td></td>
<td>Completed high school and not studying at TAFE or tertiary degree</td>
<td>24 (4%)</td>
<td>(2%-5%) —+</td>
</tr>
<tr>
<td></td>
<td>Still studying or completed TAFE</td>
<td>27 (4%)</td>
<td>(3%-6%) —+</td>
</tr>
<tr>
<td></td>
<td>Still studying tertiary or Bachelor’s degree or higher</td>
<td>582 (88%)</td>
<td>(85%-90%) 20%</td>
</tr>
<tr>
<td>Women: any same-sex partners</td>
<td>Yes</td>
<td>38 (8%)</td>
<td>(5%-10%) 10% (16-19 yrs) 12% (20-29 yrs)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>449 (92%)</td>
<td>(90%-95%)</td>
</tr>
<tr>
<td>Men: any same-sex partners</td>
<td>Yes</td>
<td>29 (17%)</td>
<td>(11%-22%) 2% (16-19 yrs) 7% (20-29 yrs)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>144 (83%)</td>
<td>(78%-89%)</td>
</tr>
<tr>
<td>Number sexual partners in prior 12 months</td>
<td>Men with no same-sex partners</td>
<td>1.62 (mean) 1 (median) 0-19 (range)</td>
<td>1.22-2.02 (mean) 1.3 (mean, 16-19 yrs) 1.5 (mean, 20-29 yrs)</td>
</tr>
<tr>
<td></td>
<td>Women with no same-sex partners</td>
<td>1.44 (mean) 1 (median) 0-12 (range)</td>
<td>1.28-1.60 (mean) 1.0 (mean, 16-19 yrs) 1.1 (mean, 20-29 yrs)</td>
</tr>
<tr>
<td>Women: Past STI diagnosis</td>
<td>Yes</td>
<td>38 (8%)</td>
<td>(5%-10%) 3% (aged 16-19 yrs) 12% (aged 20-29 yrs)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>449 (92%)</td>
<td>(90%-95%)</td>
</tr>
<tr>
<td>Men: Past STI diagnosis</td>
<td>Yes</td>
<td>4 (2%)</td>
<td>(0%-5%) 1% (aged 16-19 yrs) 11% (aged 20-29 yrs)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>169 (98%)</td>
<td>(95%-100%)</td>
</tr>
</tbody>
</table>

¹Demographic data were compared to the Census data and the Australian Study of Health and Relationships data for similarly aged men and women.
²Census data [21]
³Remoteness defined in accordance with the Australian Standard Geographical Classification-Remoteness Area System in 2010. Major city in the study is defined as RA1; Non-major city is defined as RA2-RA5 [29].
⁴Census data [30]
⁵Direct comparisons to data provided when available from Census data [21,31]. Symbol denotes comparable data are not available.
⁶Australian Study of Health and Relationships data [22]
⁷Median and range not available for the Australian Study of Health and Relationships data [32].
⁸Australian Study of Health and Relationship data [33].
Table 2: Responses to questions regarding access to sexual health services and views on telemedicine by gender

<table>
<thead>
<tr>
<th>Question</th>
<th>Women</th>
<th>Men</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difficulty accessing a doctor for a sexual health concern</td>
<td>Easy</td>
<td>229 (47%)</td>
<td>78 (45%)</td>
</tr>
<tr>
<td></td>
<td>Neither easy nor difficult</td>
<td>117 (24%)</td>
<td>52 (30%)</td>
</tr>
<tr>
<td></td>
<td>Difficult</td>
<td>141 (29%)</td>
<td>43 (25%)</td>
</tr>
<tr>
<td>Access to a webcam</td>
<td>Yes</td>
<td>333 (68%)</td>
<td>118 (69%)</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>154 (32%)</td>
<td>55 (32%)</td>
</tr>
<tr>
<td>Willingness to have an in-person sexual health consultation</td>
<td>Willing</td>
<td>410 (84%)</td>
<td>151 (87%)</td>
</tr>
<tr>
<td></td>
<td>Unwilling</td>
<td>77 (16%)</td>
<td>22 (13%)</td>
</tr>
<tr>
<td>Willingness to have a telephone sexual health consultation</td>
<td>Willing</td>
<td>297 (61%)</td>
<td>119 (69%)</td>
</tr>
<tr>
<td></td>
<td>Unwilling</td>
<td>190 (39%)</td>
<td>54 (31%)</td>
</tr>
<tr>
<td>Willingness to have a webcam sexual health consultation</td>
<td>Willing</td>
<td>127 (26%)*</td>
<td>63 (36%)*</td>
</tr>
<tr>
<td></td>
<td>Unwilling</td>
<td>360 (74%)</td>
<td>110 (64%)</td>
</tr>
<tr>
<td>Top preference for type of sexual health consultation if imagining one lived 20 minutes from a clinic</td>
<td>In person</td>
<td>407 (84%)</td>
<td>138 (80%)</td>
</tr>
<tr>
<td></td>
<td>Telephone</td>
<td>73 (15%)</td>
<td>28 (16%)</td>
</tr>
<tr>
<td></td>
<td>Webcam</td>
<td>7 (1%)</td>
<td>7 (4%)</td>
</tr>
<tr>
<td>Top preference for type of sexual health consultation if imagining one lived 2 hours from a clinic</td>
<td>In person</td>
<td>188 (39%)</td>
<td>66 (38%)</td>
</tr>
<tr>
<td></td>
<td>Telephone</td>
<td>258 (53%)</td>
<td>82 (47%)</td>
</tr>
<tr>
<td></td>
<td>Webcam</td>
<td>41 (8%)</td>
<td>25 (15%)</td>
</tr>
<tr>
<td>Willingness to receive testing kits/treatment in post</td>
<td>Willing</td>
<td>430 (88%)</td>
<td>148 (86%)</td>
</tr>
<tr>
<td></td>
<td>Unwilling</td>
<td>57 (12%)</td>
<td>25 (14%)</td>
</tr>
<tr>
<td>Prefers another mode to speak to a doctor</td>
<td>No</td>
<td>240 (49%)</td>
<td>90 (52%)</td>
</tr>
<tr>
<td></td>
<td>Instant messaging</td>
<td>79 (16%)</td>
<td>39 (23%)</td>
</tr>
<tr>
<td></td>
<td>E-mail</td>
<td>158 (32%)</td>
<td>43 (25%)</td>
</tr>
<tr>
<td></td>
<td>SMS</td>
<td>4 (1%)</td>
<td>1 (1%)</td>
</tr>
</tbody>
</table>

*Indicates a statistically significant difference-chi-square test
Table 3 Factors associated with willingness to have a sexual health consultation by different media

| Variable         | In person |             |             | Telephone |             |             | Webcam |             |             | p value* |             |             |             | p value* |             |             |             |
|------------------|-----------|-------------|-------------|-----------|-------------|-------------|        |             |             |         |             |             |         |             |             |             |
| Gender           |           |             |             |           |             |             |        |             |             |         |             |             |         |             |             |             |             |
| Male             | 151 (87%) | 22 (13%)    | 0.33        | 119 (69%) | 54 (31%)    | 0.07        | 63 (36%) | 110 (64%)   | 0.01     |         |             |             |         |             |             |             |
| Female           | 410 (84%) | 77 (16%)    |            | 297 (61%) | 190 (39%)   |            | 127 (26%) | 360 (74%)   |           |         |             |             |         |             |             |             |
| Age              |           |             |             |           |             |             |        |             |             |         |             |             |         |             |             |             |             |
| 16 to 19         | 173 (81%) | 41 (19%)    | 0.04        | 147 (69%) | 67 (31%)    | 0.04        | 63 (29%) | 151 (71%)   | 0.86     |         |             |             |         |             |             |             |
| 20 to 24         | 390 (87%) | 58 (13%)    |            | 270 (60%) | 178 (40%)   |            | 129 (29%) | 319 (71%)   |           |         |             |             |         |             |             |             |
| Remoteness       |           |             |             |           |             |             |        |             |             |         |             |             |         |             |             |             |             |
| Major city       | 474 (87%) | 74 (14%)    | 0.03        | 353 (64%) | 195 (36%)   | 0.06        | 158 (29%) | 390 (71%)   | 1        |         |             |             |         |             |             |             |
| Non-major city   | 87 (78%)  | 24 (22%)    |            | 61 (55%)  | 50 (45%)    |            | 32 (29%) | 79 (71%)    |           |         |             |             |         |             |             |             |
| Had an STI test  |           |             |             |           |             |             |        |             |             |         |             |             |         |             |             |             |             |
| Yes              | 207 (93%) | 16 (7%)     | <0.01       | 140 (63%) | 83 (37%)    | 0.9         | 66 (30%) | 157 (70%)   | 0.8      |         |             |             |         |             |             |             |
| No               | 351 (84%) | 82 (16%)    |            | 274 (63%) | 159 (37%)   |            | 124 (29%) | 309 (71%)   |           |         |             |             |         |             |             |             |
| Any same-sex partners | | |             |           |             |             |        |             |             |         |             |             |         |             |             |             |             |
| No               | 498 (84%) | 95 (16%)    | 0.03        | 365 (62%) | 228 (38%)   | 0.02        | 160 (27%) | 433 (73%)   | <0.01    |         |             |             |         |             |             |             |
| Yes              | 63 (94%)  | 4 (6%)      |            | 51 (76%)  | 16 (24%)    |            | 30 (45%) | 37 (55%)    |           |         |             |             |         |             |             |             |
| Yearly visits to a doctor | | |             |           |             |             |        |             |             |         |             |             |         |             |             |             |             |
| 0 to 3           | 303 (89%) | 55 (15%)    | 0.75        | 234 (65%) | 124 (35%)   | 0.17        | 111 (31%) | 247 (69%)   | 0.22     |         |             |             |         |             |             |             |
| 4+               | 260 (86%) | 44 (15%)    |            | 183 (60%) | 121 (40%)   |            | 81 (27%) | 223 (73%)   |           |         |             |             |         |             |             |             |
| Partner total    |           |             |             |           |             |             |        |             |             |         |             |             |         |             |             |             |             |
| 0-2              | 440 (84%) | 86 (16%)    | 0.05        | 327 (62%) | 199 (38%)   | 0.39        | 141 (27%) | 385 (73%)   | 0.01     |         |             |             |         |             |             |             |
| 3+               | 123 (90%) | 13 (10%)    |            | 90 (66%)  | 46 (34%)    |            | 51 (38%) | 85 (63%)    |           |         |             |             |         |             |             |             |

*Chi-square test

Table 4 Free text examples of perceived advantages and disadvantage of telemedicine consultations

Advantages of telephone consultations

1. Patients can remain anonymous
   "By communicating over the phone i'd probably be more willing to discuss private details and be able to feel somewhat anonymous." (Female, aged 23)

2. Telephone consultations are less embarrassing and more convenient than in-person consultations
   "Over the phone is far less embarrassing." (Female, aged 20)
   "Telephone consults would help a lot, especially if there was a short waiting time. I hate GP waiting rooms." (Male, aged 21)

3. Time saving
   "The idea of communicating from home would in many cases be easier- no travel, less time wasted." (Female, aged 24)

Disadvantages of telephone consultations

1. Difficulty verifying the doctor's credentials and the potential for eavesdropping
   "Over the phone is probably a less appealing option because you don't know who exactly you are talking to, or if others are listening in." (Male, aged 19)

Advantages of webcam consultations

1. Enables face-to-face engagement with the doctor
   "I would be much more comfortable with a webcam than over the phone as there's much more of a sense of face-to-face contact." (Female, aged 20)
   "Great idea. Comfort [sic] of your own home, but you would be able to see that the doctor is in their office in a confidential environment." (Female, aged 21)

2. No need to travel to a clinic
   "I think [a webcam consultation is] a great idea, it would save people having to make the trip to the medical centre." (Female, aged 18)

Disadvantages of webcam consultations

1. Privacy and security concerns
   "The reason I would feel uncomfortable about using a webcam would be that I would fear someone could hack into my computer and access the chat between my GP and I. Obviously for confidentiality reasons this would be disastrous [sic]." (Female, aged 24)
   "I would be concerned about the retention of webcam data. The Doctor would need to have a policy about this. Preferable [sic] the policy would be never keep [sic] any permanent record of any data ever. If enough of this data exists it is inevitable that some of it will be misplaced or stolen at some point." (Male, aged 23)

2. Viewing webcam consultations as unnecessary
   "I don't see the point of using a webcam - if it's something that can be discussed at a distance, then the telephone should suffice. If it's something that needs to be done with visual interaction, surely it should be done in person." (Female, aged 23)
because, unlike the telephone, webcam consultations enabled face-to-face engagement with the doctor.

Preferred medium for an asymptomatic consultation
If imagining they lived 20 minutes from a doctor, 83% of respondents reported their first preference as an in-person consultation with a doctor, 15% preferred telephone, and only 2% webcam (Table 2). Multivariate analysis revealed that respondents who had never had an STI test had an increased odds of choosing speaking to a doctor by telephone as their top preference when compared to an in-person consultation (OR 1.84; 95% CI 1.08-3.14) (Table 5). Additionally, respondents with three or more partners had increased odds of preferring to speak to a doctor over webcam compared with an in-person consultation (OR 4.24; 95% CI 1.24-14.42). No other association was found.

When respondents were asked to imagine that they lived two hours from a doctor, most (51%) preferred a telephone consultation (Table 2). Thirty nine percent indicated that an in-person consultation was their top preference and 10% preferred webcam. No associations were found (Table 6).

Other modes of communication with a doctor
Respondents were asked if there was another mode of communication they would prefer to use to speak to a doctor about a sexual health matter. Fifty percent (n = 330) said no, 31% (n = 202) said they preferred email, 18% preferred instant messaging, and 1% preferred SMS.

Home STI testing kits
Eighty eight percent (n = 580) of the sample was willing to receive testing kits and/or treatment through the post.

Discussion
This study is the first we are aware of to seek the views of young people on telemedicine and access to sexual health services for STI care. The survey revealed that most young people would not use webcam consultations, because they had strong concerns about the inherent confidentiality and security. However, a minority did express a more favourable view. Men, respondents with same-sex partners, and respondents with three or more sexual partners reported finding webcam consultations more acceptable. Respondents overall were more favourably disposed to telephone consultation and most were willing to receive home STI tests and treatment through the post.

These results highlight the value of offering a variety of options for accessing sexual health services in order to cater to heterogeneous needs. While only about a third of respondents were willing to consult by webcam, such a service may be invaluable for youth who may not otherwise access a sexual health service. More research is needed to improve understanding of the circumstances in which particular subsets of young people find webcam consultations most acceptable.

The acceptability of webcam consultations may be increased by mitigating perceived privacy and security concerns. For example, clinics could clearly state their policy that webcam consultations would never be recorded or saved. Consultations could be conducted over an encrypted Internet connection for increased security. A comprehensive and comprehensible security and privacy policy would ideally be easily located on the medical centre’s website as well as reiterated before every online consultation [23]. Additionally, an attractive, professionally designed website may increase people’s trust in the medical centre’s service [23]. Such tactics have been successfully used to increase the acceptability of and trust in other types of sensitive online transactions, such as online banking [23], and may similarly increase the acceptability of online medical consultations.

It is possible that security concerns could be lessened if the consultation were not directly between the doctor and the patient in their home, but rather, as in other telemedicine services, between a distant specialist and a health care professional together with a patient in a local clinic. In this situation the service may be perceived as more legitimate and the doctor on the screen as more trustworthy because the consultation is validated by taking place in a clinic.

The privacy and security concerns expressed about webcam consultations are not specific to sexual health. The larger telemedicine literature reveals that patients commonly express privacy and security concerns about using this technology to consult a doctor [8,12]. It has also been argued that patients may be more apprehensive about their privacy during a webcam consultation compared with an in-person consultation, because there are no standards currently in place to guarantee patients’ privacy and security when their health information is transmitted online [8]. A qualitative study examining people’s pre-use views of webcam consultations for general health matters also reported that participants were concerned that, once the consultation was transmitted online, security measures could be breached and the footage could become accessible to anybody [24].

Results from other telemedicine studies suggest that webcam consultations for sexual health may be most successful in two scenarios. The first is using webcam consultations for follow-up appointments, in which the client would likely already have a trusting relationship with the health care professional. A qualitative study examining HIV/AIDS patients’ use of home telemedicine, for example, reported that, although patients were willing
Table 5 Respondents’ preferred medium for consulting a doctor if hypothetically living 20 minutes from a clinic

<table>
<thead>
<tr>
<th>Variable</th>
<th>Telephone¹</th>
<th>Unadjusted OR (95% CI)</th>
<th>p value²</th>
<th>Adjusted OR³ (95% CI)</th>
<th>p value³</th>
<th>Webcam¹</th>
<th>Unadjusted OR (95% CI)</th>
<th>p value²</th>
<th>Adjusted OR³ (95% CI)</th>
<th>p value³</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>487 (74%)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Male</td>
<td>173 (26%)</td>
<td>1.13 (0.70-1.82)</td>
<td>0.61</td>
<td>1.08 (0.65-1.79)</td>
<td>0.78</td>
<td>2.95 (1.02-8.56)</td>
<td>0.05</td>
<td>2.26 (0.68-7.57)</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 to 19</td>
<td>214 (32%)</td>
<td>1.10 (0.70-1.72)</td>
<td>0.69</td>
<td>0.97 (0.61-1.54)</td>
<td>0.89</td>
<td>2.16 (0.75-6.26)</td>
<td>0.16</td>
<td>3.22 (0.97-10.75)</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>20 to 24</td>
<td>448 (68%)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Remoteness</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major city</td>
<td>548 (83%)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Non-major city</td>
<td>111 (17%)</td>
<td>0.84 (0.46-1.51)</td>
<td>0.55</td>
<td>0.83 (0.45-1.50)</td>
<td>0.53</td>
<td>0.80 (0.18-3.62)</td>
<td>0.77</td>
<td>0.77 (0.16-3.68)</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td><strong>Had an STI test</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>223 (34%)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>No</td>
<td>433 (65%)</td>
<td>1.91 (1.16-3.13)</td>
<td>0.01</td>
<td>1.84 (1.08-3.14)</td>
<td>0.03</td>
<td>0.90 (0.29-2.79)</td>
<td>0.85</td>
<td>0.75 (0.19-2.92)</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td><strong>Any same-sex partners</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>67 (10%)</td>
<td>0.75 (0.35-1.63)</td>
<td>0.47</td>
<td>0.94 (0.42-2.11)</td>
<td>0.88</td>
<td>2.38 (0.65-8.80)</td>
<td>0.19</td>
<td>1.13 (0.25-5.07)</td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>933 (90%)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td><strong>Yearly visits to a doctor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-3</td>
<td>358 (54%)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4+</td>
<td>304 (46%)</td>
<td>0.97 (0.64-1.49)</td>
<td>0.90</td>
<td>1.10 (0.70-1.73)</td>
<td>0.67</td>
<td>0.65 (0.21-1.95)</td>
<td>0.44</td>
<td>0.86 (0.26-2.91)</td>
<td>0.81</td>
<td></td>
</tr>
<tr>
<td><strong>Partner total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-2</td>
<td>526 (80%)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>3+</td>
<td>136 (21%)</td>
<td>0.66 (0.37-1.19)</td>
<td>0.66</td>
<td>0.78 (0.42-1.44)</td>
<td>0.42</td>
<td>3.80 (1.31-11.05)</td>
<td>0.01</td>
<td>4.24 (1.24-14.42)</td>
<td>0.02</td>
<td></td>
</tr>
</tbody>
</table>

¹Reference category is in-person consultation
²Chi-square test
³Multinominal regression. Adjusted for sex, age, remoteness, past STI test, same-sex partners, doctor visits, and partner total.
⁴Reference category
Table 6 Respondents’ preferred medium for consulting a doctor if hypothetically living 2 hours from a clinic

<table>
<thead>
<tr>
<th>Variable</th>
<th>n%</th>
<th>Telephone</th>
<th>Unadjusted OR (95% CI)</th>
<th>p value</th>
<th>Adjusted OR (95% CI)</th>
<th>p value</th>
<th>Webcam</th>
<th>Unadjusted OR (95% CI)</th>
<th>p value</th>
<th>Adjusted OR (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>487 (74%)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>173 (26%)</td>
<td>0.91 (0.62-1.32)</td>
<td>0.60</td>
<td>0.85 (0.57-1.27)</td>
<td>0.43</td>
<td>1.74 (0.98-3.07)</td>
<td>0.06</td>
<td>1.63 (0.87-3.08)</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16 to 19</td>
<td>214 (32%)</td>
<td>0.92 (0.65-1.30)</td>
<td>0.63</td>
<td>0.90 (0.63-1.29)</td>
<td>0.56</td>
<td>0.98 (0.55-1.73)</td>
<td>0.94</td>
<td>1.02 (0.55-1.90)</td>
<td>0.94</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20 to 24†</td>
<td>448 (68%)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Remoteness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Major city</td>
<td>548 (83%)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-major city</td>
<td>111 (17%)</td>
<td>0.97 (0.63-1.49)</td>
<td>0.87</td>
<td>1.02 (0.66-1.60)</td>
<td>0.92</td>
<td>0.85 (0.40-1.80)</td>
<td>0.68</td>
<td>0.98 (0.46-2.11)</td>
<td>0.97</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Had an STI test</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes†</td>
<td>223 (34%)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>433 (65%)</td>
<td>0.95 (0.67-1.34)</td>
<td>0.77</td>
<td>0.92 (0.63-1.34)</td>
<td>0.92</td>
<td>0.73 (0.42-1.28)</td>
<td>0.27</td>
<td>0.64 (0.34-1.20)</td>
<td>0.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Any same-sex partners</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes†</td>
<td>67 (10%)</td>
<td>0.87 (0.51-1.50)</td>
<td>0.62</td>
<td>0.85 (0.48-1.50)</td>
<td>0.85</td>
<td>1.16 (0.50-2.69)</td>
<td>0.73</td>
<td>0.84 (0.33-2.16)</td>
<td>0.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No†</td>
<td>593 (90%)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Yearly visits to a doctor</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-3†</td>
<td>358 (54%)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4+</td>
<td>304 (46%)</td>
<td>0.77 (0.55-1.06)</td>
<td>0.77</td>
<td>0.72 (0.51-1.02)</td>
<td>0.06</td>
<td>0.82 (0.48-1.40)</td>
<td>0.46</td>
<td>0.88 (0.49-1.58)</td>
<td>0.67</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Partner total</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0-2†</td>
<td>526 (80%)</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3+</td>
<td>136 (21%)</td>
<td>0.99 (0.66-1.49)</td>
<td>0.98</td>
<td>0.99 (0.64-1.52)</td>
<td>0.95</td>
<td>1.13 (0.59-2.16)</td>
<td>0.72</td>
<td>0.98 (0.48-1.99)</td>
<td>0.95</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1Reference category is in-person consultation
2Chi-square test
3Multinominal regression. Adjusted for sex, age, remoteness, past STI test, same-sex partners, doctor visits, and partner total.
4Reference category
to have webcam consultations, they preferred first consultations to be in person in order to develop a relationship with the health care professional, which was perceived as difficult to do over a webcam [25]. Using webcam consultation in a similar manner for STI care may increase its acceptability.

The second situation where there may be value in a webcam consultation is psychological counselling following the diagnosis of an STI; mental health is one field where telemedicine has frequently been applied [10]. Reviews have found that mental health services provided to patients over video are highly reliable in comparison with in-person consultations, and that patients report high levels of satisfaction with these services [10]. Webcam consultations could thus be used when informing patients of a positive diagnosis.

The current study has some limitations. First, the results are from a self-selected convenience sample, not a representative sample. Some recruitment strategies were more successful than others. However, we have no evidence to explain why some organisations were more willing than others to advertise. It is possible that some organisations were deterred by the sensitive topic of youth’s sexual health; STI services have been perceived as “unmentionable” or controversial topics in advertising [26,27]. Comparison to the Census data reveals that women were overrepresented. Given that women in the study reported being less willing than men to have a webcam consultation, the general population may find webcams slightly more acceptable. Most respondents also had high levels of education. In Australia, people with high levels of education have higher rates of home Internet access [28]. Greater access to and familiarity with the Internet could result in a sample more able and willing to have a webcam consultation than the general population. However, such a sample may also be more aware than the general population of the security and confidentiality risks posed by an online service. The second limitation is that the study asked people’s hypothetical views on using a telemedicine service. People’s views on such service may vary if they were actually using the service. However, pre-use views are important in helping to determine whether such services should be implemented. Finally, the results from the study cannot be generalized beyond the field of sexual health.

Conclusion
While the majority of respondents were willing to have a telephone consultation, only 29% were willing to have a webcam consultation for sexual health. Although the acceptability of webcam consultations is currently low, efforts to reduce privacy and security concerns may help to augment the acceptability of such services and will influence whether webcam consultations are eventually adopted on a large scale. Furthermore, the value of webcam services to an important minority of youth should not be overlooked.

Acknowledgements and Funding
The authors thank all participants for their generous contributions of time. CCG was awarded a Melbourne International Research Scholarship, Melbourne International Fee Remission Scholarship, and an Institute for a Broadband-Enabled Society PhD Top-Up Scholarship to undertake her doctoral research. The funders exercised no influence on the research.

Author details
1Centre for Women’s Health, Gender and Society, Melbourne School of Population Health, The University of Melbourne, Victoria 3010 Australia.
2Melbourne Sexual Health Centre, Alfred Hospital, 580 Swanston Street, Carlton, Victoria 3053 Australia.
3Melbourne School of Population Health, The University of Melbourne, Victoria 3010 Australia.
4The Jean Hailes Research Unit, School of Public Health and Preventative Medicine, Monash University, Clayton, Victoria, 3168 Australia.

Authors’ contributions
CCG designed the study, advertised, collected the data, conducted the data analysis and drafted the manuscript under the guidance of her supervisors. Supervisors JH, MYC, CKF, and MK (principal supervisor) contributed to the design of the study, assisted in the interpretation of the data, and revised the manuscript. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Received: 11 June 2011 Accepted: 25 October 2011
Published: 25 October 2011

References
14. Gaydos CA, Barnes M, Aumakhan B, Quinn N, Agreda P, Whittle P, Hogan T: Can e-technology through the Internet be used as a new tool to
address the Chlamydia trachomatis epidemic by home sampling and vaginal swabs? Sex Transm Dis 2009, 36:577-580.

Pre-publication history
The pre-publication history for this paper can be accessed here:
http://www.biomedcentral.com/1471-2334/11/285/ prepub

Cite this article as: Garrett et al: Young people’s views on the potential use of telemedicine consultations for sexual health: results of a national survey. BMC Infectious Diseases 2011 11:285.
Research Article

How Do Low-Income Urban African Americans and Latinos Feel about Telemedicine? A Diffusion of Innovation Analysis

Sheba George, Alison Hamilton, and Richard S. Baker

1 Center for Biomedical Informatics, Charles R. Drew University of Medicine and Science, 2594 Industry Way, Lynwood, CA 90262, USA
2 Department of Psychiatry, UCLA and VA Greater Los Angeles Healthcare System, 11301 Wilshire Boulevard, Los Angeles, CA 90073, USA
3 College of Medicine, Charles R. Drew University of Medicine and Science, 1731 E. 120th Street, Los Angeles, CA 90059, USA

Correspondence should be addressed to Sheba George, shebaghome@aol.com

Received 13 January 2012; Accepted 31 July 2012

Academic Editor: Yunan Chen

Copyright © 2012 Sheba George et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction. Telemedicine is promoted as a means to increase access to specialty medical care among the urban underserved, yet little is known about its acceptability among these populations. We used components of a diffusion of innovation conceptual framework to analyze preexperience perceptions about telemedicine to assess its appeal among urban underserved African Americans and Latinos.

Methods. Ten focus groups were conducted with African American (n = 43) and Latino participants (n = 44) in both English and Spanish and analyzed for key themes. Results. Both groups perceived increased and immediate access to multiple medical opinions and reduced wait time as relative advantages of telemedicine. However, African Americans expressed more concerns than Latinos about confidentiality, privacy, and the physical absence of the specialist. This difference may reflect lower levels of trust in new health care innovations among African Americans resulting from a legacy of past abuses in the US medical system as compared to immigrant Latinos who do not have this particular historical backdrop. Conclusions. These findings have implications for important issues such as adoption of telemedicine, patient satisfaction, doctor-patient interactions, and the development and tailoring of strategies targeted to each of these populations for the introduction, marketing, and implementation of telemedicine.

1. Introduction

Telemedicine involves using computer information and telecommunication technologies to provide health care when the provider and care recipient are in separate geographic locations. It has been promoted as a vehicle to increase access to specialty care among the urban underserved minorities, yet little is known about its acceptability among such populations. The literature on the adoption and diffusion of new technology, such as telemedicine, suggests that stakeholders' perceptions about a new innovation and the extent to which they see it as a “relative advantage” are central to the rate of diffusion and adoption [1]. The objective of this study is to explore perceptions regarding telemedicine among African Americans and Latinos in South Central Los Angeles.

It is well documented that racial/ethnic minorities and socioeconomically disadvantaged individuals face significant barriers to receiving basic health care [2–6]. African Americans and Latinos make up the largest proportion of minority populations who experience the most severe and concentrated types of health disparities. Much of this disparity in health is thought to be due to lack of timely access to appropriate health care [3]. Medically underserved populations experiencing health disparities tend to be concentrated in either inner city or rural areas. These areas are plagued by low physician-to-population ratios, limited specialty care, and health care facilities that suffer from overcrowding, inadequate infrastructure, and inefficient organizational structures [2–5, 7–11]. Given that the Institute of Medicine's report on quality of health care has already identified
illiteracy and distrust of technology as potential barriers to the delivery of telemedicine in urban underserved settings, it is important to assess community perceptions of this technology [8]. South Central Los Angeles serves as a prime example of such an inner city setting, making it an excellent location for a case study.

Telemedicine has been promoted as an innovative approach to bridging the health care delivery gap by increasing access to services for medically underserved communities. The role of telemedicine in facilitating increased access to care has traditionally been framed in terms of its ability to mitigate geographic barriers. Accordingly, remote rural communities have been the primary beneficiaries of telemedicine implementation [12]. However, limited access to appropriate medical care, particularly specialty care, is a major challenge for inner city communities as well.

Although telemedicine has the potential of redressing the health care delivery problems of the inner city, there is little in the existing literature on telemedicine or health care in general that sheds light on perceptions about telemedicine among the general population and, more specifically, the urban underserved population [9, 10, 13, 14]. It is important to examine the concurrence (or divergence) between the medical aims that drive such solutions and “on-the-ground” perceptions of those receiving care, particularly among inner city African American and Latino populations. In a study of an urban urgent care dermatology clinic, while patients generally reported high levels of satisfaction, 36% of the study sample expressed self-consciousness around the camera and 17% were uncomfortable having facial pictures taken [15]. In terms of outcomes, the Informatics for Diabetes Education and Telemedicine (IDEATel) project [5] found that African American and Hispanic American participants were less adherent to the diabetes self-care intervention than white participants, suggesting the need for culturally tailored interventions [16, 17]. The issue of community acceptance of such new techniques has yet to be resolved.

Diffusion of innovation (DOI) theory is useful in understanding the importance of assessing perceptions about a new technology such as telemedicine among a population before its introduction in order to promote likelihood of adoption. Though there are several components to DOI theory, here we focus on its applicability to characteristics of the innovation itself, that is, how an innovation spreads from innovators to others within a social system. Rogers’ classic DOI model points to five factors that shape the rate of diffusion of new innovations among stakeholders: (a) the perception of relative advantages, (b) the compatibility with past experiences and existing values, (c) the complexity of the innovation, (d) observability of benefits, and (e) trialability of the innovation on a limited basis. For example, according to this framework, if patients perceive the relative advantages of using telemedicine to be greater than existing options with regard to savings in time/money, increases in comfort, social status, and so forth, they will be more likely to adopt telemedicine. The compatibility factor points to the importance of consistency of telemedicine use with “past experiences, existing values and needs of potential adopters” [18]. Whereas the first two factors focus on the stakeholders’ needs, experiences, and values, the latter three factors (complexity, observability of benefits, and trialability) focus on aspects of the innovation. It seems important to understand which of these factors maybe most relevant at baseline for specific type of populations within particular geographical contexts vis-à-vis a new innovation in order to best promote the diffusion of the innovation.

Most of the studies that examine patient perceptions about telemedicine tend to question participants on their past experiences of receiving health care through telemedicine [19, 20]. However, preexperience perceptions are important to the success of telemedicine adoption since they shape a patient’s initial decision to (a) sample a telemedicine service and (b) use the service on a continual basis [21]. There is scant research on viewpoints about telemedicine among the target population before the introduction of telemedicine. Some exceptions include studies by Bashshur [22], Brick et al. [23], and Turner et al. [21]. The first two studies found that patients do not perceive telemedicine as preferable to seeing a doctor in person, even though they appreciated the usefulness of telemedicine for emergency situations and minor problems. Turner and colleagues found that the greater the perceived relative advantage and the greater the perceived compatibility of the innovation, the greater the intent to adopt it with varying levels of openness depending on the task situation (e.g., their respondents were more open to telemedicine care in emergency situations than for specialist care). However, none of these studies examines the perceptions of urban inner-city populations regarding telemedicine and the specificities of their care contexts.

Given that there is little research on the perceptions about telemedicine among African American and Latino underserved populations, we examined the pretelemedicine perceptions of these groups and the differences between them. In addition to our focus on these two populations, we were interested in identifying the differences between elders in these groups (over 65) and younger adults (parents of school-aged children) since these are the two groups most likely to utilize and benefit from telemedicine services with a clear source of health care reimbursement. We hypothesized that the elderly would be less amenable to the idea of new technology and parents would be more willing to try the new technology to meet the needs of their children.

2. Methods

2.1. Setting: South Central Los Angeles. The research was conducted in South Central Los Angeles, which is home to more than 1.4 million individuals, most of whom are racial/ethnic minorities (62.7% Hispanic, 33.4% African American). South Central is the most socioeconomically disadvantaged community in Los Angeles, with 28% of the population living below the federal poverty level [24]. The population faces several barriers to receiving timely care: in 2005, 40.2% reported that they could not afford to see a physician when needed and 27.5% of adults reported transportation problems as a barrier that kept them from obtaining needed medical care [25].
2.2. Procedures. Focus group methodology was utilized to explore the range of individual opinions within relatively homogeneous groups (described below), using a standardized set of questions [26–29]. The research team consisted of the authors and two research associates who assisted with the recruitment and moderation of the focus groups.

Community-based recruiting efforts were used to develop a sample population for the focus groups. Flyers about the focus groups were posted in community centers and public housing complexes. Interested individuals called the number on the flyer. When 8–10 individuals from the priority populations (African American and Latino parents of school-aged children and seniors) responded to these efforts, focus groups were assembled (see Table 1 for group composition). The focus groups took place in community and senior centers. Informed consent was obtained from all participants. All participants completed a background questionnaire and were paid $20 at the end of the focus group.

After introductions, participants were asked for their definitions of the word “telemedicine.” After a short discussion, a brief video presentation—a dramatization of a patient, receiving care for ear pain at a telemedicine clinic—was shown to focus group participants. Groups that were Spanish speaking (5 of the 10 groups) were shown a Spanish version of the video. In the video, the patient’s ear pain is assessed by a physician’s assistant (PA) who contacts an ear, nose and throat (ENT) specialist using videoconferencing. This ENT specialist is depicted as being distant from the clinic; he examines the patient using an otoscope with a camera at the end, which transmits images of the patient’s ear to the specialist. All parties (patient, specialist, PA) are able to see each other through videoconferencing technology.

The video was followed by a focus group discussion about participants’ reactions to and perceptions about receiving medical care through telemedicine. The moderator used a semistructured interview script that covered reactions to the video, perceived advantages and disadvantages of telemedicine, diagnoses/health conditions for which telemedicine would be appropriate, and general experiences in receiving health care services (Table 2).

2.3. Data Analysis. All interviews were audio- and videotaped and transcribed, and all Spanish-language transcripts were translated into English by a professional transcription and translation agency. Atlas, it was used for data analysis. The transcripts were analyzed using the constant comparative method of data analysis [30]. Transcripts were initially deductively coded by the second author with questions from the interview script guiding the predominant themes. These themes were summarized and discussed by the research team, and then the data underwent another level of more inductive coding to explicate the range of issues that were raised in response to each question and to compare across categorical groupings (parents versus elders, African American participants versus Latino participants). Through an iterative process of immersion in the data and refining the categories, key themes and DOI theoretical insights were identified and interpreted collaboratively by the authors.

3. Results

Participants emphasized two DOI factors: relative advantages and compatibility with experiences and existing values. There were some differences between African Americans and Latinos in how they viewed these factors. While the two groups had similar perceptions of the relative advantages of telemedicine, they had differing perceptions of the compatibility of telemedicine with their experiences and existing values, resulting in different types of concerns (see Table 3). Participants were understandably less prone to raise innovation-focused factors that are important to rate of adoption (complexity, observability of benefits, and trialability) because they were not familiar with telemedicine. We did not identify consistent discernible differences by age.

3.1. Relative Advantages. For both African Americans and Latinos, there were several relative advantages to telemedicine as compared to their usual modes of health care. The main advantages noted in all of the focus groups were: (1) reduced waiting time, (2) immediate feedback as to diagnosis and course of action, (3) increased access to specialists, and (4) increased access to multiple medical opinions. It is important to note that these perceived advantages are not necessarily correct perceptions of how telemedicine operates, but they do illustrate the values that participants associated with this type of system.

With regard to speed and accuracy of diagnosis, one Latino participant said that telemedicine would be a “novelty” because “it can give you the diagnosis right away cause they’re consulting the specialist so you can get your diagnosis instantly. I think that’s good.” Another Latino participant in another group said: “Science is more advanced and you will be able to see everything through the Internet… It will be like having the doctor in front of you but you won’t have to go to his office. The laboratory won’t take a lot of time and you will really know what you have.”

Telemedicine’s potential convenience in terms of these issues and in terms of logistics (such as location) was perceived to be very appealing. The African American participants felt that telemedicine would be particularly beneficial for children and the elderly. For example, one participant said, “I can see it going places. I can see where people will like it. Young people will love it. Their families, I can see my children, you know, loving it for their children, you know, in many cases. First, because they do not have enough time to do whatever, you know, because they are so busy all the time. So that helps to get an immediate feedback and to give a diagnosis and a solution to a problem.”

While the same four major advantages were discussed in all of the groups, Latino participants also noted several additional advantages and seemed, overall, more positive and enthusiastic about the prospect of telemedicine. They felt that telemedicine could potentially cut down on misdiagnoses, particularly because the computer gives “exact data.” This idea of the precision of computers was raised in three of the Latino focus groups. One group felt that telemedicine might result in more choice over which doctor is assessing
responses were in order to gain more in-depth information. More probing questions were asked of participants depending on what their perceptions about the compatibility of this innovation with their experiences and existing values. Participants’ main concerns about telemedicine were confidentiality/privacy (considering the use of the Internet for the transmission of personal information) and the process of diagnosis (considering the use of scopes rather than actual clinical observation, i.e., physician’s physical presence). Overall, African Americans were more concerned about these issues, and were especially concerned about the physical absence of the physician and the perceived inability to monitor the (distant) specialist’s qualifications and level of attention. Latino participants were substantially less concerned about these issues and in some cases felt very differently about them. They did, however, express concerns about whether telemedicine would be made accessible to uninsured and undocumented individuals.

3.1.2. Technology Issues: Confidentiality and Privacy. For both African American and Latino participants, the technology critical to telemedicine posed some problems. On a technical level, some participants in both sets of groups discussed the possibility that the computer could go down or the system could fail. More important than this concern, however, was that personal information could be obtained by individuals other than those involved in the telemedicine encounter. For example, there was discussion among African American participants that one’s identity could be stolen and that one’s pictures would be “floating around.” The Internet was perceived as “insecure” and “for anybody.” One African American participant noted, “Internet is the Internet. So that was perceived as “insecure” and “for anybody.” One African American participant noted, “Internet is the Internet. So that was perceived as “insecure” and “for anybody.” One African American participant noted, “Internet is the Internet. So that was perceived as “insecure” and “for anybody.” One African American participant noted, “Internet is the Internet. So that was perceived as “insecure” and “for anybody.” One African American participant noted, “Internet is the Internet. So that was perceived as “insecure” and “for anybody.”

Table 1: Focus group composition.

<table>
<thead>
<tr>
<th>Groups</th>
<th>N</th>
<th>Groups</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>African Americans N = 43</td>
<td></td>
<td>Latinos N = 44</td>
<td></td>
</tr>
<tr>
<td>Seniors, N = 37 (average age = 67; range 61–83 years)</td>
<td>1</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Parents, N = 50 (average age = 34; range 21–55 years)</td>
<td>5</td>
<td>10</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Focus group script—interview themes and examples of questions*.

<table>
<thead>
<tr>
<th>Broad themes</th>
<th>Example questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>A telemedicine clinic in your community</td>
<td>(i) How do you feel about it? (ii) How did you form this impression? (iii) From what particular experiences?</td>
</tr>
<tr>
<td>Perceived advantages and disadvantages of telemedicine</td>
<td>(i) What are specific benefits? (ii) What are potential challenges? (iii) Would telemedicine address any specific gaps/issues you have with your present form of health care?</td>
</tr>
<tr>
<td>Ideal recipients of telemedicine care</td>
<td>(i) Would you use telemedicine yourself? (ii) Would you recommend it to a friend? (iii) Would it be particularly suitable for older people/young children?</td>
</tr>
<tr>
<td>Conditions and context of use</td>
<td>(i) For what types of health conditions would you be most comfortable using telemedicine? (ii) How often and under what conditions (e.g., weekends only) would you want to use such a clinic?</td>
</tr>
</tbody>
</table>

*These are examples of only some of the initiating questions used. Other more probing questions were asked of participants depending on what their responses were in order to gain more in-depth information.

you and might provide better doctors. One group pointed out that telemedicine would result in more jobs for nurses.

The location and convenience of the clinics were also discussed more extensively in the Latino focus groups, because they felt that the clinics would be easier with children and with transportation. For example, “I would love something like this to open as soon as possible, because we need it. We need it for all of our children, because sometimes we take them all in when one has an appointment. You save time seeing the specialist that one of your children needs, or if another specialist is needed, you don’t waste any time; you save time and see the doctor you want to see and it would be great if Medi-care would pay for these services.”

3.1.1. Concerns about Compatibility with Past Experiences and Existing Values. African Americans and Latinos had very different perceptions about the compatibility of this innovation with their experiences and existing values. Participants’ main concerns about telemedicine were confidentiality/privacy (considering the use of the Internet for the transmission of personal information) and the process of diagnosis (considering the use of scopes rather than actual clinical observation, i.e., physician’s physical presence). Overall, African Americans were more concerned about these issues, and were especially concerned about the physical absence of the physician and the perceived inability to monitor the (distant) specialist’s qualifications and level of attention. Latino participants were substantially less concerned about these issues and in some cases felt very differently about them. They did, however, express concerns about whether telemedicine would be made accessible to uninsured and undocumented individuals.
for the most part, expressed that maintenance of confidentiality was the physician’s responsibility, and that the physician would not risk his license with a questionable system: “I don’t think [the doctor] would risk his degree to give out the files of all the patients that are in the computer because he would be responsible.” In another group, participants discussed asking for confidentiality, and they felt that by asking, confidentiality would be assured: “If you tell the person who’s going to carry out the treatment that you want confidentiality in your case, I don’t think there would be any problem. But you must ask for it. It won’t come on its own.”

For some Latino participants, the technology assured more privacy. One Latino participant stated, “I feel there’s more privacy. I really like the idea because the computer gives you exact data. It makes me feel better, you know? “Cause the fact that you’re being looked at through the computer, it removes the self-consciousness, shame, or whatever of talking openly to a doctor. Like this, without being face to face, I can say whatever I wanted.” There was some concern about identity theft, but, overall, Latino participants felt confident that transmitted information would remain confidential. As noted above, even for those who were not convinced of the confidentiality, there was typically little concern. One Latino participant said, “It doesn’t matter to me that people should see me because the whole world has to know what science is doing.”

3.1.3. Diagnosis and the Physical Absence of the Specialist. One of the main topics addressed in the focus groups was the physical absence of the specialist, which is one of the main distinguishing features of telemedicine. Discussions on this topic revealed the complexity of the doctor-patient interaction, illustrated by the multiple layers of meaning that participants attach to such interactions. Because of the richness of this set of findings, our results regarding the importance of physical presence and touch in telemedicine for these populations will be elaborated upon in a separated publication and here we will provide a summary of the findings. In general, the participants associated the physical presence or absence of the specialist to their (1) satisfaction with a medical encounter, (2) level of assurance that appropriate information was being conveyed, and (3) ability to accurately gauge the reactions of the specialist and monitor the latter’s activities.

Several of the African American participants’ concerns about not being physically with the specialist seemed related to sensory experiences of vision and touch, that is, being unable to “see” the specialist in person and/or not having the specialist touch the patient. For example, the physical absence of the physician was related to concerns about being able to assess if “the truth” was being told to the patient. For others, it was about being able to monitor the activities of the specialist (e.g., “How do I know that the doctor ain’t on the other side and he’s getting high?”). The importance of the physical presence of the specialist, particularly sight and touch, was often related to the specialist’s capacity to make accurate diagnoses.

The Latino participants seemed less concerned about the physical absence of the physician in the telemedicine clinic. Having the doctor physically present did not equate with better care for these participants, as expressed in such statements as, “It makes no difference having the doctor in front of you.” Participants expressed that even when the doctor is present, they tend to “only ask questions,” whereas it is the nurse who provides care. The doctor “bases his opinions on what the nurse tells him,” so diagnosis could take place just as well from a distance.

Interestingly, some Latino participants expressed a preference for telemedicine because of the physical absence of the physician. The reasons for this preference seemed to be centered on embarrassment about gender, age, and class differences between the provider and patient. As one participant explained, she preferred gynaecological exams by telemedicine because it would help her avoid in-person interactions with “young, attractive” (male) gynaecologists.

3.1.4. Qualifications and Qualities of the Physicians. As noted above, some African American participants were concerned that the telemedicine physician would not be giving the patient his/her undivided attention. This relates to an issue
that came up in several of the focus groups, which is how do you trust in the physician who is not in the room with you? How do you know he is qualified and certified?

One African American participant wondered how experienced the telemedicine physicians would be: “How many years of experience have they had? You know, some of them might not even have but six months, some might not even have a year. So you have to take all that into consideration because I myself don’t want anything that hasn’t been in medicine over a year to be looking at me… I still prefer an experienced doctor, whether he’s on telemedicine or I see him in person.” There was suspicion that the physician might not be who s/he claims to be, as expressed by an African American participant in the following question: “What is the reassurance that we have that this so-called specialist that’s on the screen really is what he’s supposed to be?”

Latino participants had more discussions about how they know the qualifications of any physicians, telemedicine or not. Most often, knowledge of a physician’s quality and qualifications came from the success of the treatment, the physician’s interpersonal qualities, or other people’s recommendations. For example, “I have been seeing my current doctor for more than seven years, and he gives you the medicine and so you don’t have to come back. And that’s how we would know if they are good doctors or not.” Another participant responded, “If I go to my doctor, I’m not 100% sure if he is a doctor or not. In terms of whether or not a doctor is good, well, you try him and see. I like the way I was treated.”

In two Latino focus groups, participants agreed that one knows of a doctor’s quality because “the medicines he gives you do you good.” The participants said that the telemedicine personnel would be responsible for assuring the quality of the physicians: “We are trusting in you like we trust in the clinics we go to. We trust that the doctor we are going to see is really a certified doctor who has gone to school and who knows medicine. I think you must take that risk, for it’s the responsibility of those who are in charge of the clinic.” Latino participants also discussed seeking information on their own as to the qualifications of physicians, for example, by looking on the Internet: “All you do is go to a website and all you have to do is fill in the doctor’s name and the clinics you’ve been to. There are many doctors that have done bad things and they are in jail, and their names are not on the list and that’s another way to find out if a doctor is good or not.” In general, while both African American and Latino participants shared concerns about the qualifications of the telemedicine physician providing care, the latter tended to think that the risks were not necessarily greater for telemedicine-based physicians as compared to physicians seen in person, and they expressed more trust that the quality of the physicians would have to be acceptable.

4. Discussion

Telemedicine has been promoted as an innovative approach to bridging the health care delivery gap particularly for underserved communities. While inner-city minority communities could potentially benefit from this innovation, there is little in the existing literature that speaks to the acceptability of such a solution among minority populations. To the best of our knowledge, this is the first study that explicitly examines perceptions about telemedicine among urban underserved minority populations, although some studies on telemedicine have included minority cultural groups [31] and studies of minority perceptions of health care in general have been done [32].

Both African American and Latino focus group participants emphasized two DOI factors that shape the rate of diffusion of an innovation: relative advantage and compatibility with past experience. Participants were less likely to discuss complexity, observability of benefits, and trialability of telemedicine, likely because these factors focus on features of the innovation, with which the study participants were not very familiar. We contend that they were more likely to talk about telemedicine’s relative advantages and compatibility since these factors were salient to their current concerns about their health care needs, lived experiences, and existing values, and they could be discussed despite their lack of firsthand experience with telemedicine.

The advantages of any health care innovation are usually assessed by potential users relative to their current experiences of receiving care. This was true regarding telemedicine for the focus group participants. Given their underserved inner-city location, the study participants overwhelmingly identified timely access to care as one of the greatest relative advantages of telemedicine. Telemedicine appears to provide some relatively efficient solutions to issues such as the challenge of transportation to get to specialist care, lack of timely access to specialists, the lack of timely diagnoses and feedback, and the lack of multiple opinions in a specialist-scarce zone.

However, the two groups had different concerns about health care received through telemedicine, reflecting differences in the compatibility of their lived experiences and values with the perceived nature of telemedicine-based care. For African Americans, their experiences as a community with a history of slavery and continuing racism in many aspects of their lives, particularly with health care, may affect their views on new and innovative medical care [33, 34]. The legacy of past abuses such as medical experimentation on slaves and the Tuskegee syphilis experiment and other types of continuing racism in health care contribute to lower levels of trust and a higher level of suspicion [34–37].

A related issue that has been studied in more detail is the attitude of minorities toward enrollment in medical research, where similar findings have been reported about African American attitudes towards research [38–46]. Among African Americans, mistrust is frequently associated with the perception that research will benefit whites or the research institution and not people of color. Furthermore, mistrust of the health care system was a primary barrier that prevented African Americans from participating in medical research [38].

For the African American participants in this study, the emphatic need to “see” and “touch” the physician seemed related to similar issues of trust. The physical absence of the physician, the instability of technology, and the inability to
monitor the specialist’s qualifications were all highlighted as concerns with telemedicine for the African American participants. All these concerns reflect a sense of vulnerability when placing trust in a medical system that historically has been unreliable and not trustworthy. African American participants expressed a need to be vigilant and monitor physicians to make sure that they would get quality care, particularly when telemedicine appeared to present greater opportunities for care to be compromised. This concern about quality of care is consistent with literature that indicates African Americans’ less than satisfying interactions with physicians [47, 48].

With regard to technology, there were many levels of concern. First, there was concern about whether the scopes used in telemedicine would perform adequately to allow physicians to make accurate diagnoses. Second, there was some concern about the computer system failing. However, the bulk of the apprehension among African Americans regarding technology was about the insecurity of transmitting personal data and images over the Internet when using telemedicine. A third issue of trust reiterated by African American participants was that of being able to trust the qualifications and qualities of the physician who is not in the room. There was concern about the level of experience of the physician, suggesting that these participants were concerned that telemedicine might be a way to unload inexperienced or second-rate doctors on them.

In contrast, the Latino participants had distinctly different responses to telemedicine, which may be explained partly by their different vantage points and lived experiences. Latinos, across age groups, appeared to have a significantly more trusting attitude towards the health care system in general and telemedicine in particular. This difference was reflected in their very different attitudes towards the telemedicine-related issues identified as problematic by African Americans, namely, the physician’s virtual presence, the usage of technology, and the qualifications and qualities of physicians. The Latino participants’ relative lack of concern about the physical absence of the physician points to the possibility that physical exams and the touch of the physician in time-pressured primary care visits are becoming less frequent [49] and consequently telemedicine is not that different from their expected standards of care.

Latino participants tended to equate the use of technology with access to scientific advances and expressed faith in the appropriate authorities to maintain confidentiality. Technology was seen by many Latinos as assuring greater accuracy (more exact data). Such optimism and openness towards technological innovations among Latinos was markedly different from the attitude found among African American participants. Despite the fact that both groups may experience what is commonly called the “digital divide,” they had noticeably different opinions about technology in general.

Latino participants also differed from African Americans in that they trusted the administrators of both telemedicine and non-telemedicine clinics to be responsible for hiring qualified doctors. Finally, the knowledge of the quality and qualifications of the physicians was determined by the success of the treatment, whether telemedicine or nontelemedicine based.

The qualitative racial/ethnic differences in attitudes about telemedicine-based health care among Latinos and African Americans point to differences in their lived experiences and values. The point of reference for many African Americans is the history of racism and medical experimentation and abuse they have experienced collectively in the United States. In contrast, immigrant Latinos encounter the US medical system without this particular historical backdrop and their point of reference maybe less than optimal health care in their home countries, along with a generally positive perception of the American health care and medical education systems. For many of the immigrant Latinos, access to American health care and especially telemedicine-based care that is perceived as scientifically and technologically cutting edge also seems to be seen as a positive improvement. Thus, in terms of the DOI framework, there appears to be good compatibility between the needs, lived experiences, and values of Latinos with the structure and delivery of telemedicine-based care.

4.1. Implications for Telemedicine. Our findings of differences in attitudes toward telemedicine suggest that it will be necessary to tailor approaches to the introduction, marketing, and implementation of telemedicine among these different populations. It is critical to gather this information before the extensive introduction of telemedicine clinics in inner-city communities for at least three reasons.

First, this information can be important for determining the best manner in which to introduce and market telemedicine among these two groups. Based on the findings from this study, it is important to identify the gaps in knowledge or the misinformation that can lead to distrust of new technology or the overestimation about the benefits of new technology and false expectations. The information gathered from this study can be used to help lower the barriers to acceptance of telemedicine by developing educational materials that address misinformation and gaps in knowledge. Marketing information could be tailored to address the specific concerns voiced by the two racial/ethnic groups, such as clearly informing African Americans about the medical qualifications of the specialists and the security procedures for maintaining confidentiality and level of diagnostic accuracy using telemedicine equipment.

Second, this information can be important in selecting the optimal ways in which to implement new telemedicine clinics. For example, for African Americans, having an initial in-person meeting with a physician may be important in helping establish trust and better preparing the patient for future virtual appointments. For real-time telemedicine consultations, cameras could be set up to make the specialist’s activities especially transparent to the patients. Physicians’ assistants or the nurses in the clinics and the specialists involved in telemedicine could be better informed about the concerns of each of these groups so that they can address these concerns (such as reassurances about confidentiality), even if the patients do not voice them.
Third, this data can also serve as a baseline point of comparison for studies that will examine changes in patient perceptions over time. As telemedicine becomes implemented in urban settings and becomes more familiar to African American and Latino populations, it will be important to have an understanding of their baseline pre-experience perceptions regarding telemedicine to gauge the changes in attitudes towards telemedicine as it spreads into different communities.

4.2. Limitations. There are several important limitations to our data and study findings. First, we have a relatively small convenience sample and our participants are not statistically representative of the wider population in inner-city settings. However, as is common to qualitative methods, they represent information-rich cases, homogenously stratified across race and age, to allow in-depth understanding of the perceptions about telemedicine among these groups. Another limitation is that for the majority of our participants, the only information about telemedicine came from the video they saw at the beginning of the focus group. While telemedicine was portrayed in a typical setting with a typical health problem, our participants’ understanding and consequent reactions to telemedicine were clearly influenced by what we were able to show them in a short video. For example, we represented telemedicine primarily as a diagnostic interaction with a specialist and did not address other potential uses, such as long-term management of chronic diseases. Our finding of no age group differences may be a reflection of the limitations of our study design. We may have needed a more sensitive interview protocol that would have more finely delineated the nuances of age differences in our sample.

5. Conclusion

Using the DOI framework regarding features of an innovation, this study contributes to an underserved area by exploring the pre-experience perceptions of telemedicine among urban, underserved African Americans and Latinos. Despite reservations, many participants indicated that they would take advantage of telemedicine clinics.

Through this study, we were able to identify components of the DOI framework that spoke to the experiences of the two minority groups—particularly with regards to compatibility with past experiences and existing values. It will be important to develop larger studies in different geographical regions with different populations to further understand the importance of these factors for the introduction/marketing, implementation, and eventual adoption of telemedicine among diverse populations.

Acknowledgments

The authors acknowledge the National Center for Research Resources (NCRR) Research Centers in Minority Institutes (RCMI) Grant G12-RR03026 at Charles Drew University; the Agency for Health Care Research and Quality (AHRQ) Grant 1R24-HS014022-01A1; the NIH-NIMHD Grant U54MD007598 (formerly U54RR026138); and the Community Technology Foundation Grant (2004-TT-002) for support during the research and writing of this paper.

References

Consumer health informatics: results of a systematic evidence review and evidence based recommendations

Michael C Gibbons, MD, MPH,1,2,3,4,5 Renee F Wilson, MA,1 Lipika Samal, MD,2 Christoph U Lehmann, MD,2,5,6 Kay Dickersin, PhD,1 Harold P Lehmann, PhD,2,5,6 Hannan Aboumatar, MD,2 Joseph Finkelstein, MD, PhD,2,5 Erica Shelton, MD, MPH,1 Ritu Sharma, MD,1 Eric B Bass, MD, MPH1,2,7

Abstract
An increasing array of technology based tools are available for patient and consumer utilization which claim to facilitate health improvement. The efficacy of these Consumer Health Informatics tools has not previously been systematically reviewed. As such a systematic evidence review of the efficacy of consumer health informatics tools was conducted. This review also sought evidence of any barriers to future widespread utilization of these tools and evidence of economic impact of these tools on health care costs. The findings of this review indicate that while more work needs to be done, the available literature does suggest a positive impact of consumer health informatics tools on select health conditions and outcomes. Many barriers remain that must be overcome prior to widespread utilization of these tools. There was insufficient data regarding economic impact of consumer health informatics tools on healthcare costs.

Keywords
Consumer health informatics, Systematic evidence review, eHealth, Health information technology

INTRODUCTION
We are living in a time, in the evolution of the US healthcare system, when significant forces, including rising costs, a growing proportion of the population is over the age of 65, and the increasing prevalence of chronic disease, are suggesting a need for significant changes in the traditional healthcare system [1]. Effective chronic medical care, unlike acute disease treatment, is often a much more collaborative process between patients and providers. It involves a much larger reliance on patient and caregiver engagement and shared decision making [1]. Recent reports on patient safety [2] and healthcare disparities [3], among others, document the inability of the current healthcare system to provide consistent high-quality care to every patient. In addition, a review of 31 national surveys found that two thirds of Americans do not actively and consistently perform the actions directly linked to benefiting from available healthcare [4]. At the same time, approximately 10 million people on any given day or by 2009, a total of approximately 175 million Americans have used the Internet to search for health information [5]. Taken together, these data suggest that there is both great patient need and consumer appetite for resources, especially electronic resources, to enable and improve their engagement in healthcare.

Although one previous review did evaluate barriers and drivers of utilization (6), prior to the present study, a comprehensive, systematic evaluation of the efficacy of these types of electronic tools has not previously been undertaken. In 2001, the term Consumer Health Informatics was introduced in an attempt to describe a focus on patient oriented technology based health behavior and health information supports. Early evidence suggest there may be value in using these tools to help patients achieve desired clinical goals.

Implications

Practice: In the future, the scope of clinical practice should include patient oriented technology based health behavior and health information supports. Early evidence suggest there may be value in using these tools to help patients achieve desired clinical goals.

Policy: Resources should be devoted to facilitate the integration of consumer health informatics tools with provider health informatics tools. The development of patient oriented meaningful use criteria may be a useful starting point.

Research: Researchers should continue to build the evidence base and further clarify the role of consumer health informatics tools particularly among among the elderly, children and medically underserved populations. There is also great need to elucidate design and development principles that can ensure the widest possible usability of the most efficacious tools.
health information from the tool application or system, and (3) is one in which the data, information, recommendations, or other benefits provided to the consumer, may be used in coordination with a healthcare professional but is not dependent on a healthcare professional [7]. Here, we distinguish patients (individuals who are already ill) from consumers (individuals who are not ill).

To understand the potential impact of these consumer health informatics tools, we conducted a Systematic Evidence Review of the available scientific literature. For the purposes of this review, we have excluded point of care devices (e.g., glucometer, remote monitoring devices), prescribed clinical devices that are part of the provision of clinical care, general information websites, message boards, and applications that are designed for use in a healthcare delivery environment. This definition has the following advantages: first, it keeps the focus of the review on how CHI applications meet the needs of consumers rather than the needs of clinicians. Second, it helps avoid a categorical disease-oriented evaluation of every clinical technological development for every disease which is not necessarily focused on the needs of consumers. Third, it helps to keep the focus of the review on studies that demonstrate impact, value or efficacy from the perspective of consumers. Finally, it facilitates categorization of CHI applications in ways that may be more meaningful for patients.

Potential categories of CHI tools or applications included in the review, but not be limited to, are (a) applications that facilitate knowing, tracking, or understanding clinical parameters (disease management); (b) applications that facilitate knowing/tracking/understanding observations of daily living; (c) applications and technologies that facilitate calendaring (lifestyle management assistance); (d) applications and tools that facilitate prevention and health promotion; (e) applications that facilitate self-care; and (f) applications that facilitate assisted care and care giving.

Primary objectives of this review were to review and synthesize the available evidence regarding the impact of currently developed CHI applications on health and healthcare processes and outcomes, identify barriers to the use of CHI applications, identify the gaps in published information on costs, benefits, and net value of these applications, and finally, to identify what critical information is needed for consumers, their families, clinicians, and developers to clearly understand the value of CHI applications.

**METHODS**

A core team of experts, who have strong expertise in clinical and health sciences informatics, clinical trials, systematic literature reviews, epidemiological studies, and general medicine was assembled along with two advisors who have done extensive research in the areas of open access, health policy, eHealth, and CHI. Additionally seven external technical experts from diverse professional backgrounds including consumer advocates, research methodologies ethics, decision aids, CHI, to CHI user acceptance were identified. Finally, two additional peer reviewers, who were not otherwise involved in the project, were identified to provide comments and feedback on the review.

The core team worked with the external advisors, technical experts, and representatives of the Agency for Healthcare Research and Quality (AHRQ), which commissioned the review, to refine a set of guiding questions. The final key questions to be answered by this study included are shown in Table 1.

To answer Key Question 1, we reviewed research employing randomized controlled trial (RCT) study designs. To answer Key Question 2, we reviewed research that was designed to look at barriers to use of CHI employing any study design. All research identified in the process of evaluating Key Questions 1 or 2 was also used to evaluate Key Questions 3 and 4.

The study team developed the conceptual model, shown in Fig. 1, to address the key questions (above). The goals of the model were to direct our review of the relevant literature and to assist reviewers in understanding which articles applied to our strict criteria for inclusion.

Searching the literature involved identifying reference sources, formulating a search strategy for each source, and executing and documenting each search. For the searching of electronic databases, we used medical subject heading (MeSH) terms. To identify

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Key Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) What evidence exists that CHI applications impact</td>
<td>a) Health care process outcomes (e.g., receiving appropriate treatment) among users,</td>
</tr>
<tr>
<td></td>
<td>b) Intermediate health outcomes (e.g., self-management, health knowledge, and health behaviors) among users,</td>
</tr>
<tr>
<td></td>
<td>c) Relationship-centered outcomes (e.g., shared decision making or clinician-patient communication) among users,</td>
</tr>
<tr>
<td></td>
<td>d) Clinical outcomes (including quality of life) among users, e) Economic outcomes (e.g., cost and access to care) among users?</td>
</tr>
<tr>
<td>2) What are the barriers that clinicians, developers, consumers, and their families or caregivers encounter that limit utilization or implementation of CHI applications?</td>
<td></td>
</tr>
<tr>
<td>3) What knowledge or evidence exists to support estimates of cost, benefit, and net value with regard to CHI applications?</td>
<td></td>
</tr>
<tr>
<td>4) What critical information regarding the impact of CHI applications is needed to give consumers, their families, clinicians, and developers a clear understanding of the value proposition particular to them?</td>
<td></td>
</tr>
</tbody>
</table>
articles that were potentially relevant to Key Question 1, we searched for terms relevant to our definition of CHI applications, combined with terms relevant to our definition of “consumer,” combined with terms identifying RCTs as the study design of interest. To identify articles that were potentially relevant to Key Question 2, we searched for terms relevant to our definition of CHI applications, combined with terms relevant to barriers; the search was not limited by study design. Eligible studies were also identified by reviewing the references in pertinent reviews and by querying our experts.

Our comprehensive search included electronic searching of peer-reviewed literature and gray literature databases as well as hand searching. On December 22, 2008, we conducted searches of the MEDLINE®, EMBASE®, The Cochrane Library, Scopus, and Cumulative Index to Nursing and Allied Health Literature (CINAHL) databases. This search was while review findings were reviewed by AHRQ and external experts to ensure we included the most current relevant articles; this search was extended to June 1, 2009. A supplemental search targeting gray literature sources was conducted on January 7, 2009; it was also extended to June 1, 2009. Sources searched were: Health Services Research Projects in Progress, Institute of Electrical and Electronics Engineers (IEEE) Conference Proceedings, Institution of Engineering and Technology (IET) Conference Proceeding, Proceedings of the American Society for Information Science and Technology (Wiley InterScience), World Health Organization (WHO)-International Clinical Trials Registry Platform, American Public Health Association 2000–2008, OpenSIGLE-System for Information on Gray Literature in Europe, and The New York Academy of Medicine—Gray Literature (see Tables 2 and 3).

Search strategies specific to each database were designed to enable the team to focus the available resources on articles that were most likely to be relevant to the Key Questions (see Table 1). We developed a core strategy for MEDLINE®, accessed via PubMed, on the basis of an analysis of the MeSH terms and text words of key articles identified a priori. The PubMed strategy formed the basis for the strategies developed for the other electronic databases.

Studies were eligible for inclusion in the review if they applied to Key Questions 1 or 2 and did not have one of the following reasons for exclusion: no health informatics application, health informatics application does not apply to the consumer, health informatics applications is for general information only (e.g., general Web site) and is not tailored to individual consumers, study of a “point of care” device (defined as requiring a clinician to use or

Key questions 3 (knowledge or evidence deficits) and 4 (critical information regarding CHI applications) are not included in this conceptual framework.

**Fig. 1** CHI conceptual model
<table>
<thead>
<tr>
<th>Database</th>
<th>Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMBASE</td>
<td>(&quot;Informatics&quot;:ti,ab OR telemedicine:ti,ab OR internet:ti,ab OR &quot;consumer health information&quot;:ti,ab) AND (&quot;consumer&quot;:ti,ab OR &quot;parents&quot;:ti,ab OR &quot;age groups&quot;:ti,ab OR caregivers:ti,ab) AND (&quot;randomized controlled trial&quot;:ti,ab OR (controlled:ti,ab AND trial:ti,ab) OR (clinical:ti,ab AND trial:ti,ab)) OR ((&quot;Informatics&quot;:ti,ab OR telemedicine:ti,ab OR internet:ti,ab OR &quot;consumer health information&quot;:ti,ab) AND (&quot;consumer&quot;:ti,ab OR &quot;parents&quot;:ti,ab OR parents[tiab] OR &quot;age groups&quot;:ti,ab OR caregivers:ti,ab) AND (access:ti,ab OR barrier:ti,ab OR facilitator:ti,ab OR compatibility:ti,ab OR incompatibility:ti,ab OR &quot;user centered&quot;:ti,ab OR &quot;work flow&quot;:ti,ab OR workflow:ti,ab OR &quot;reimbursement mechanisms&quot;:mh) OR reimbursement:ti,ab OR &quot;attitude to computers&quot;:mh OR &quot;health knowledge, attitudes, practice&quot;:mh OR computer literacy[mh] OR (computer:ti,ab AND literacy:ti,ab)) AND English NOT (editorial[pt] OR letter[pt] OR comment[pt]) NOT (&quot;animals&quot;[MeSH Terms] NOT &quot;humans&quot;[MeSH Terms]) AND ((&quot;1900/01/01&quot;[PDat] : &quot;2009/06/01&quot;[PDat])) AND ((&quot;1900/01/01&quot;[PDat] : &quot;2009/06/01&quot;[PDat])) AND (&quot;1900/01/01&quot;[PDat] : &quot;2009/06/01&quot;[PDat])</td>
</tr>
<tr>
<td>Cochrane library</td>
<td>(&quot;Medical Informatics applications&quot;:ti,ab,kw OR &quot;Informatics&quot;:ti,ab,kw or (telemedicine):ti,ab,kw or (internet):ti,ab,kw or &quot;Consumer Health Information&quot;:ti,ab,kw or &quot;Support systems&quot;:ti,ab,kw) AND ((&quot;consumer&quot;:ti,ab,kw or &quot;Parents&quot;:ti,ab,kw or (&quot;parents&quot;:ti,ab,kw) OR &quot;age groups&quot;:ti,ab,kw or (Caregivers):ti,ab,kw) AND (&quot;randomized controlled trial&quot;:ti,ab,kw or (controlled trial):ti,ab,kw) OR</td>
</tr>
<tr>
<td>SCOPUS</td>
<td>(TITLE-ABS-KEY(&quot;Medical Informatics applications&quot;) OR TITLE-ABS-KEY(telemedicine) OR TITLE-ABS-KEY(internet)) AND (TITLE-ABS-KEY(&quot;Consumer Health Information&quot;)) AND (TITLE-ABS-KEY(consumer) OR TITLE-ABS-KEY(&quot;Patients&quot;) OR TITLE-ABS-KEY(caregivers)) AND (TITLE-ABS-KEY(&quot;randomized controlled trial&quot;) OR TITLE-ABS-KEY(&quot;clinical trial&quot;)) OR ((TITLE-ABS-KEY(&quot;Medical Informatics applications&quot;) OR TITLE-ABS-KEY(telemedicine) OR TITLE-ABS-KEY(internet)) AND (TITLE-ABS-KEY(consumer) OR TITLE-ABS-KEY(&quot;Patients&quot;) OR TITLE-ABS-KEY(caregivers)) AND (TITLE-ABS-KEY(access) OR TITLE-ABS-KEY(facilitator) OR TITLE-ABS-KEY(&quot;user centered&quot;)) OR TITLE-ABS-KEY(&quot;computer literacy&quot;) OR TITLE-ABS-KEY(&quot;health knowledge, attitudes, practice&quot;)) AND (LIMIT-TO(DOCTYPE, &quot;ar&quot;) OR LIMIT-TO(DOCTYPE, &quot;re&quot;) OR LIMIT-TO(DOCTYPE, &quot;rp&quot;) OR LIMIT-TO(LANGUAGE, &quot;English&quot;))</td>
</tr>
<tr>
<td>CINAHL</td>
<td>(&quot;Medical Informatics&quot; OR TX telemedicine OR TX internet OR TX &quot;Consumer Health Information&quot; OR TX &quot;Support systems&quot;) AND (TX consumer OR TX &quot;Patients&quot; OR TX parents OR TX &quot;age groups&quot; OR TX Caregivers) AND (TX &quot;randomized controlled trial&quot; OR TX &quot;controlled trial&quot; OR TX &quot;clinical trial&quot;) OR ((&quot;Informatics&quot; OR TX telemedicine OR TX internet OR TX &quot;Consumer Health Information&quot; OR TX &quot;Support systems&quot;) AND (TX consumer OR TX &quot;Patients&quot; OR TX parents OR TX &quot;age groups&quot; OR TX Caregivers) AND (TX Access OR TX barrier OR TX facilitator OR TX compatibility OR TX incompatibility OR TX &quot;user centered&quot; OR TX &quot;work flow&quot; OR TX Reimbursement OR TX Attitude OR TX &quot;computer literacy&quot;) NOT (PT editorial) OR (PT letter) OR (PT comment))</td>
</tr>
</tbody>
</table>
obtain and is part of the regular provision of care), or no original data.

We assessed the eligible studies on the basis of the quality of their reporting of relevant data. For the RCTs, we used the study quality scoring system developed by Jadad et al. [8]. For the other studies, we used a form to identify key elements that should be reported when reporting results. The quality assessments were done independently by paired reviewers.

We then created a set of detailed evidence tables containing information extracted from the eligible studies. We stratified the tables according to the applicable Key Question and subquestion (for Key Question 1). We did not quantitatively pool the data for any of the outcomes because of the marked heterogeneity of target conditions of interest and the wide variety of outcomes studied. Data were abstracted by one investigator and entered into online data abstraction forms using SRS (Mobius Analytics, Inc., Ottawa, Ontario, CA). Second reviewers were generally more experienced members of the research team, and one of their main priorities was to check the quality and consistency of the first reviewers’ answers.

At the completion of our review, we graded the quantity, quality, and consistency of the best available evidence for each type of outcome in each clinical area, using an evidence grading scheme recommended by the GRADE Working Group and modified for use by the Evidence-Based Practice Centers Program [9]. For each outcome of interest, two investigators independently assigned a grade, and then the entire team discussed their recommendations and reached a consensus.

RESULTS

As shown in Fig. 2, the literature search process identified 24,794 citations that were deemed poten-
6,673 additional articles were identified relevant to Key Questions 1 and/or 2 and 6,673 additional articles were identified through hand searching. We identified no additional eligible articles in the gray literature. We excluded 8,943 duplicate citations from the electronic search results. Most duplicates came from concurrently searching MEDLINE®, The Cochrane Library, EMBASE®, CINAHL, and SCOPUS. The search strategy used in all search engines was modeled after that which we used in MEDLINE®, with similar search terms. Additionally, the EMBASE® search engine allows the user to search the MEDLINE® database as well as EMBASE®, a strategy that often yields many duplicates between the two search sites but improves the sensitivity of the search.

In the title review process, manuscript titles were reviewed by two investigators to preliminarily assess manuscript relevance. We excluded 19,377 citations that clearly did not apply to the Key Questions. In the abstract review process, we excluded 2,642 citations that did not meet one or more of the eligibility criteria. At the article review phase, we excluded an additional 357 articles that did not meet one or more of the eligibility criteria. Two more articles were removed from the pool of articles because the articles could not be located through any of the cooperating libraries or the manuscript authors. Ultimately, we were left with 146 articles that were eligible for inclusion in this report: One hundred and twenty-one for Key Question 1 and 31 for Key Question 2; six articles were eligible for both Key Questions 1 and 2. The complete list of publications has been published in the full report [7].

In terms of types of applications studied, 55% of studies evaluated interactive web-based applications or tailored educational websites. Another 15% of studies evaluated computer-generated tailored feedback applications. Interactive computer programs and personal monitoring devices were evaluated in approximately 8% of studies each. Finally, health risk assessments, decision aids, discussion or chat groups, and computer-assisted imagery were evaluated in less than 5% of studies each.
Ninety-nine studies reported user age, 77% (76/99) of those studies reporting age of participants targeted adults, approximately 12% of studies targeted adolescents, 3% of studies targeted seniors, and another 3% of studies targeted children. Five percent of studies targeted participants from overlapping age groups. Among studies reporting the race of the participants (n=53), 92% (49/53) of the studies employed populations that were greater than 50% Caucasian. There was only one study with greater than 50% African-American participants and no studies with a majority of participants who were Hispanic, American Indian/Alaska Native, or Asian/Pacific Islander. Fifty-eight percent of studies reporting delivery location evaluated CHI applications that were used in the home or residence. A minority of evaluations were completed in schools (15%), clinical settings (17%), communities (3%), online (5%), or kiosks (2%).

The impact of CHI applications on health outcomes

The impact of CHI applications on healthcare process outcomes

There were only six studies that met the inclusion-exclusion criteria and thus were available to shed light on this question. Five of these studies focused on asthma and one additional study focused on contraceptive medication utilization. All of the asthma studies showed a significant positive effect of the CHI application on at least one healthcare process measure. The oral contraceptive medication use application failed to reduce contraceptive discontinuation. No study found any evidence of harm.

The impact of CHI applications on intermediate health outcomes (Key Question 1b)

This review identified 108 studies that addressed the influence of CHI applications on intermediate health outcomes in the context of nine categories of diseases or health conditions. These were breast cancer in three studies, diet, exercise, and physical activity (not obesity) in 32 studies, alcohol abuse in seven studies, smoking cessation in 19 studies, obesity in 11 studies, diabetes mellitus (or diabetes with associated conditions) in seven studies, mental health in eight studies, asthma/chronic obstructive pulmonary disease (COPD) in four studies, and miscellaneous health conditions in another 15 studies.

With regard to breast cancer, evaluated intermediate outcomes included self-management, knowledge, program adherence, and change in health behaviors. Eighty-nine percent of these studies demonstrated significant positive effect on at least one intermediate health outcome related to diet, exercise, and physical activity. No study found any evidence of harm.

Evaluated intermediate outcomes related to alcohol abuse included self-management, knowledge attainment, and change in health behaviors. All seven studies found significant positive effect on at least one intermediate outcome related to alcohol abuse. No study found any evidence of harm.

With regard to smoking cessation, intermediate outcomes assessed in these smoking cessation CHI trials included self-management, knowledge attainment, and change in health behaviors. Fifty-seven percent of these studies demonstrated a positive effect on at least one intermediate outcome related to smoking cessation. No study found any evidence of harm.

Evaluated intermediate outcomes of interest related to obesity included weight loss behaviors and body composition. Only 36% of studies demonstrated positive effect on intermediate outcomes related to obesity. No study found any evidence of harm.

Seven studies were identified to evaluate the influence of CHI on intermediate outcomes related to diabetes mellitus. Intermediate outcomes of interest included perceived self-efficacy, satisfaction, and readiness to change, perceived competence, exercise minutes per day, and self-reported global health. All seven studies found evidence of effect of CHI applications on one or more intermediate outcomes related to diabetes mellitus. No study found any evidence of harm.

Eight studies were identified to evaluate the effect of CHI applications on intermediate outcomes related to mental health issues. Intermediate outcomes of interest included work and social adjustment, perceived stress, self-rated self-management, sleep quality, mental energy, and concentration. Seven of the eight studies found evidence of positive effect of CHI applications on at one or more intermediate outcomes related to mental health. No study found any evidence of harm.

Four studies were identified to evaluate the effect of CHI applications on intermediate outcomes related to asthma/COPD. Intermediate outcomes of interest included adherence, knowledge, change in behavior, dyspnea knowledge, and self-efficacy. Only one of the four studies demonstrated a significant effect on any intermediate outcome related to asthma/COPD. No study found any evidence of harm.

Two studies were identified to evaluate the effect of CHI applications on intermediate outcomes related to menopause or hormone replacement therapy (HRT). Only one study found evidence of significant effect on an intermediate outcome related to menopause/HRT utilization.
Finally, an additional 15 studies were identified to evaluate the influence of intermediate health outcomes in other clinical areas. These intermediate outcomes were in health areas related to arthritis, back pain, behavioral risk factor control, contraception, cardiovascular disease, cancer, caregiver decision making, fall prevention, health behavior change, headache, HIV/AIDS, and adolescent risk behaviors. Each of these studies found evidence of significant effect of the CHI application on intermediate outcomes related to the health condition under study. No study found evidence of harm.

The effect of CHI applications on relationship-centered outcomes (Key Question 1c)
Eight studies were identified that met the inclusion–exclusion criteria. Relationship centered outcomes of interest included social support, quality of life, decision making skill, social support, positive interaction with the provider, and satisfaction with care. These relationship-centered outcomes were evaluated in the context of care for HIV/AIDS, cancer, osteoarthritis, and pregnancy. Five of eight studies demonstrated significant effect of CHI on at least one aspect of relationship-centered care. No study found any evidence of harm.

The impact of CHI applications on clinical outcomes (Key Question 1d)
Twenty-eight studies addressed this question in the context of care for cancer (three studies), diabetes mellitus (three studies), mental health (seven studies), diet, exercise, or physical activity [five studies], and Alzheimer’s disease, arthritis, asthma, back pain, aphasia, COPD, HIV/AIDS, headache, obesity, and pain [one study each]. Over 80% of the studies found significant influence of CHI applications on at least one clinical outcome.

Three studies evaluated the effect of CHI applications on breast cancer clinical outcomes, but only one found any evidence of significant CHI impact. Of the five studies that evaluated the effect of CHI applications on clinical outcomes related to diet, exercise, or physical activity, four studies found a significant positive effect on one or more clinical outcomes. Among the seven studies that evaluated the effect of CHI applications on mental health clinical outcomes, all seven found evidence of significant effect of CHI on one or more clinical outcomes. Three studies evaluated the effect of CHI applications on diabetes mellitus clinical outcomes. All three studies found evidence of significant effect of CHI on at least one clinical outcome. The remaining nine studies evaluated a CHI application in different health areas including Alzheimer’s disease, arthritis, asthma, back pain, aphasia, COPD, headache, HIV/AIDS, and general pain. With the exception of the general pain study, the eight remaining studies all found evidence of significant effect of CHI on one or more clinical outcomes. None of these 27 studies found any evidence of harm attributable to a CHI application.

The impact of CHI applications on economic outcomes (Key Question 1e)
Three studies addressed this question. Economic outcomes evaluated in these studies included cost of program delivery, cost of computer information system with manual data extraction versus cost of the computer system with use of an electronic patient record, materials costs, total costs, and incremental cost effectiveness. These outcomes were evaluated in the context of care for asthma, cancer, and obesity. Each of these studies used different economic metrics and methodologies. One study failed to provide any cost estimates for the control group. One study was done in an adult population, another in a pediatric population, and the third study did not provide any details regarding the age of study participants.

Barriers that limit utilization or implementation of CHI applications for clinicians, developers, consumers, and their families or caregivers encounter
Thirty-one studies addressed the question of barriers to CHI application use. Studies focused on a wide variety of clinical conditions including cancer, HIV/AIDS (and sexually transmitted disease), mental health, physical activity/diet/obesity, smoking cessation, prostate cancer, and hypertension. Because CHI applications involve the participation of consumers, their caregivers, may also include clinicians, and developers, this analysis included barriers that impede participation of any of the above groups. Identified barriers were grouped into healthcare system-level barriers and individual-level barriers. Six studies addressed systems-level barriers including Internet access at home or in the community and all six found this to be a barrier. One study identified hardware requirements and another study identified mobile device shape, design, or configuration as a systems-level barrier to use. Another five studies cited CHI tool incompatibility with current healthcare as a barrier.

Identified individual-level barriers included clinic staff who feared increased workloads, lack of built-in social support, forgotten passwords, automated data entry inability to allow for back entry of old data, lack of adequate user customization, and substantial financial investment. Nineteen studies queried application usability or user-friendliness and all 19 found evidence of lack of usability as a barrier to use. Eleven studies explored how patient knowledge, literacy, and computer skills could impact the use the CHI application. Ten found deficits in knowledge, literacy, or computer skills to be barriers. Six studies considered the possibility that users would find the application too time-consuming. Five of
these studies simply cited this possibility in the “RESULTS” section of the manuscripts, while the one of these studies actually reported that “too many emails to participants” was found to be a barrier.

Utilization fees were also identified as a barrier. Five studies sought information about privacy concerns and four reported concerns over privacy as a barrier. These studies also found concerns over the control of information or lack of trust in the technology to be barriers. Only two studies queried for potential cultural barriers to use and only one study found evidence of this. The expectations of consumers including acceptability, usefulness, credibility, expectations, and goals were found to be barriers in eight studies. Cost was mentioned as a barrier in only one study, and only one study found evidence that physical or cognitive impairment resulted in barriers to the use of CHI applications. Finally, anxiety over the use of computers, complaints about lack of personal contact with clinicians, and the belief that health IT would not be an improvement to current care were mentioned in two studies as barriers.

Knowledge or evidence needs to support estimates of cost, benefit, and net value with regard to consumer health informatics applications

The identified studies indicate that the available literature is at a very early stage of development. Many questions have only been evaluated by one study. Thus, confirmatory studies have generally not been done. In addition, no high-quality studies have been conducted regarding several important questions. Broadly, these questions can be grouped into at least one of four categories: patient-related questions; CHI utilization factors; technology-related issues (i.e., hardware, software, and platform related issues and health-related questions); and health-related questions.

**Patient-related questions**—The literature is relatively silent on the question of whether or not significant differences in patient preferences, knowledge, attitudes, beliefs, needs, utilization, and potential benefits exists across gender, age, and race/ethnicity. Beyond these demographic differences, the field of CHI is developing within the context of a global emergence of technology-based realities including the emergence of Web 2.0/Web 3.0 and ubiquitous computing, which are enabling an unprecedented level of user-determined content, interactivity, and functionality. The degree to which this functionality could be harnessed for the health benefit of consumers is unknown. The targeted uses of CHI applications must increasingly be focused on more than just the index patient. The role of sociocultural and community factors will likely exert significant effect on access, usability, desirability, and benefit of CHI applications. Issues related to trust, security, and confidentiality need to be further explored. Because the bulk of the currently available research has been conducted on the 18- to 65-year-old adult population, more work needs to be done among the populations that may have the most potential for using CHI applications. Seniors may stand to benefit from those applications that reduce social isolation to independence.

Adolescents are some of the most intense technology users. Their natural affinity for technology may prove advantageous to CHI applications that could be developed in the future. Finally, most of the current CHI research is being conducted among predominately white/Caucasian populations. Early evidence suggests that differential utilization patterns and preferences exist by race [10, 11]. Such differences could potentially lead to differential efficacy of emerging CHI applications. This could have the unintended consequence of enhancing rather than reducing some racial and ethnic disparities in healthcare. Age, gender, and race/ethnicity subgroup differences need to be better understood and those differences incorporated into the development of emerging applications to ensure efficacy among all population subgroups.

**CHI utilization factors**—Despite a rapid increase in access to broadband services among all population groups, age groups, and geographic regions of the country, differential access to broadband Internet access may have significant implications in terms of health benefits that may be derived from these tools to applications. While many in the younger generations become very technically savvy at an early age, many Americans still have limited computer literacy. These CHI utilization factors suggest the need for a more robust evaluation of the epidemiology of broadband access and technology literacy in the USA [11, 12].

**Technology-related issues**—The majority of CHI applications are designed for use on personal computers as web-based applications. Many more potential types of platforms exist that have not been evaluated. In addition, emerging evidence is suggesting that the CHI applications and functionality that consumers want and need are not always what healthcare practitioners think they need. As a result, important human factors considerations (graphics vs. text based interfaces, mobile vs. desktop devices) may not get incorporated into emerging CHI applications and, therefore, lead to CHI applications with limited efficacy.

**Health-related questions**—Finally, most CHI applications that have been evaluated tend to focus on one or more domains of chronic disease management. Insufficient attention has been given to the role of CHI applications in addressing acute health problems. The role of CHI applications in primary, secondary, and tertiary prevention also needs to be more adequately explored. Sociocultural factors are increasingly important determinants of healthcare outcomes. The potential influence on social factors including social isolation and social support and perhaps even broader social determinants of health...
need to be evaluated and may prove useful in helping consumers address specific health concerns in the home and community-based setting.

Information needed to give consumers, their families, clinicians, and developers a clear understanding of the value proposition of CHI particular to them

Several critical information needs must be addressed to enable a clear understanding of the value proposition of CHI applications. It is likely that the knowledge gaps needed to establish a value proposition, while overlapping, are not identical across all potential stakeholders. Because providers are often most concerned about clinical outcomes and costs, it seems reasonable that questions of the impact of CHI applications on provider or healthcare processes, costs, and outcomes as addressed in this report will need to be more definitively characterized. In addition, the potential liability a provider might incur from a patient using a CHI application will also need to be addressed.

Patients often cite convenience and anonymity as the primary reasons the Internet has become such a major source of health information [10]. It is likely that the more these elements can be incorporated into emerging CHI applications, the more likely they will be considered of value by consumers. Other related factors such as usability, portability, and patient-centered functionality are likely important characteristics of CHI applications that may help drive utilization. Those technologies that exist and enable consumers to accomplish tasks (empower) without further complicating individuals’ lives may ultimately prove to be the most widely valued CHI applications. By expanding the number of and types of platforms available to consumers, CHI applications may become more appealing to a broader consumer base and thus prove valuable to those consumers who could most benefit but may not otherwise use a more traditional CHI application.

DISCUSSION

The results of this SER suggest several emerging themes. First, there may be a role for CHI applications in reaching consumers at a low cost and also in obviating the need for some activities currently performed by professionals. In addition, the data suggest that CHI applications may also be used to enhance the efficacy of interventions currently delivered by professionals. Several studies compared the use of a CHI application and traditional therapy against traditional therapy alone and found that the group receiving traditional therapy with a CHI application had more benefit than traditional therapy alone. Thirdly, the studies evaluated in this review tended to support the finding that at least three critical elements are most often found in effective CHI tools and applications including [7] individual tailoring, [1] personalization, and [2] behavioral feedback. Personalization involves designing the intervention to be delivered in a way that makes it specific for a given individual. Tailoring refers to building an intervention in part on specific knowledge of actual characteristics of the individual receiving the intervention. Finally, behavioral feedback refers to providing consumers with messages regarding their progression through the intervention. Interestingly, it is not clear from this literature that CHI-derived behavioral feedback is any better than feedback originating from human practitioners or others. Rather, it appears that the feedback must happen with an appropriate periodicity, in a format that is appealing and acceptable to the consumer.

Finally, despite the previously cited limitation of the available scientific literature, the body of the available scientific evidence suggests that CHI applications may hold significant future promise for improving outcomes across a variety of diseases and health issues. In terms of healthcare processes and relationship-centered outcomes, the literature is positive but very limited. Most of the currently available research has evaluated the impact of CHI applications on intermediate health outcomes. The literature appears strongest for CHI applications targeting intermediate outcomes related to smoking cessation. In terms of clinical outcomes, the weight of the evidence appears strongest for the use of CHI applications on mental health outcomes. Evidence-based conclusions regarding economic outcomes cannot be made at this time.

Study limitations

This review has several important limitations. First, our initial search for eligible studies proved to be challenging because of inconsistent use of terminology in the literature. We minimized this problem by searching multiple databases and supplementing our search with a review of selected journals and querying experts. The most important limitation was marked heterogeneity of interventions, populations, and outcomes, making synthesis across studies difficult and precluding meta-analysis. Inconsistent definitions and reporting of outcome measures further limited our ability to synthesize data, as many studies did not report enough data to support calculation of effect sizes. Methodologic (limited sample size, randomization scheme not specified, poor adjustment for potential confounders, etc.) limitations of many of the RCTs limit the strength of conclusions. Usually because of the relatively small number of available studies, but sometimes also due to variability in the quality of available studies, the strength of the body of evidence was often graded as low. Finally, there are several ongoing CHI studies that have not yet reported. This evidence report may need to be updated when the results of these studies are available.
Acknowledgments: This project was funded under contract No. HHSA 290-2007-10061-I from the Agency for Healthcare Research and Quality (AHRQ), US Department of Health and Human Services. The opinions expressed in this document are those of the authors and do not reflect the official position of AHRQ or the US Department of Health and Human Services.

Age at Menarche, Schooling, and Sexual Debut in Northern Malawi

Judith R. Glynn1,*, Ndoliwe Kayuni2, Sian Floyd1, Emmanuel Banda2, Monica Francis-Chizororo1,2, Clare Tanton3, Anna Molesworth1, Joanne Hemmings1, Amelia C. Crampin1,2, Neil French1

1 Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom, 2 Karonga Prevention Study, Chilumba, Malawi, 3 Centre for Sexual Health and HIV Research, Research Department of Infection and Population Health, University College London, London, United Kingdom

Abstract

Background: Age at sexual debut is a key behavioural indicator used in HIV behavioural surveillance. Early age at menarche may precipitate early sex through perceived readiness for sex, or through school drop-out, but this is rarely studied. We investigated trends and circumstances of sexual debut in relation to schooling and age at menarche.

Methods and Findings: A cross-sectional sexual behaviour survey was conducted on all individuals age 15–59 within a demographic surveillance site in Karonga District, Malawi. Time trends were assessed using birth cohorts. Survival analysis was used to estimate the median age at menarche, sexual debut and first marriage. The 25th centile was used to define "early" sex, and analyses of risk factors for early sex were restricted to those who had reached that age, and were done using logistic regression. Of the 8232 women and 7338 men resident in the area, 88% and 78%, respectively, were seen, and, 94% and 92% of these were interviewed. The median reported age at first sex was 17.5 for women and 18.8 for men. For women, ages at menarche, sexual debut and first marriage did not differ by birth cohort. For men, age at sexual debut and first marriage decreased slightly in later birth cohorts. For both men and women increased schooling was associated with later sexual debut and a longer delay between sexual debut and first marriage, but the associations were stronger for women. Earlier age at menarche was strongly associated with earlier sexual debut and marriage and lower schooling levels. In women early sexual debut (<16 years) was less likely in those with menarche at age 14–15 (odds ratio (OR) 0.31, 95%CI 0.26–0.36), and ≥16 (OR 0.04, 95%CI 0.02–0.05) compared to those with menarche at <14. The proportion of women who completed primary school was 46% in those with menarche at <14, 60% in those with menarche at 14–15 and 70% in those with menarche at ≥16. The association between age at menarche and schooling was partly explained by age at sexual debut. The association between age at menarche and early sex was not altered by adjusting for schooling.

Conclusions: Women with early menarche start sex and marry early, leading to school drop-out. It is important to find ways to support those who reach menarche early to access the same opportunities as other young women.

Introduction

Age at first sexual intercourse is one of the key sexual behaviour indicators recommended in second generation HIV surveillance [1]. Age at sexual debut has been shown to correlate with subsequent risk behaviour: on average those with younger ages at debut have more partners [2] and a higher risk of HIV [3]. Another important indicator of risk is the delay between sexual debut and first marriage, as this can be a time of high partner change, and is also associated with a higher number of partners later [4].

Largely missing from the discussion on risk factors for early sexual intercourse, at least in the biomedical literature, is age at menarche. Yet an association between early menarche and early sex has been noted where it has been measured [5,6,7]. Earlier menarche may lead to earlier sex because of the girl’s desires but also due to social pressures and expectations. Girls who reach menarche may be regarded as “ready” to start sex and marry [8]. And for those who do start sex, pregnancy is more likely if they are physically mature.

There is particular interest in the role of schooling in sexual behaviour change [9,10]. The association between level of schooling achieved and sexual debut is complex. Schooling is a measure of socio-economic status and of education itself. Socio-economic conditions which may limit school access or contribute to poor performance may also pre-dispose to early sexual activity [11]. And poor achievement may lead to drop out and sexual activity [11,12]. For young women there is an additional complexity, since pregnancy may lead to school exclusion [11,12]. Menarche may also precipitate early sexual debut through its affects on schooling. Menstruation can make school attendance difficult where sanitary arrangements are limited, leading to poor grades and school drop-out [13,14].
Using data from a cross-sectional survey in a demographic surveillance site in northern Malawi we assessed secular trends in sexual debut and characteristics of first partners, and the links between menarche, schooling and sexual debut.

Methods

The study was conducted as part of the Karonga Prevention Study in Karonga District, Malawi. A demographic surveillance system was set up in 2002 in a rural population of about 32,000 [15]. A biennial census started in 2004, which was replaced by annual census rounds, together with a socio-economic status update and HIV serosurveillance in 2007. A sexual behaviour survey was started in 2008, including all individuals aged 15–59. HIV prevalence in the area rose from less than 2% in the late 1980s to around 10% now [16].

Ethics statement: Ethics approval for the study was received from the Health Sciences Research Committee, Malawi, and the ethics committee of the London School of Hygiene & Tropical Medicine, UK. Before the start of the study the Traditional Authority that covers the area, and all village headmen and traditional advisors in the study area were informed about the aims of the study and the nature of the data to be collected, and their approval and verbal consent was sought. All household members were given a similar explanation and interviews were only conducted if verbal consent was given by the household head and by the respective household members. The consent for the demographic surveillance was recorded by the interview sheet being filled. Refusals were recorded in field registers. During the baseline census 15 households did not provide verbal consent and have consequently been excluded from the study. The socio-demographic data for this study come from the basic demographic surveillance for which the ethics committees agreed that written consent was not needed. For the sexual behaviour survey individual written consent was sought.

Schooling level achieved was asked for all individuals. In those aged up to 30 years at the time of interview, questions about schooling included the reason for leaving (asked as an open question: more than one reason could be recorded). Socio-economic status was only available at the time of interview, not historically. Parental education level, available for those aged up to 30 years at the time of interview, was used as a proxy of socio-economic status in adolescence.

Whether sexual debut occurred before menarche was asked throughout the study. A question on age at menarche was added in mid-October 2008, so was only available on about half of the women. Other questions included age at first intercourse, and information on characteristics of the first sex partner, and age at first marriage.

Statistical analysis

Data from the first round of the sexual behaviour survey were used. Median age at menarche, first sexual intercourse and first marriage were determined using survival analysis, to allow for right censoring of those who had not yet experienced these events. The data were smoothed by adding a random fraction of a year, since age in whole years was recorded for events, not dates [17,18]. The 25th centile of age at sexual debut was used to define “early” sex, and subsequent analyses of determinants of early sex only included individuals over these cut-off ages, and was done using logistic regression. A delay between sexual debut and marriage was defined as “long” if the age at first marriage (or the current age for those not yet married) was more than one year older than the age at first sex, and analysis was restricted to individuals more than one year older than their age at first sex. This period of delay was chosen to distinguish marriage happening soon after, and perhaps related to, sexual debut from that happening later.

The analysis explored secular trends (using birth cohorts, <1965, 1965–74, 1975–84, 1985–94) in ages at first sexual intercourse and first marriage and in partner types; and risk factors for early sex and for a long delay between first sex and first marriage. For women, analyses also explored the relationship between age at menarche, schooling and sexual debut. In particular we assessed whether an association between age at menarche and sexual debut was affected by adjusting for schooling (which would suggest that schooling was on the causal pathway between menarche and sexual debut); and whether an association between age at menarche and schooling was affected by adjusting for sexual debut and marriage (which would suggest that an association between menarche and schooling was mediated via sexual debut and/or marriage).

Results

At the time of the survey there were 8232 women and 7338 men aged 15–59 resident within the demographic surveillance area. 987 (12%) women and 1613 (22%) men were not found and seen by the interviewers. Of those who were seen 6825 (94%) women and 5283 (92%) men agreed to be interviewed about their sexual behaviour, and 6796 women and 5253 men were interviewed.

At the time of interview 89% of the women had ever had sexual intercourse, and 85% had ever been married. Equivalent figures for the men were 80% and 62%. Median age at first sexual intercourse for the women was 17.5, and for first marriage was 18.5 years. Using birth cohorts, both values were unchanged over time (figure 1). For the men the median age at first sexual intercourse was 18.8 years, and at marriage 23.7 years: both sexual debut and marriage occurred at slightly younger ages in the more recent birth cohorts (figure 1).

Using the 25th centile as the cut-off, early sex was defined as sexual debut at <16 years for women and <17 years for men.

Women

There was a strong association between age at menarche and age at sexual debut, with 55% of those with menarche at <14 years having had early sex, compared with 27% of those with menarche at 14 or 15, and 4% of those with menarche aged ≥16 years (table 1). Sexual debut before menarche was unusual: it was reported by 2.8% of women overall, but the proportion increased over time (figure 1). The interval between age at menarche and sexual debut was 3.5 years for those with menarche age <14, 2.7 years for those with menarche at 14–15, and 2.5 years for those with menarche at 16 or older. The association between age at menarche and early sex was not changed by adjusting for schooling or birth cohort (table 1).

Age at menarche was similar in each birth cohort (median 15.1 years, figure 1). The interval between age at menarche and age at sexual debut was also similar in each birth cohort: median 2.8, 3.0, 2.9, 2.8 years for the 4 birth cohorts. There was weak evidence of effect modification between the associations of age at menarche and birth cohort with age at sexual debut (p for interaction 0.2). In those with menarche at 14 years or older there was no association between birth cohort and early sex, but in those with early menarche, the likelihood of early sex increased with later birth cohort (48.1%, 47.3%, 57.1% 60.4%, for the 4 birth cohorts respectively, p = 0.03).
Early sex was more common in those with less schooling. It was also more common in those whose parents had had less schooling (only asked for those aged 30 and under). Although overall there was no association between birth cohort and early sex, after adjusting for schooling level, early sex was more likely in those born more recently. There was no evidence of effect modification: within each level of schooling, there was a trend towards earlier sex in the younger cohorts. In those aged ≤30 years, the association of schooling and early sex was not altered by adjusting for parental schooling level. In this age group, 45% gave pregnancy or marriage as a reason for leaving school, and 33% were still in school. Those at higher levels of schooling were more likely to give pregnancy or marriage as the reason for leaving than those with more basic schooling. Menstruation was never mentioned as a reason for leaving school.

Among the oldest women 73% said their first sex partner was their husband, compared to 42% of the youngest women (table 2). The proportion reporting that their first partner was a boyfriend increased over time, from 25% to 56%. Many women later married this boyfriend but among ever married women there was still a decrease by birth cohort in the proportion who married their first sex partner (table 2). Condom use with the first partner,
excluding those whose first partner was their husband, rose from 0% in the oldest women to 41% in the youngest cohort.

A quarter of the women delayed more than one year between sexual debut and marriage. Delaying more than a year was less common in those born before 1965, but was similar in the other birth cohorts. Delay was more common in those with more schooling and in those with early sexual debut, and all these associations persisted after adjusting for each other (table 3). The estimates were similar when also adjusted for age at menarche, and, in those 30 and under, when adjusted for parental education. Delay was longer in those with later menarche, and this association was stronger after adjustment for age at sexual debut and weaker after adjustment for schooling.

There was a strong association between age at menarche and school level achieved (table 4). Standard 8, the end of primary school, was reached by 46% (383/836) of women with menarche, 14 years, 60% (1024/1713) of those with menarche at 14 or 15, and 70% (476/680) with menarche at ≥16 years, giving odds ratios, compared to those with menarche at <14, of 1.7 (1.5–2.0) and 2.7 (2.2–3.4) for those with menarche at 14/15 and ≥16 years respectively. This association was unchanged when adjusted for birth cohort, but was partly explained by adjusting for age at sexual debut (aOR 1.5 (1.2–1.8) for age at menarche of 14/15 and aOR 1.7 (1.3–2.1) for age at menarche ≥16, compared with those with menarche at <14 years old), and further when also adjusted for age at first marriage (aOR 1.3 (1.1–1.6) and 1.3 (1.0–1.8) for age at menarche of 14/15 and ≥16 respectively, table 4).

Men

Early sex (before age 17) was less common in the oldest cohort and in those with secondary or more schooling (table 1). These associations did not change when adjusted for each other. Among men aged 30 and under, 8% reported that they had left school because of marriage or because their girlfriend or wife was pregnant. The proportion of men who reported that their first partner was their wife decreased from 29% in the oldest group to 6% in the youngest (table 2). Most of the other first partners were described as girlfriends, and there was only a slight decrease over time in the proportion who married their first sex partner (from 38% to 33%). Condom use with the first partner (excluding spouses) rose to 54% in the youngest cohort.

Most men (75%) delayed more than a year between sexual debut and first marriage (table 3). Delay was more common in the more recent cohorts and among those with more schooling, and much more common (96%) in those with early sexual debut. Among men aged 30 and under, there was no association between parental education and early sex (table 1), but those whose fathers had had secondary schooling were more likely to delay between sexual debut and first marriage (table 3). Adjusting for parental education made little difference to the associations.
between a man's own schooling and either early sex or the delay between sexual debut and first marriage.

**Discussion**

For women there was little change over calendar time in age at menarche, sexual debut or first marriage. For men, there was a slight decrease in age at debut and first marriage, and an increase in delay between sexual debut and marriage. Changes over time were more marked in the type of first partner, with a decrease in the proportion reporting that sexual debut occurred within marriage for both men and women. Condom use with the first partner, where this was not the spouse, rose to 54% and 41% for the youngest groups of men and women. Using these cross-sectional data, secular trends can only be examined using birth cohorts. These will under-represent the higher risk individuals in older age groups, some of whom will have died of HIV. This may account for the downward shift in age at sexual debut for men, and obscure any trends for women.

The results relied on recall of age at events. Several studies have examined the accuracy of recall of age at sexual debut and first marriage by comparing results from the same individual over survey rounds. These have shown that 30–50% of reports are inconsistent, but there are no particular trends towards under or over estimation of ages, so aggregate trends can only be examined using birth cohorts. These will under-represent the higher risk individuals in older age groups, some of whom will have died of HIV. This may account for the downward shift in age at sexual debut for men, and obscure any trends for women.

The extent to which education influences behaviour or behaviour influences education is unclear, but it is likely that the effect is in both directions. Nearly half the women (and 8% of the men) aged 30 and under gave pregnancy or marriage as the reason for leaving school. This is a high percentage compared with other reports from sub-Saharan Africa [11]. From antenatal clinic (ANC) surveys in the same setting, 31% of teenage women attending ANC with their first pregnancy were at school when they became pregnant. 8% of teenage women attending ANC with subsequent pregnancies had also been at school when they became pregnant, confirming that some women are able return to education after the birth of a child (unpublished data).

For both men and women, those who achieved a higher level of schooling had later sexual debut and a longer delay between sexual debut and first marriage. The associations were much stronger for women. Schooling is also a marker of socio-economic status. We did not know socio-economic status at the time of sexual debut or marriage, but the associations found with an individual’s own schooling were not altered by adjusting for parental education status, which is a proxy for socio-economic status early in life. While the level of parental schooling may not be very accurately known or recalled, and thus there could be residual confounding, this suggests that the associations are with schooling itself.

The extent to which education influences behaviour or behaviour influences education is unclear, but it is likely that the effect is in both directions. Nearly half the women (and 8% of the men) aged 30 and under gave pregnancy or marriage as the reason for leaving school. This is a high percentage compared with other reports from sub-Saharan Africa [11]. From antenatal clinic (ANC) surveys in the same setting, 31% of teenage women attending ANC with their first pregnancy were at school when they became pregnant. 8% of teenage women attending ANC with subsequent pregnancies had also been at school when they became pregnant, confirming that some women are able return to education after the birth of a child (unpublished data).

For women, the onset of menstruation may be a major – and neglected – factor influencing both schooling and sexual debut. Menstruation marks physical maturity, the transition to womanhood, and, in some cultures, is marked by initiation rituals. In many and diverse settings girls are seen as “ready” for sexual activity soon after menarche (and boys as “needing” sexual activity once pubescent) [20]. Menstruation also brings practical problems with school attendance in managing the blood loss where latrine

---

**Table 2.** Characteristics of sexual debut by birth cohort.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Women</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Before menstruation</td>
<td>N 997</td>
<td>1202</td>
<td>2001</td>
<td>1551</td>
<td></td>
</tr>
<tr>
<td>% (n)</td>
<td>1.3 (13)</td>
<td>2.6 (31)</td>
<td>2.6 (52)</td>
<td>4.1 (64)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Type of partner</td>
<td>N 1035</td>
<td>1276</td>
<td>2089</td>
<td>1624</td>
<td></td>
</tr>
<tr>
<td>Husband % (n)</td>
<td>73.3 (759)</td>
<td>59.2 (755)</td>
<td>48.4 (1010)</td>
<td>42.1 (683)</td>
<td></td>
</tr>
<tr>
<td>Boyfriend % (n)</td>
<td>24.5 (254)</td>
<td>39.2 (500)</td>
<td>49.6 (1036)</td>
<td>55.9 (907)</td>
<td></td>
</tr>
<tr>
<td>Other/unknown % (n)</td>
<td>2.1 (22)</td>
<td>1.7 (21)</td>
<td>2.1 (43)</td>
<td>2.1 (34)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Married 1st partner (if ever married)</td>
<td>N 1026</td>
<td>1253</td>
<td>2046</td>
<td>1422</td>
<td></td>
</tr>
<tr>
<td>% (n)</td>
<td>84.2 (864)</td>
<td>76.3 (956)</td>
<td>70.3 (1439)</td>
<td>71.2 (1012)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Married 1st partner (if ever married)</td>
<td>N 271</td>
<td>513</td>
<td>1063</td>
<td>933</td>
<td></td>
</tr>
<tr>
<td>% (n)</td>
<td>0 (0)</td>
<td>2.0 (10)</td>
<td>16.8 (179)</td>
<td>41.2 (384)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td><strong>Men</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of partner</td>
<td>N 733</td>
<td>886</td>
<td>1513</td>
<td>1085</td>
<td></td>
</tr>
<tr>
<td>Wife % (n)</td>
<td>29.1 (213)</td>
<td>21.4 (190)</td>
<td>15.5 (235)</td>
<td>6.0 (65)</td>
<td></td>
</tr>
<tr>
<td>Girlfriend % (n)</td>
<td>67.1 (492)</td>
<td>74.6 (661)</td>
<td>79.6 (1205)</td>
<td>86.4 (937)</td>
<td></td>
</tr>
<tr>
<td>Other/unknown % (n)</td>
<td>3.8 (28)</td>
<td>4.0 (35)</td>
<td>4.8 (73)</td>
<td>7.7 (83)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Married 1st partner (if ever married)</td>
<td>N 721</td>
<td>866</td>
<td>1324</td>
<td>332</td>
<td></td>
</tr>
<tr>
<td>% (n)</td>
<td>37.6 (271)</td>
<td>30.4 (263)</td>
<td>29.9 (396)</td>
<td>32.8 (109)</td>
<td>0.003</td>
</tr>
<tr>
<td>Married 1st partner (if ever married)</td>
<td>N 514</td>
<td>690</td>
<td>1262</td>
<td>1009</td>
<td></td>
</tr>
<tr>
<td>% (n)</td>
<td>0.7 (5)</td>
<td>5.0 (36)</td>
<td>40.8 (292)</td>
<td>53.5 (383)</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

doi:10.1371/journal.pone.0015334.t002
facilities are poor, disposable pads are unaffordable, and it is difficult to wash in private [14]. Girls may skip school rather than risk discovery [13,14]. Physical maturity can also bring unwanted attention and teasing from boys, adding to the problems of school attendance [8,13]. An association between having reached puberty and school drop-out has been found previously [21].

Early menarche could therefore lead to early sexual debut via school drop-out, or more directly, in response to individual and societal pressures. While there is no scope for intervention in the timing of menarche, there are possibilities for intervening on the consequences. It is therefore important to understand the dominant pathways. If the practical problems of menstruation and schooling are the key, then the emphasis should be on improving the facilities and attitude of schools [5,8,13]. There are already some initiatives to do that internationally (eg UNICEF [22]) and nationally (Lieza du Preez, personal communication). However, if the direct pathway, through individual and societal pressures, is prominent, reducing early sexual debut following menarche will also require a shift in expectations.

The median age at menarche of 15 years is the same as measured previously in northern Malawi [23]. In Karonga District, unlike southern Malawi [23,24], there are no initiation rituals. At menarche girls are traditionally sent to stay with an aunt or other female relative, for instruction, and it is likely to become known in the community. In the current analysis there was a strong association between age at menarche and school level reached, consistent with an effect of menstruation on schooling. But this association was much less strong after adjusting for early sexual debut, suggesting that it was the early sex that led to the lower schooling level, not primarily the menstruation itself. Conversely the association between age at menarche and early sexual debut was not changed by adjusting for schooling level. This suggests that the major pathway by which earlier menarche leads to earlier sexual debut is not through the effect on schooling.

Sexual Debut in Northern Malawi

| Table 3. Factors associated with delay of more than one year between sexual debut and marriage, restricted to those seen more than one year since sexual debut. |
|------------------|------------------|------------------|
|                  | Women            | Men              |
|                  | n/N      | %     | OR | aOR  | n/N      | %     | OR | aOR |
| All              | 152/15479 | 27.8  |    |     | 2809/3749 | 74.9  |    |     |
| Birth cohort     |            |        |    |     |            |        |    |     |
| <1965            | 150/988   | 15.2  | 1  |     | 438/706   | 62.0  | 1 |     |
| 1965–74          | 321/1244  | 25.8  | 1.9 | 1.6–2.3 | 598/844  | 70.9  | 1.5 | 1.2–1.8 |
| 1975–84          | 627/2034  | 30.8  | 2.5 | 2.0–3.0 | 1061/1427 | 74.4  | 1.8 | 1.5–2.2 |
| 1985–94          | 423/1213  | 34.9  | 3.0 | 2.4–3.7 | 712/772   | 92.2  | 7.3 | 5.4–9.8 |
| Schooling        |            |        |    |     |            |        |    |     |
| None/P1–5       | 170/970   | 17.5  | 1  |     | 250/367   | 68.1  | 1 |     |
| P6–7            | 327/1421  | 23.0  | 1.4 | 1.1–1.7 | 373/542   | 68.8  | 1.0 | 0.78–1.4 |
| P8              | 439/1675  | 26.2  | 1.7 | 1.4–2.0 | 844/1165  | 72.5  | 1.2 | 0.95–1.6 |
| Secondary+       | 534/1237  | 43.2  | 3.6 | 2.9–4.4 | 1271/1584 | 80.2  | 1.9 | 1.5–2.4 |
| Early sex        |            |        |    |     |            |        |    |     |
| No              | 814/3636  | 22.4  | 1  |     | 1444/2323 | 62.2  | 1 |     |
| Yes             | 699/1832  | 38.2  | 2.1 | 1.9–2.4 | 1309/1370 | 95.6  | 13.1 | 10.0–17.1 |
| Age at menarche  |            |        |    |     |            |        |    |     |
| <14             | 215/758   | 28.4  | 1  |     | 1260/1533 | 82.2  | 1 |     |
| 14–15           | 446/1399  | 31.9  | 1.2 | 0.97–1.4 | 143/164   | 87.2  | 1.5 | 0.92–2.4 |
| ≥16             | 192/575   | 33.4  | 1.3 | 1.0–1.6 | 2.1 (1.6–2.8) |
| Mother schooling|            |        |    |     |            |        |    |     |
| ≤Primary        | 745/2274  | 32.8  | 1  |     | 1260/1533 | 82.2  | 1 |     |
| Secondary       | 117/290   | 40.3  | 1.4 | 1.1–1.8 | 143/164  | 87.2  | 1.5 | 0.92–2.4 |
| Father schooling|            |        |    |     |            |        |    |     |
| ≤Primary        | 489/1634  | 29.9  | 1  |     | 922/1153  | 80.0  | 1 |     |
| Secondary       | 346/870   | 39.8  | 1.5 | 1.3–1.8 | 443/494 | 89.7  | 2.2 | 1.6–3.0 |

OR = odds ratio, aOR = adjusted odds ratio.
1Adjusted for early sex, birth cohort and schooling level.
2Adjusted for age and early sex.
doi:10.1371/journal.pone.0015334.t003
similar to that of boys in the community. That reached by girls with earlier menarche falls far short of this.

Age at puberty is falling in many societies [25]. It will become increasingly important to find ways to stop girls who reach menarche early from being disadvantaged for the rest of their lives.

**Acknowledgments**

We thank the Government of the Republic of Malawi for their interest in this Project and the National Health Sciences Research Committee of Malawi for permission to publish the paper.

We thank Isolde Birdthistle for comments on an earlier version.

**Author Contributions**

Conceived and designed the experiments: JRG JH CT SF ACC NF. Performed the experiments: NK EB CT MFC AM. Analyzed the data: JRG SF. Wrote the paper: JRG ACC NK SF EB MFC CT AM JH NF.

**Table 4.** Associations with age at menarche.

<table>
<thead>
<tr>
<th>Age at menarche</th>
<th>N</th>
<th>14–15</th>
<th>≥16</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>836</td>
<td>1713</td>
<td>680</td>
</tr>
<tr>
<td>P6–7 (%)</td>
<td>18.2</td>
<td>12.8</td>
<td>11.3</td>
</tr>
<tr>
<td>P8 (%)</td>
<td>35.7</td>
<td>27.4</td>
<td>18.7</td>
</tr>
<tr>
<td>Secondary+ (%)</td>
<td>27.6</td>
<td>34.5</td>
<td>30.9</td>
</tr>
</tbody>
</table>

**Age at sexual debut**

<table>
<thead>
<tr>
<th>N</th>
<th>median, IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>15.7 (14.6–17.8)</td>
</tr>
<tr>
<td>P6–7 (%)</td>
<td>17.3 (15.9–18.8)</td>
</tr>
<tr>
<td>P8 (%)</td>
<td>18.9 (17.9–20.6)</td>
</tr>
</tbody>
</table>

**Age at first marriage**

<table>
<thead>
<tr>
<th>N</th>
<th>median, IQR</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>16.9 (15.5–18.9)</td>
</tr>
<tr>
<td>P6–7 (%)</td>
<td>18.5 (16.9–20.1)</td>
</tr>
<tr>
<td>P8 (%)</td>
<td>20.3 (18.8–21.9)</td>
</tr>
</tbody>
</table>

**Reason for leaving school (if age ≥30)**

<table>
<thead>
<tr>
<th>N</th>
<th>ref</th>
<th>1.7 (1.5–2.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other (%)</td>
<td>2.7 (2.2–3.4)</td>
<td></td>
</tr>
<tr>
<td>In school</td>
<td>2.7 (2.1–3.3)</td>
<td></td>
</tr>
</tbody>
</table>

**Finished primary school**

<table>
<thead>
<tr>
<th>N</th>
<th>ref</th>
<th>1.4 (1.2–1.7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Other (%)</td>
<td>1.6 (1.2–2.1)</td>
<td></td>
</tr>
<tr>
<td>In school</td>
<td>1.4 (1.0–1.8)</td>
<td></td>
</tr>
</tbody>
</table>

**References**


Challenges in ethics, safety, best practices, and oversight regarding HIT vendors, their customers, and patients: a report of an AMIA special task force

Kenneth W Goodman,1 Eta S Berner,2 Mark A Dente,3 Bonnie Kaplan,4 Ross Koppel,5 Donald Rucker,6 Daniel Z Sands,7,8 Peter Winkelstein,9 for the AMIA Board of Directors

ABSTRACT

The current commercial health information technology (HIT) arena encompasses a number of competing firms that provide electronic health applications to hospitals, clinical practices, and other healthcare-related entities. Such applications collect, store, and analyze patient information. Some vendors incorporate contract language whereby purchasers of HIT systems, such as hospitals and clinics, must indemnify vendors for malpractice or personal injury claims, even if those events are not caused or fostered by the purchasers. Some vendors require contract clauses that force HIT system purchasers to adopt vendor-defined policies that prevent the disclosure of errors, bugs, design flaws, and other HIT-software-related hazards. To address this issue, the AMIA Board of Directors appointed a Task Force to provide an analysis and insights. Task Force findings and recommendations include: patient safety should trump all other values; corporate concerns about liability and intellectual property ownership may be valid but should not over-ride all other considerations; transparency and a commitment to patient safety should govern vendor contracts; institutions are duty-bound to provide ethics education to purchasers and users, and should commit publicly to standards of corporate conduct; and vendors, system purchasers, and users should encourage and assist in each others’ efforts to adopt best practices. Finally, the HIT community should re-examine whether and how regulation of electronic health applications could foster improved care, public health, and patient safety.

INTRODUCTION AND BACKGROUND

The health information technology (HIT) industry, currently in the midst of extraordinary growth, actively transforms the way that we collect, store, use, and analyze health information. Correspondingly, the corporations that develop and sell electronic health record systems, associated devices, and health-related software applications face a complex suite of obligations—to patients, clinicians, shareholders, and society.

This article presents a report commissioned and approved by the AMIA Board of Directors. It briefly surveys the challenges that HIT vendors face; discusses the roles that ethics and related considerations can play in health informatics; and makes a number of recommendations regarding vendor contracts, ethics education, health information system user groups, best practices, marketing of health information systems, and regulation and oversight of the industry.

A concluding section, “Next Steps,” makes the case that the issues identified in the report deserve continued attention. The conclusion itemizes many issues that our field must address, ranging from stakeholder responsibilities and defect reporting, to meaningful use standards and unintended consequences. An appendix provides suggestions for further reading (see www.jamia.org).

Framing the issue

Developers and vendors of computer applications have enabled, fostered, and influenced the growth and evolution of HIT, and how healthcare organizations adopt and deploy it. Developers and researchers in academia, government, and industry have contributed substantially to improving information collection, analysis, transmission, and use in clinical care and research. Commercial vendors of HIT systems have an unprecedented dual opportunity to satisfy concurrently the demands of investors and markets while contributing to the improved health of individuals and of populations.

Complexity characterizes and makes matters difficult for a process that relies on competition to produce new, or at least better, devices and drugs. Not only must a corporation that sells HIT-related devices seek to turn a profit, but it must do so in competition with others. This competition hinges on quality, price, and other considerations of the marketplace. It is further shaped by other factors, requiring skillful management of liability and risks, increasing system complexity, public relations, industry standards, and best practices.

Some HIT vendors incorporate contract language that commits their customers to particular stances with respect to indemnity and error management. Specifically, hospitals and other purchasers of HIT systems are sometimes contractually obligated to indemnify vendors for malpractice or personal liability and intellectual property ownership may be valid but should not over-ride all other considerations; transparency and a commitment to patient safety should govern vendor contracts; institutions are duty-bound to provide ethics education to purchasers and users, and should commit publicly to standards of corporate conduct; and vendors, system purchasers, and users should encourage and assist in each others’ efforts to adopt best practices. Finally, the HIT community should re-examine whether and how regulation of electronic health applications could foster improved care, public health, and patient safety.
injury claims against hospitals or clinicians, even when those events are not caused or fostered by the purchasers. Some purchasers must contractually agree to adopt vendor-defined policies that prevent the disclosure of HIT system errors, design flaws, and other hazards. A publication laying out these concerns has aroused intense interest and discussion. Moreover, that report elicited renewed discussion about a major and longstanding issue in the HIT domain, namely, the extent to which the HIT industry should be subject to various additional kinds of oversight, regulation, or control, and by whom.

The AMIA Board of Directors appointed a task force in September 2009 to provide an assessment of these issues and to make recommendations to AMIA leadership—and, by extension, to the HIT community. This document contains the resulting analysis and recommendations.ii

**Ethical, legal, and social issues**

For-profit manufacturers of healthcare products are bound by values which may at times conflict. For instance, as entities in a marketplace, they are duty-bound to provide a financial return to those investors who have contributed resources in anticipation of their success. Yet, as developers and manufacturers of products that affect the health of people, they are no less obligated to ensure, to the extent possible, that their products are safe and effective, and beneficially support patients and those who treat and care for them.

The makers of HIT systems therefore confront many of the same challenges as counterparts in the pharmaceutical and medical device industries. Indeed, the intersection of corporate duty and patient-centered obligations involves two large areas of normative analysis, namely, business ethics and bioethics. Responsible companies foster an internal culture of ethics and at the same time recognize duties to customers and shareholders.

Ethics is the study of morality, or judgments, or standards about the rightness or wrongness of actions. Within specific professions, applied ethics serves to identify, clarify, and resolve moral issues, conflicts, and controversies that arise in professional practice. Applied ethics can teach or hone skills used to address these issues, especially when values are in conflict and when reasonable people (might) disagree. “Ethics” is not synonymous with “virtue.” For instance, information about a couple with discordant HIV statuses might be drawn from a database and used to warn the partner who is HIV-negative—or the information might be withheld to protect the privacy of the partner who is HIV-positive. Being a good or virtuous person will not help identify the correct action. In such situations, applied ethics provides tools for critical analysis and decision-making by professionals and others.

In civil society, morality and ethics guide the law. Ethics precedes the law. The reason murder, for instance, is illegal is because it is recognized to be wrong. It would be wrong even in the absence of a legal or criminal justice system. Ethics guides public policy in the same way. How, for instance, should businesses balance duties to shareholders, employees, and society? Applied ethics provides ways to answer such questions.

In health informatics, ethical issues address appropriate uses and users of decision-support systems, privacy and confidentiality, consent for secondary use of clinical and genetic information in databases, accountability or responsibility for errors, and so on. A thoroughgoing commitment to ethics should influence standards for education, practice, and business applications. Many issues in bioethics and business ethics arise for HIT professionals. Policy issues include efforts to balance the forces that drive a free-market system with the needs of clinicians, patients, researchers, public health workers and officials, and others.

**Ethical and policy issues for electronic health application vendors and users**

The recommendations in this report constitute an effort by AMIA to address and help resolve issues surrounding vendor-user contracts and subsequent interactions. These issues include:

- the identification of vendor and user duties to protect patient safety and improve healthcare quality;
- responsibility for and mechanisms of identifying and correcting errors in product design and manufacture, device installation, and subsequent modification and use;
- marketing practices;
- the extent to which governments should regulate electronic health systems and software.

The Task Force met in person and by teleconference six times between September 2009 and June 2010. Discussions were candid and wide-ranging. While there was unanimity among members about the primacy of patient safety, there was disagreement about the extent to which recommendations might be regarded as onerous by the HIT vendor community, in part by virtue of adding to what some regarded as an excessive, regulatory burden.

The recommendations here, however, were approved by consensus of the Task Force members and, subsequently, by vote of the AMIA Board of Directors. One of the recommendations is that the panel, or some successor, should undertake analyses aimed at identifying best practices, if not ethical standards, for those who manufacture, sell, and use the tools of health information technology.

The recommendations fall under the following headings:

- contract language;
- education and ethics;
- user groups;
- best practices;
- marketing;
- regulation and oversight of the industry.

The report also includes a bibliography and an online appendix itemizing suggestions for further reading (see www.jamia.org). The authors welcome comments and questions about this report; please direct them to the Task Force chair, whose email address appears on the first page of this article.

**RECOMMENDATIONS**

**Contract language**

The standard of transparency, in conjunction with uncontroversial duties to increase the growth of biomedical knowledge, entails that certain provisions should or should not be included in contracts governing the sale, lease, or use of HIT systems. The Task Force finds and/or recommends that:

a. Contracts should not contain language that prevents system users, including clinicians and others, from using their best judgment about what actions are necessary to protect patient safety. This includes freedom to disclose system errors or flaws, whether introduced or caused by the vendor, the client,

b. Electronic Health Applications comprise healthcare information and practice management systems that are intended to affect the care of patients and could have an impact on the quality and safety of care.

---

ii The authors of this report constituted the members of the task force, and their analysis and recommendations have been reviewed by the Board of Directors and approved as an AMIA Position Statement.
or any other third party. Disclosures made in good faith should not constitute violations of HIT contracts. This recommendation neither entails nor requires the disclosure of trade secrets or of intellectual property.

b. Hospitals, physician purchasers, and other users should understand that commercial products’ screen designs and descriptions of software-supported workflows represent corporate assets developed at a cost to software vendors. Unless doing so would prematurely prevent disclosure of flaws, users should consider obligations to protect vendors’ intellectual property and proprietary materials when disclosing (potential) flaws. Users should understand and accept their obligation to notify vendors before disclosing such features, and be aware of the range of remedies available to both the purchaser and the vendor in addressing safety issues. Equally, or more important, users should consider obligations to protect patient safety via such disclosures.

c. Because vendors and their customers share responsibility for patient safety, contract provisions should not attempt to circumvent fault and should recognize that both vendors and purchasers share responsibility for successful implementation. For example, vendors should not be absolved from harm resulting from system defects, poor design or usability, or hard-to-detect errors. Similarly, purchasers should not be absolved from harm resulting from inadequate training and education, inadequate resourcing, customization, or inappropriate use.

d. While vendors have legitimate corporate interests and duties (e.g., to shareholders), contract language should make explicit a commitment by all parties to patient care and safety, and, as applicable, to biomedical research and public health.

e. Vendors should be protected from claims in which a facility (hospital, medical office, practitioner, etc) causes errors that cannot reasonably be attributed to a defect in the design or manufacture of a product, or to vendor-related problems in installation, updating, or configuration processes. Similarly, vendors should not be held responsible for circumstances in which users make foolish or intentional errors.

f. “Hold harmless” clauses in contracts between Electronic Health Application vendors and purchasers or clinical users, if and when they absolve the vendors of responsibility for errors or defects in their software, are unethical. Some of these clauses have stated in the past that HIT vendors are not responsible for errors or defects, even after vendors have been informed of problems.

g. A collaborative system or process of third- or neutral-party dispute resolution should be developed. Contracts should contain language describing a process for timely and, as appropriate, transparent conflict resolution.

h. Contracts should make explicit a mechanism by which users/clients can communicate problems to the company; and vendors should have a mechanism for dealing with such problems (compare in this regard the processes in place for adverse event and device failure tracking by implantable medical device manufacturers).

i. Contracts should require that system defects, software deficiencies, and implementation practices that threaten patient safety should be reported, and information about them be made available to others, as appropriate. Vendors and their customers, including users, should report and make available salient information about threats to patient safety resulting from software deficiencies, implementation errors, and other causes. This should be done in a way easily accessible to customers and to potential customers. This information, when provided to customers, should be coupled with applicable suggested fixes, and should not be used to penalize those making the information available. Disclosure of information should not create legal liability for good-faith reporting. Large HIT systems undergo thousands of revisions when looked at on a feature-by-feature basis. Requirements that the vendor notify every customer of every single feature change on a real-time basis would have the unintended result of obscuring key safety risks, as customers would have to bear the expense of analyzing thousands of notifications about events which are typically rare. Therefore, vendors should notify customers as soon as possible about any product or configuration issues (1) of which they are aware and (2) which pose a risk to patients.

Education and ethics

HIT vendors and their clients should understand that safe and successful uses of HIT systems require significant education about how to install, configure, and use the products. This education should involve a collaboration between vendors and customers, and should be targeted to users, programmers, analysts, and others as needed.

Safe and successful HIT systems further require ethics education, which has become a standard part of professional development in the corporate world. Vendors of Electronic Health Application systems and their clients should adopt enterprise-wide ethics education to parallel that required of healthcare organizations for accreditation. The size of the institution should determine how it implements this requirement. Smaller institutions, including small clinical practices, cannot always mount and sustain an internal ethics and HIT training program, especially if they have no other ethics training available. Contrarily, large and mid-size entities should incorporate ethics education into their ongoing continuing education efforts. Smaller institutions should be encouraged to develop partnerships with other health systems, as well as academic linkages that might assist with providing ethics education. This education should emphasize business ethics and corporate compliance (not just the latter), and should address the following topics, among others:

▸ general business ethics and corporate social responsibility;
▸ ethics of contracts and agreements;
▸ ethical issues arising in the development, manufacture, sale, and maintenance of hospital- and other institution-specific IT systems, made appropriate for relevant employees—isues to be addressed might include patient safety; workflow and workarounds; disclosure of system defects versus intellectual property; privacy, confidentiality and security challenges.

Ethical standards

Vendors and clients should hew to—and publicly announce their commitment to—widely recognized and uncontroversial standards for corporate conduct, and education about these standards. These principles include transparency, veracity, and accountability. It is recognized that these values are sometimes vague and often require elaboration. Thus, for instance, a commitment to transparency does not entail a duty to divulge trade secrets or intellectual property. These standards sometimes

---

*The HIT community should seek support to develop online and other ethics curricula to accommodate the needs of entities of varying sizes. We note in passing the existence of a number of online education programs used by small practices participating in human subjects research (generally in compliance with federal law) and, in some states, ethics requirements for continuing medical and nursing education.*
even conflict with each other, or with other values. It follows that for a public and credible commitment to these standards, vendors and clients (healthcare organizations) should put in place processes (including education, as above) to ensure a consistent stance. In healthcare organizations, these commitments—in conjunction with issues in clinical ethics—confer obligations on ethics committees. It is therefore also recommended that:

- Vendors and clients create internal ethics processes and entities to take responsibility for HIT-related education, consultation, and policy creation and review when they engage in these activities. These functions parallel those identified by the Joint Commission and the American Society for Bioethics and the Humanities for healthcare organizations.
- These ethics processes and entities should be distinct from existing mechanisms for corporate compliance (under Sarbanes-Oxley, for instance).
- If appropriate for their size and mission, vendors and client institutions contribute to the growth of biomedical knowledge by conducting HIT research—analogously to the research missions of pharmaceutical and medical device manufacturers. The results of this research should be published according to standards and conventions for biomedical sciences. Authors of scientific reports should not be prevented from identifying devices, tools, and systems by name in publications.

**User groups**

Purchasers and users of Electronic Health Application systems, as well as other stakeholder groups, comprise user groups to share information, data, and news about successes, challenges, failures, and so on. Many user groups are product- or vendor-specific. In some cases, these groups may be seen as mechanisms for collaborating with stakeholders, which can include members of professional groups and provider organizations, as well as users. The Task Force recommends that:

- The user community should identify, or develop and provide, resources such as pointers to (i) contract tool kits, (ii) organizations or consortia which smaller practices and institutions could join for group negotiation, and (iii) similar organizations or practices with which to share experiences.
- Processes for ensuring the fair, reliable, and, as appropriate, transparent reporting of defects should be established.
- The HIT community should identify a “trusted broker” to recommend or develop such processes.

**Best practices**

Health IT vendors, system purchasers, users, and others comprising the HIT community should encourage and assist institutions and clinical practices in their efforts to adopt systems optimized for high-quality healthcare and patient safety. The means by which this might be accomplished include:

1. Creating vendor-supplied information technology that flexibly integrates with varying workflows in client organizations.
2. Adapting institutional workflow to match the changes associated with information technology adoption, giving adequate resources to staff training, and assuring that software configurations are based on clear requirements and are thoroughly tested.
3. Developing sample contracts, with a “menu” of choices for wording and explanation of the pros and cons of each.
4. Encouraging a consensus development process or standard setting process for vendor contracts, reporting forms, and kindred instruments.
5. Establishing a bulletin board, database, listserv, etc, where questions could be posted, experiences could be shared, etc.
6. Identifying common or widely recognized risks and harms and the means to reduce or prevent them.
7. Publicizing and improving on existing resources that function like a “Consumer Reports” for HIT systems.
8. Collaborating with other organizations on these issues.

**Marketing**

There are situations in which HIT vendors pursue joint marketing agreements with institutions that adopt vendors’ products and by which these institutions become a part of the vendors’ marketing program, often in exchange for discounts, payments, stock options, or favorable treatment by the vendor. In at least some cases, these agreements include provisions whereby healthcare institutions that serve as demonstration sites for particular products receive compensation when other institutions adopt products from the same vendor. The Task Force notes that such agreements might place the “referring” institutions in a conflict of interest, and therefore recommends that:

- Any such conflicts should be eliminated or managed, including disclosure, according to current standards.
- Where such agreements are made, they should include a provision whereby any payment or other compensation contingent on the sale of a system to another party must be disclosed to that other party.
- Payments or gifts to individuals and institutions, including institutional officials, clinicians, etc, should be disclosed. Alternatively, they should be addressed by entities’ internal mechanisms for managing conflicts of interest and commitment, perhaps along the lines of the “rebuttable presumption” standard endorsed by the Association of American Medical Colleges. The goal of the standard is “to ensure that institutions systematically review any financial interest that might give rise to the perception of a conflict of interest, and further, that they limit the conduct of human subjects research by financially interested individuals to those situations in which the circumstances are compelling.”

**Regulation and oversight of the industry**

The idea of additional regulation or oversight of HIT vendors continues to be a source of intense controversy. The US Food and Drug Administration has recently asked a network of hospitals to report data on safety issues raised by adoption of HIT systems; the American Reinvestment and Recovery Act (ARRA) of 2009 calls for increased data evaluation to accompany the allocation of economic stimulus funds directed to HIT adoption; Sen. Charles Grassley of the US Senate Committee on Finance has requested information from hospitals and vendors about HIT adoption and interactions with the HIT vendor community; the Health Information Technology Policy Committee, a federal advisory committee, has made recommendations regarding certification to the Office of the National Coordinator (ONC); and the ONC has proposed the establishment of certification programs “for purposes of testing and certifying health information technology.”

In 1997, AMIA, in conjunction with other leading organizations, called for local oversight of clinical software systems and adoption by HIT system developers of a code of good...
It is clear that health information systems are increasingly large and complex, and that both vendors and users of their systems share responsibility for product safety and effectiveness. The challenge, as ever, involves identifying the appropriate amount of regulation, simultaneously to foster innovation and to protect and improve patient safety.

Since 1997, the extraordinarily rapid growth of HIT system adoption, the direct digital links between patient physical monitoring systems and HIT, and, moreover, the role of government in fostering that adoption, suggests that additional work—if not ongoing assessment—is necessary. Therefore, the Task Force recommends that AMIA join with other stakeholders to revisit the role of governmental and other formal regulation and governance of institutions that manufacture and use health information systems—including, but not limited to, electronic health records, personal health records, computerized provider order entry systems, electronic medication administration record systems, and laboratory systems.

**NEXT STEPS**

The recommendations herein are wide-ranging, and some should be regarded as first approximations of future, more comprehensive and perhaps sustained analyses. Moreover, the Task Force recognizes the increased attention these issues now receive from governmental, professional, and scientific organizations. Therefore, the Task Force recommends that AMIA ensure that either this group or a successor group undertakes the analysis of determining the next steps. Concurrently, AMIA should identify appropriate stakeholders to participate in the process. Given the potential expense involved, AMIA should seek funding to support the undertaking.

The analysis should address the following areas of major concern:

- Identify and define a framework of best practices which should reflect the importance of patient safety and the needs of clinicians, researchers, the public health community, vendors, and start-up initiatives.
- Give special attention to the role of legislation governing, and requiring the regulation of, HIT systems.
- Develop “practice guidelines” to ensure that HIT systems continuously function as claimed, including ongoing in situ testing, evaluation, and other quality control responsibilities with regard to all products and upgrades as they are actually used in each institution over the product life.
- The analysis should also take the following into consideration:
  - Defect and hazard reporting and management.
  - Development of tools to identify loci of defects and hazards (eg, in manufacture, implementation, use, etc).

- Responsibilities of manufacturers and vendors.
- Responsibilities of hospitals and other institutions, clinicians and other users and, regarding personal health records, patients.
- Possible effects on HIT purchasers and HIT vendors of any new FDA regulations when combined with the numerous in-process ARRA “Meaningful Use” and temporary and permanent certification regulations.
- Time commitments for HIT/EHS users to meet “meaningful use” and “Evaluation & Management” documentation requirements and the implications of this time and effort on patient safety and access, as well as fair apportionment of legal and moral responsibility and accountability.
- Indemnification of vendors and users for good faith actions and disclosures.
- Parallels to “disclose-and-apologize” mechanisms for medical error reduction.
- Unintended consequences

Task Force members are mindful of the well-motivated controversy surrounding the idea of (increased) government regulation of medical and health-related software. A majority of members, however, are of the view that (1) given the patient safety concerns on the table, (2) in light of noteworthy cases involving adverse events, and (3) because of the need to foster public trust in a rapidly expanding use of electronic health records with embedded decision-support functionality, personal health records, and research tools, some system of government oversight or regulation of health information technology needs to be given serious consideration. Although the form of any such oversight was left for further discussion, we hope that a useful framework will emerge from the additional studies recommended here.

**Acknowledgments** The AMIA Board of Directors endorsed this position paper as representing the position of the Board on these topics on September 26, 2010.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**REFERENCES**


Education and debate

Should we screen for gestational diabetes?

R J Jarrett wrote to us arguing that gestational diabetes was a muddled concept that provided uncertain benefits for mother and infant. We invited Richard Beard and colleagues to put the other side of the argument and make the case for screening all pregnant women for gestational diabetes.

“The concept of gestational diabetes was popularised before considerations of evidence based medicine came on the scene”

R J Jarrett

Much confusion surrounds the topic of screening for glucose intolerance—hyperglycaemia during pregnancy in terms of who should be screened, how to screen, and the management of those with positive results.\(^1\) Confusion arises from lack of or poor quality evidence, compounded in this instance by a concept (gestational diabetes mellitus) founded on risk of subsequent non-insulin dependent diabetes mellitus rather than outcome of the index pregnancy.\(^5\) In addition the criteria for gestational diabetes prescribe a minimum, but not a maximum, level of glucose intolerance, so that any group of women labelled as having gestational diabetes might contain some with glycaemia in the range that qualifies for a diagnosis of non-insulin dependent diabetes, rendering comparisons of different series impossible. Coustan, whose comment is reproduced in my title, suggested four questions which required answers to achieve resolution:\(^1\):

1. How severe must maternal hyperglycaemia be to measurably worsen pregnancy outcome?
2. Can we intervene to prevent adverse outcomes?
3. Is such intervention cost effective?
4. If so, what is the most appropriate way of screening and detecting the problem?

Severity of maternal hyperglycaemia

Women with pre-existing diabetes, either insulin dependent or non-insulin dependent, undoubtedly have an increased risk of bearing a child with a congenital abnormality and this risk is related to the degree of hyperglycaemia during embryogenesis.\(^6,7\) However, this is not relevant to screening at typical booking times as embryogenesis is complete by week seven of gestation. By contrast, gestational diabetes as defined\(^8\) is not associated with risk of congenital abnormalities,\(^9\) despite the presence of some women with glucose intolerance sufficient to qualify them for a diagnosis of non-insulin dependent diabetes.

The only commonly (though not absolutely consistently) reported “complication” of gestational diabetes is macrosomia, a rather emotive description of a newborn infant with a birth weight in the upper centiles (variously defined) of the distribution. To what extent birth weight is determined by maternal glycaemia is debated, but the relation is confounded by maternal fatness.\(^10\) A very large baby is more likely to give rise to obstetric problems and to acquire a birth injury, but one estimate suggested that about 4% of women with untreated gestational diabetes would deliver infants weighing 4500 g or more compared with about 2% of the general obstetric population.\(^10\) While there is some evidence that treatment can reduce fetal weight,\(^11\) this cannot be automatically assumed to be justified given the data showing an inverse association between birth weight and the incidence of disorders in later life.\(^12\)

Are there any adverse effects of the diagnosis of gestational diabetes? Women with gestational diabetes are more likely to be delivered by caesarean section. This has been attributed to their higher proportion of large babies, but in a recent study the section rate was higher even though the proportion of large babies was not,\(^1\) supporting the view that the diagnostic label sensitises obstetricians. The gestational diabetes label also leads to the necessity of self monitoring of blood glucose and possibly insulin injections. The possible distress due to screening and treatment in someone who previously thought herself to be healthy has not been investigated.

Intervention

There is only one clinical trial of any merit.\(^11\) In this 66 women, mostly Hispanic and including an appreciable (though unstated) number with undiagnosed non-insulin dependent diabetes, were treated at random with either a more or a less intensive insulin regimen. Birth weights were, on average, nearly 400 g less in the intensively treated group, but caesarean section rates were not significantly different. Whether treatment influences any outcome of pregnancy in women discovered to have non-insulin dependent diabetes during pregnancy has not otherwise been subject to clinical trial. Indeed, pregnancy associated with non-insulin dependent diabetes has attracted little research interest.
Screening tests and cost effectiveness

Screening for hyperglycaemia is bedevilled by the lack of a suitable screening test. If sensitivity is important then some kind of glucose tolerance test is essential to identify gestational impaired glucose tolerance. Oral glucose tolerance tests are, however, tedious to perform and poorly reproducible. Single blood tests, such as glycated haemoglobin and fructosamine, cannot even identify the lower range of non-insulin dependent diabetes glycaemia, let alone gestational impaired glucose tolerance, though they could identify more florid hyperglycaemia.

“Screening is bedevilled by the lack of a suitable test.”

The only relevant data available on cost effectiveness concern the yield of screening for non-insulin dependent diabetes using the World Health Organisation’s “epidemiological” criterion—a plasma glucose value over 11.0 mmol/l two hours after a 75 g oral glucose load. Two studies provide minimal estimates of incidence of 4/10 000 for Europid women


Conclusions

The ethics of screening require the screener to show the likelihood of benefit from screening. No clear benefit has been shown from screening for glucose intolerance—hyperglycaemia (at least for the woman being screened) during pregnancy, and there are disadvantages, which include the acquisition of disease status and an increased risk of caesarean section. It is argued that screening to identify someone at risk of subsequent non-insulin dependent diabetes or with undiagnosed non-insulin dependent diabetes is a good thing. If so, it should be available to all adults, not only pregnant women. However, the most recent review of population screening for non-insulin dependent diabetes advocates extensive and varied further research on all aspects of the question. This was in the context of non-pregnant adults, but the same requirements apply to screening in pregnancy before it can be regarded as justified.

Since submitting this article I have noted two American groups which do not recommend screening for gestational diabetes. The American College of Obstetricians and Gynecologists, which in 1986 recommended selective screening, in 1994 noted the absence of data to support screening and did not make a specific recommendation. The US Preventive Services Task Force cites insufficient evidence for or against screening for gestational diabetes. In contrast, an expert committee of the American Diabetes Association continued to recommend screening, though no longer without some degree of selection.

The case for screening for gestational diabetes

Jacqueline de A C Soares, Anne Dornhorstt, Richard W Beard

Screening for gestational diabetes is imperative but we need to refocus away from short term improvements in obstetric outcomes to more important medium and long term health benefits. Today 5% of United Kingdom and 12% of United States total healthcare expenditure is spent on diabetes and its complications. There is a global epidemic of non-insulin dependent diabetes, and radical preventive measures are required if morbidity and mortality from diabetes are to fall. We would ask whether we can afford not to screen for gestational diabetes.

Identifying future diabetics

Detecting gestational diabetes identifies women at risk of future non-insulin dependent diabetes. The success of treating non-insulin dependent diabetes is severely hampered by the high percentage of diabetic complications present at diagnosis, so earlier diagnosis is important in preventing complications. Half of all people with non-insulin dependent diabetes are clinically undiagnosed, and diagnosis takes, on average, seven years from onset of the disorder.


(Accepted 4 March 1997)
The rate of progression to non-insulin dependent diabetes mellitus after a pregnancy where the woman has had gestational diabetes depends predominantly on ethnicity and the degree of glucose intolerance both in pregnancy and immediately afterwards. Other contributing factors are weight during pregnancy and subsequent weight gain, age, parity, and family history. In high risk populations, such as Hispanic American women, about 40% of women with gestational diabetes develop diabetes within six years, which rises to 70% among those with impaired glucose tolerance (by World Health Organisation criteria) after birth. In white Europeans the rate of progression to diabetes is slower—20–40% within 20 years. Identifying women who are at risk while they are still young provides an opportunity to identify the disease in subsequent pregnancies and to modify the natural history of non-insulin diabetes mellitus.

The onset of non-insulin dependent diabetes in women can be delayed by weight control and exercise, the benefits being greatest among obese women with a family history of diabetes. Whether further benefit can be obtained with a change of lifestyle or drugs is currently being evaluated in the United States. If this study does show that intervention can delay diabetes in these women the long term benefits of screening for gestational diabetes will be further established.

Effects on the child
The consequences of gestational diabetes on the health of the child have until recently received little attention. The recent explosion of studies underlining the importance of the intrauterine environment for future adult chronic diseases has challenged the sceptics’ view that in pregnancy lesser degrees of maternal glucose intolerance, not justifying treatment of the non-pregnant woman, are irrelevant. The concept that diabetes begets diabetes through an intrauterine effect on the fetal pancreas, additional to any genetic effect, is strongly supported by animal and human epidemiological studies.

In support of this hypothesis are studies within the Pima population, where children of a diabetic mother are at greater risk of diabetes and childhood obesity than older siblings born before their mother became diabetic. Maternal carbohydrate metabolism may also influence future human fetal insulin secretion and function as suggested by the studies on black and white American adolescents. Those born to diabetic mothers have greater insulin resistance and are more likely to be glucose intolerant during puberty. The literature supports an effect of maternal hyperglycaemia on an infant’s future susceptibility to abnormalities of carbohydrate metabolism. However, the critical threshold of hyperglycaemia is currently not known.

Poor maternal diabetic control is associated with an increased risk of large for gestational age infants. In certain ethnic groups up to half of all pregnancies where gestational diabetes is present have evidence of accelerated fetal growth. The apparent paradox of low birth weight associated with an increased risk of future diabetes and high birth weight with a decreased risk has distracted attention from the knowledge of the potential harm in later life of accelerated intrauterine fetal growth associated with gestational diabetes. In a high risk population with a 5% prevalence of gestational diabetes up to half the infants above the 90th birthweight percentile theoretically could be from a diabetic pregnancy. In such populations both high and low birth weights are associated with an increased risk of diabetes in later life. By contrast, when the prevalence of gestational diabetes is low (0.5–1%) at most only 5–10% of all infants above the 90th birthweight percentile could be the result of maternal diabetes. Thus in low risk populations this argument for universal screening is not so strong.

Demonstrating benefit
The short term benefits of screening and treating gestational diabetes have focused on pregnancy outcome. In high risk populations, with a high background prevalence of diabetes combined with limited access to medical and perinatal care, perinatal mortality can be seen to improve after screening for and treatment of gestational diabetes. Retrospective studies suggest a benefit on stillbirth rates after the introduction of screening and treating gestational diabetes in low risk populations, but demonstrating a benefit on perinatal mortality in prospective trials has proved more difficult. In Western populations, with a low prevalence of diabetes, good access to medical care, and low perinatal mortality and morbidity rates, there are ethical constraints in mounting randomised trials with sufficient power to test whether treating gestational diabetes reduces perinatal morbidity. Prospective studies in these populations have therefore assessed pregnancy outcome using surrogate markers of diabetic control. These include macrosomia, need for caesarean section, and fetal hypoglycaemia. None of these end points are specific for diabetes and many are influenced by the practice of individual obstetricians, maternal obesity, age, and parity.

These difficulties should not, however, detract from the fact that maternal hyperglycaemia is the cause of a diabetic fetopathy syndrome of Pedersen. With the knowledge that a baby showing evidence of this syndrome may well develop diabetes in later life, there

Screening for gestational diabetes: should we do it?
is a good case for early detection and treatment of the mother. Evidence of increased visceral fat and enlargement of the liver, spleen, and heart may be apparent on ultrasound as early as 28 weeks’ gestation. These accelerated growth patterns, associated with gestational diabetes, can be corrected with diet or insulin, which results in fewer large for gestational age infants and fewer operative deliveries. To what extent, if any, abnormal fetal growth patterns due to hyperinsulinemia reflect aberrant fetal cell programming and what influence this may have on the future insulin sensitivity and adult health remains speculative. To ignore such a link in the face of mounting animal evidence would, however, be short sighted.

Who should be screened?

Should all pregnant women be screened or only those at risk? The answer needs to reflect the ethnicity of the population, the availability of health care, and the economic and medicolegal expectations of the country. Once the decision has been made to screen a reproducible screening test needs to be chosen that is sensitive, specific, and easily applied.

The most universally researched screening test is the O’Sullivan test, which involves a one hour timed blood glucose sample after a 50 g oral glucose load, a value >7.8 mmol/l being positive. This test has a ~95% sensitivity and ~85% specificity for detecting pregnancy induced glucose intolerance that occurs at 20-28 weeks’ gestation. A first trimester test is advisable in high risk populations, in which more women will have gestational diabetes before 20 weeks. This can be done with either the 50 g oral glucose load or a timed plasma glucose value, which will identify women with glucose intolerance likely to require insulin treatment—namely, a fasting plasma glucose concentration >6 mmol/l or a 2 hour postprandial value >9 mmol/l Other screening tests—which include random glucose values, glucosuria, fructosamine, diurnal glucose profiles, and glucose responses to mixed random glucose values, glucosuria, fructosamine, diurnal glucose profiles, and glucose responses to mixed meals—have been less extensively evaluated in pregnancy than the O’Sullivan test, which remains the gold standard. The sensitivity of purely clinical risk factors is poor, <70%, especially in multiethnic populations, since they do not include ethnicity.

The confirmatory diagnostic test for gestational diabetes remains controversial. Gestational diabetes is usually diagnosed on the basis of an oral glucose tolerance test. However, the exact load administered (50, 75, or 100 g) varies between centres. The need for one test and one set of diagnostic criteria is recognised. Epidemiologically the 75 g oral glucose tolerance test has the advantage that it is internationally used outside pregnancy. However, the diagnostic limits at which treatment is required still need to be defined.

A dogmatic stand against screening for gestational diabetes not only ignores the proved benefits of treatment on perinatal outcome but also denies affected mothers the possibility to reduce their own and their babies’ risk of later diabetes.

How to read a paper
Papers that go beyond numbers (qualitative research)

Trisha Greenhalgh, Rod Taylor

What is qualitative research?

Epidemiologist Nick Black has argued that a finding or a result is more likely to be accepted as a fact if it is quantified (expressed in numbers) than if it is not.¹ There is little or no scientific evidence, for example, to support the well known “facts” that one couple in 10 is infertile, or that one man in 10 is homosexual. Yet, observes Black, most of us are happy to accept uncritically such simplified, reductionist, and blatantly incorrect statements so long as they contain at least one number.

Researchers who use qualitative methods seek a deeper truth. They aim to “study things in their natural setting, attempting to make sense of, or interpret, phenomena in terms of the meanings people bring to them,”¹¹ and they use “a holistic perspective which preserves the complexities of human behaviour.”¹²

Questions such as “How many parents would consult their general practitioner when their child has a mild temperature?” or “What proportion of smokers have tried to give up?” clearly need answering through quantitative methods. But questions like “Why do parents worry so much about their children’s temperature?” and “What stops people giving up smoking?” cannot and should not be answered by leaping in and measuring the first aspect of the problem that we (the outsiders) think might be important. Rather, we need to listen to what people have to say, and we should explore the ideas and concerns which the subjects themselves come up with. After a while, we may notice a pattern emerging, which may prompt us to make our observations in a different way. We may start with one of the methods shown in box 1, and go on to use a selection of others.

Box 2 summarises (indeed, overstates) the differences between the qualitative and quantitative approaches to research. In reality, there is a great deal of overlap between them, the importance of which is increasingly being recognised.⁴

Quantitative research should begin with an idea (usually articulated as a hypothesis), which then, through measurement, generates data and, by deduction, allows a conclusion to be drawn. Qualitative research, in contrast, begins with an intention to explore a particular area, collects “data” (observations and interviews), and generates ideas and hypotheses from these data largely through what is known as inductive reasoning.⁵ The strength of the quantitative approach lies in its reliability (repeatability)—that is, the same measurements should yield the same results time after time. The strength of qualitative research lies in validity (closeness to the truth)—that is, good qualitative research, using a selection of data collection methods, really should touch the core of what is going on rather than just skimming the surface. The validity of qualitative methods is greatly improved by using a combination of research methods, a process known as triangulation, and by independent analysis of the data by more than one researcher.

The so called iterative approach (altering the research methods and the hypothesis as the study progresses, in the light of information gleaned along the way) used by qualitative researchers shows a commendable sensitivity to the richness and variability of the subject matter. Failure to recognise the legitimacy of this approach has, in the past, led critics to accuse qualitative researchers of continually moving their own goalposts. Though these criticisms are often misguided, there is, as Nicky Britten and colleagues have observed, a real danger “that the flexibility [of the iterative approach] will slide into sloppiness as the researcher ceases to be clear about what it is (s)he is investigating.”¹⁶ These authors warn that qualitative researchers must, therefore, allow periods away from their fieldwork for reflection, planning, and consultation with colleagues.

Summary points

Qualitative methods aim to make sense of, or interpret, phenomena in terms of the meanings people bring to them

Qualitative research may define preliminary questions which can then be addressed in quantitative studies

A good qualitative study will address a clinical problem through a clearly formulated question and using more than one research method (triangulation)

Analysis of qualitative data can and should be done using explicit, systematic, and reproducible methods

Box 1

Examples of qualitative research methods

Documents—Study of documentary accounts of events, such as meetings
Passive observation—Systematic watching of behaviour and talk in natural occurring settings
Participant observation—Observation in which the researcher also occupies a role or part in the setting, in addition to observing
In depth interviews—Face to face conversation with the purpose of exploring issues or topics in detail. Does not use preset questions, but is shaped by a defined set of topics
Focus groups—Method of group interview which explicitly includes and uses the group interaction to generate data
Evaluating papers that describe qualitative research

By its very nature, qualitative research is non-standard, unconfined, and dependent on the subjective experience of both the researcher and the researched. It explores what needs to be explored and cuts its cloth accordingly. It is debatable, therefore, whether an all-encompassing critical appraisal checklist along the lines of the Users’ Guides to the Medical Literature could ever be developed. Our own view, and that of a number of individuals who have attempted, or are currently working on, this very task, is that such a checklist may not be as exhaustive or as universally applicable as the various guides for appraising quantitative research, but that it is certainly possible to set some ground rules. The list which follows has been distilled from the published work cited earlier; and also from our own research and teaching experiences. You should note, however, that there is a great deal of disagreement and debate about the appropriate criteria for critical appraisal of qualitative research, and the ones given here are likely to be modified in the future.

Question 1: Did the paper describe an important clinical problem addressed via a clearly formulated question?
A previous article in this series explained that one of the first things you should look for in any research paper is a statement of why the research was done and what specific question it addressed. Qualitative papers are no exception to this rule; there is absolutely no scientific value in interviewing or observing people just for the sake of it. Papers that cannot define their topic of research more closely than “We decided to interview 20 patients with epilepsy” inspire little confidence that the researchers really knew what they were studying or why.

You might be more inclined to read on if the paper stated in its introduction something like, “Epilepsy is a common and potentially disabling condition, and up to 20% of patients do not remain free of fits while taking medication. Antiepileptic medication is known to have unpleasant side effects, and several studies have shown that a high proportion of patients do not take their tablets regularly. We therefore decided to explore patients’ beliefs about epilepsy and their perceived reasons for not taking their medication.”

Question 2: Was a qualitative approach appropriate?
If the objective of the research was to explore, interpret, or obtain a deeper understanding of a particular clinical issue, qualitative methods were almost certainly the most appropriate ones to use. If, however, the research aimed to achieve some other goal (such as determining the incidence of a disease or the frequency of an adverse drug reaction, testing a cause and effect hypothesis, or showing that one drug has a better risk-benefit ratio than another), a case-control study, cohort study, or randomised trial may have been better suited to the research question.

Question 3: How were the setting and the subjects selected?
The second box contrasts the statistical sampling methods of quantitative research with theoretical methods of qualitative research. In quantitative research, it is vital to ensure that a truly random sample of subjects is recruited so that the results reflect, on average, the condition of the population from which that sample was drawn.

In qualitative research, however, we are not interested in an “on average” view of a patient population. We want to gain an in depth understanding of the experience of particular individuals or groups; we should therefore deliberately seek out individuals or groups who fit the bill. If, for example, we wished to study the experience of non-English speaking British Punjabi women when they gave birth in hospital (with a view to tailoring the interpreting or advocacy service more closely to the needs of this patient group), we would be perfectly justified in going out of our way to find women who had had a range of different birth experiences—an induced delivery, an emergency caesarean section, a delivery by a medical student, a late miscarriage, and so on—rather than a “random” sample of British Punjabi mothers.

Question 4: What was the researcher’s perspective, and has this been taken into account?
It is important to recognise that there is no way of abolishing, or fully controlling for, observer bias in qualitative research. This is most obviously the case when participant observation is used, but it is also true for other forms of data collection and of data analysis. If, for example, the research concerns the experience of asthmatic adults living in damp and overcrowded housing and the perceived effect of these surroundings on their health, the data generated by techniques
such as focus groups or semistructured interviews are likely to be heavily influenced by what the interviewer believes about this subject and by whether he or she is employed by the hospital chest clinic, the social work department of the local authority, or an environmental pressure group. But since it is inconceivable that the interviews could have been conducted by someone with no views at all and no ideological or cultural perspective, the most that can be required of the researchers is that they describe in detail where they are coming from so that the results can be interpreted accordingly.

**Question 5:** What methods did the researcher use for collecting data—and are these described in enough detail?
I once spent two years doing highly quantitative, laboratory-based experimental research in which around 15 hours of every week were spent filling or emptying test tubes. There was a standard way to fill the test tubes, a standard way to spin them in the centrifuge, and even a standard way to wash them up. When I finally published my research, some 900 hours of drudgery was summed up in a single sentence: “Patients’ serum rhubarb levels were measured according to the method described by Bloggs et al [reference to Bloggs et al’s published paper].”

The methods section of a qualitative paper often cannot be written in shorthand or dismissed by reference to someone else’s research techniques. It may have to be lengthy and discursive since it is telling a unique story without which the results cannot be interpreted. As with the sampling strategy, there are no hard and fast rules about exactly what details should be included in this section of the paper. You should simply ask, “have I been given enough information about the methods used?”, and, if you have, use your common sense to assess, “are these methods a sensible and adequate way of addressing the research question?”

**Question 6:** What methods did the researcher use to analyse the data—and what quality control measures were implemented?
The data analysis section of a qualitative research paper is where sense can most readily be distinguished from nonsense. Having amassed a thick pile of completed interview transcripts or field notes, the genuine qualitative researcher has hardly begun. It is simply not good enough to flick through the text looking for “interesting quotes” which support a particular theory. The researcher must find a systematic way of analysing his or her data, and, in particular, must seek examples of cases which appear to contradict or challenge the theories derived from the majority.

One way of doing this is by content analysis: drawing up a list of coded categories and “cutting and pasting” each segment of transcribed data into one of these categories. This can be done either manually or, if large amounts of data are to be analysed, via a tailor-made computer database. The statements made by all the subjects on a particular topic can then be compared with one another, and more sophisticated comparisons can be made such as “did people who made statement A also tend to make statement B?”

In theory, the paper will show evidence of “quality control”—that is, the data (or at least, a sample of them) will have been analysed by more than one researcher to confirm that they are both assigning the same meaning to them, although in practice this is often difficult to achieve. Indeed, when researching this article, we could find no data on the interobserver reliability of any qualitative study to illustrate this point.

**Question 7:** Are the results credible, and if so, are they clinically important?

We obviously cannot assess the credibility of qualitative results through the precision and accuracy of measuring devices, nor their significance via confidence intervals and numbers needed to treat. It usually takes little more than plain common sense to determine whether the results are sensible and believable, and whether they matter in practice.

One important aspect of the results section to check is whether the authors cite actual data. Claims such as “general practitioners did not usually recognise the value of audit” would be infinitely more credible if one or two verbatim quotes from the interviewees were reproduced to illustrate them. The results should be independently and objectively verifiable—after all, a subject either made a particular statement or (s)he did not—and all quotes and examples should be indexed so that they can be traced back to an identifiable subject and setting.

**Question 8:** What conclusions were drawn, and are they justified by the results?
A quantitative research paper should clearly distinguish the study’s results (usually a set of numbers) from the interpretation of those results (the discussion). The reader should have no difficulty separating what the researchers found from what they think it means. In qualitative research, however, such a distinction is rarely possible, since the results are by definition an interpretation of the data.

It is therefore necessary, when assessing the validity of qualitative research, to ask whether the interpretation placed on the data accords with common sense and is relatively untainted with personal or cultural perspective. This can be a difficult exercise, because the language we use to describe things tends to impugn meanings and motives which the subjects themselves may not share. Compare, for example, the two statements, “three women went to the well to get water” and “three women met at the well and well to get water.” It is becoming a cliché that the conclusions of qualitative studies, like those of all research, should be “grounded in evidence”—that is, that they should flow from what the researchers found in the field. Mays and Pope suggest three useful questions for determining whether the conclusions of a qualitative study are valid:

- how well does this analysis explain why people behave in the way they do?
- how comprehensible would this explanation be to a thoughtful participant in the setting?
- how well does the explanation cohere with what we already know?

**Question 9:** Are the findings of the study transferable to other clinical settings?
One of the commonest criticisms of qualitative research is that the findings of any qualitative study pertain only to the limited setting in which they were obtained. In
fact, this is not necessarily any truer of qualitative research than of quantitative research. Look back at the example of British Punjabi women described above. You should be able to see that the use of a true theoretical sampling frame greatly increases the transferability of the results over a “convenience” sample.

Conclusion

Doctors have traditionally placed high value on numerical data, which may in reality be misleading, reductionist, and irrelevant to the real issues. The increasing popularity of qualitative research in the biomedical sciences has arisen largely because quantitative methods provided either no answers or the wrong answers to important questions in both clinical care and service delivery. If you still feel that qualitative research is necessarily second rate by virtue of being a “soft” science, you should be aware that you are out of step with the evidence.

In 1993, Pope and Britten presented a paper to the BSA Medical Sociology Group entitled “Barriers to qualitative methods in the medical mindset,” in which they showed their collection of rejection letters from biomedical journals. The letters revealed a striking ignorance of qualitative methodology on the part of reviewers. In other words, the people who had rejected the papers often seemed to be incapable of distinguishing good qualitative research from bad. Somewhat ironically, qualitative papers of poor quality now appear regularly in some medical journals, whose editors have climbed on the qualitative bandwagon without gaining an ability to appraise such papers. Note, however, that the critical appraisal of qualitative research is a relatively under-developed science, and the questions posed in this chapter are still being refined.

Thanks to Professor Nick Black for advice on this article.

Any questions

Use of a statin for reducing cholesterol levels

If a patient with coronary disease already drinks a glass of wine and eats a piece of fruit daily, eats fish several times a week, dresses salads with olive oil, exercises regularly, and takes a β blocker and vitamin E is it still worth while prescribing a statin to lower “normal” cholesterol concentrations for five or more years? Even if there is a reduction of 30% in relative mortality what is the absolute advantage?

All the habits of this patient may reduce the risk of coronary heart disease, but when serum cholesterol is above 5.0 mmol/l or the total ratio of cholesterol to high density lipoprotein is above 5 more specific treatment should be considered. A statin will reduce the absolute risk in such patients by about 7% over a five year period and more if the cholesterol concentrations are higher but less if they are lower. In other words, for 100 patients treated with a statin for a “normal” cholesterol there will be one coronary event less a year.

Michael Oliver, emeritus professor of cardiology, London

The articles in this series are excerpts from How to read a paper: the basics of evidence based medicine. The book includes chapters on searching the literature and implementing evidence based findings. It can be ordered from the BMJ Publishing Group: tel 0171 383 6185/6245; fax 0171 383 6662. Price £13.95 UK members, £14.55 non-members.


QUOTE-HIV: an instrument for assessing quality of HIV care from the patients’ perspective

C F Hekkink, H J Sixma, L Wigersma, C J Yzermans, J T M van der Meer, P J E Bindels, K Brinkman, S A Danner

Background: An HIV-specific version of the QUOTE questionnaire was developed to measure the quality of care of patients infected with HIV from the patients’ perspective. The consistency and validity of the questionnaire was assessed.

Methods: Focus group discussions were held to select aspects for inclusion in the questionnaire that are important to patients with HIV. Item and inter-item analysis, factor analysis, and reliability analysis were performed to test the internal consistency and validity of the questionnaire.

Results: Twenty seven items (13 generic and 14 HIV specific) were used in the QUOTE-HIV questionnaire. Separate factor analyses of the generic and HIV specific aspects indicated that each loaded onto a single factor. The internal consistency of the total questionnaire was good (Cronbach’s alpha ≥ 0.80). Feasibility of the questionnaire was shown by the diversity of importance and performance scores for general practitioners as well as for HIV specialists and AIDS nursing consultants.

Conclusion: The QUOTE-HIV questionnaire is a useful instrument for measuring the quality of care from the perspective of HIV infected patients.
To make sure that we were dealing with subjects who were representative of those infected with HIV, participants were members of and were recruited through the Dutch HIV Patient Society at the end of 1998. Letters were sent to their members in which patients were asked to participate in the focus group discussions. Two focus groups were used to determine subjects' opinions about what constitutes good quality of care from their point of view and the bottlenecks encountered in the current healthcare system. A flow chart detailing the development of the questionnaire is shown in fig 1.

The first session was an open group discussion of the definition of good quality of care and the problems to be resolved. These sessions were tape recorded and translated into formal statements of care aspects mentioned by either one or both groups. The first focus group session resulted in 56 statements of care aspects. In the second session the statements of care aspects derived from the first session were grouped and labelled by the participants and prioritised into five levels of importance ranging from 1 (relatively unimportant) to 5 (extremely important). This was done according to the concept mapping process to quantify the relative value of quality of care aspects derived from the focus groups. Concept mapping is a type of structured conceptualisation which can be used by groups to develop a conceptual framework to guide evaluation of planning.

The QUOTE questionnaire contains 10 generic aspects. These aspects were included in the questionnaire regardless of their importance rating. Further selection of the aspects by the researchers was based on the sorting and rating of the aspects derived by the participants in the second focus group session. These were classified by the researchers into the following sub-dimensions of quality of care: continuity of care, treatment, accessibility, autonomy, costs, information, competence, efficiency, privacy, and accommodation. The most important aspect within each sub-dimension was selected for the questionnaire. The list was completed using all aspects with a relatively high importance score (>3.5) as derived from the concept mapping procedure and regardless of their sub-dimension.

Empirical testing
Based on the patients' ratings, 17 items were selected and used in the questionnaire together with the 10 permanent generic QUOTE items (see table 2 for complete list). Quality aspects varied from items generally used in patient satisfaction scales to items tailored to the wishes of this specific group of patients. The final questionnaire contained 13 generic items, including the 10 generic items of the already existing QUOTE questionnaires, and 14 HIV specific items.

In order to pilot the questionnaire, two general practitioners who were not eligible for participation in the prospective study sent it to the patients with HIV in their practices. No reminder was sent. To verify whether the focus groups were successful in determining all relevant quality of care items, an open question asking about additional aspects was included in the questionnaire.

Measures
In the final questionnaire quality of care aspects were formulated as importance and performance statements. These statements refer to three healthcare providers: general practitioner (GP), medical specialist (SP), and AIDS nursing consultant (AC). Respondents were asked to rate all 27 aspects with respect to the three categories of healthcare providers. For the AIDS nursing consultant four items were not relevant and were left out (see table 4). Importance and performance were measured using a 4-point Likert scale. For the importance aspects scores were calculated for the following categories (1="not important", 2="fairly important", 3="important", and 4="extremely important") by linear transformation of standardised values (Z-scores) to values between 0 and 10. For the performance aspects the response categories (1="no", 2="not really", 3="on the whole yes", and 4="yes") were dichotomised into percentages "yes" and "no". The performance score (P) represents the proportion of respondents who were not satisfied with the care received.

The concept “quality of health care from the patients’ perspective” was operationalised as the product of importance and (perceived) performance according to the formula: \( Q_{ij} = I_{ij} \times P_{ij} \). The quality improvement score (Q) on a health service (j) by an individual patient (i) is equal to the importance score (1) multiplied by the (perceived) performance score (P). This formula is based on the PES model suggested by Zastowny et al. The conceptual framework of the formula is further explained by Sixma et al. A high quality improvement score means that better care is recommended.

Statistical analyses
Data analyses included item and inter-item analyses (non-response, skewness, correlations) to test internal consistency, explorative factor analysis (principal component analysis with varimax rotation and Kaiser normalisation), and reliability analysis. Factor analysis was performed separately on the 13 generic and 14 HIV specific aspects. Reliability scores refer to the total scale as well as to the different subscales within the generic and HIV specific items. Reliability and validity of the scales were evaluated primarily by looking at the importance scores. These importance ratings are assumed to be more stable and less situation dependent than performance scores. Feasibility of the new QUOTE-HIV patients’ instrument was established by a comparison of the quality improvements scores within and between the different healthcare services evaluated. The analyses were performed using SPSS 10.0.7 for Windows (SPSS, Chicago, IL, USA).
RESULTS

Patients

The two focus panels met at the end of 1998 and comprised three women and 12 men infected with HIV of mean age 49 years (range 30–62) and a mean duration since diagnosis of HIV infection of 9 years (range 1–15). Eleven were using anti-retroviral therapy. Six were diagnosed with AIDS and three others had symptomatic HIV infection.

The questionnaire was sent to 80 people infected with HIV; 44 questionnaires were returned (55%). Responders were predominantly homosexual men educated to higher level with a variable time since HIV diagnosis (table 1), which is a good representation of the HIV infected population in Amsterdam.

Test characteristics of the questionnaire

With respect to the validity of the conceptual framework of the QUOTE instruments, separate analyses of the generic and HIV specific items indicated that each loaded onto a single factor (table 2).

For the total questionnaire the internal consistency as represented by Cronbach’s alpha (≥0.80) was good (table 3). The generic and HIV specific parts were found to have sufficiently reliable scales (≥0.70) except for the HIV specific part of the GP importance scale (0.60). Inter-item correlations indicating the discrimination values of the items are also shown in table 3.

To test the feasibility of the QUOTE-HIV instrument the 27 items were applied to the quality of care given by GPs, HIV specialists, and AIDS nursing consultants. Table 4 shows three scores for each care provider: the importance score (I), the performance score (P), and the quality improvement score (Q).

The mean importance score shows what is important to patients with each care provider together with their expectations. For GPs, patients rated aspect 3 (“to be taken seriously”) of the highest importance (I=8.5), and for specialists and AIDS nursing consultants they rated aspect 16 (“having special knowledge about HIV”) of the highest importance (I-SP=10; I-AC=8.6). Importance scores ranged from 8.5 to 2.7 for GPs, from 10.0 to 3.7 for specialists, and from 8.6 to 4.2 for AIDS nursing consultants on a scale from 0 (“unimportant”) to 10 (“extremely important”).

The performance score shows the percentage of patients who were not satisfied with the delivered care—for example, the performance score of the specialist for aspect 7 (“never keeps me waiting longer than 15 minutes”) is 0.27, which means that 27% of the respondents reported having to wait longer than 15 minutes in the specialist’s waiting room.

The quality improvement score shows the relative score of priorities and performances on quality. The importance of aspect 6 (“can be easily reached by telephone”) is valued at 6.5 with 31% of patients reporting that they could...
not reach their specialist by telephone. The quality improvement score of 2.0 is rather high as a result of the product of a relatively high performance score (P=0.31) and a medium importance score (I=6.5). Thus, according to patients, quality improvement is needed on this aspect.

The quality improvement score can also be used to compare within and between professions and institutions on both individual aspects and combinations of aspects. The quality improvement score of one aspect over several types of healthcare services can be considered. The quality improvement score for aspect 4 (“being aware of my situation at home and at work”) differed for the GP (Q=1.1), specialist (Q=2.0), and AIDS nursing consultant (Q=2.7). This indicates that, from the perspective of patients with HIV, quality improvement

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Reliability of scales for importance and performance components for general practitioners (GP), medical specialists (SP), and AIDS nursing consultants (AC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GP importance</td>
</tr>
<tr>
<td>Mean score</td>
<td>Overall</td>
</tr>
<tr>
<td></td>
<td>General</td>
</tr>
<tr>
<td></td>
<td>HIV specific</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>Overall</td>
</tr>
<tr>
<td></td>
<td>General</td>
</tr>
<tr>
<td></td>
<td>HIV specific</td>
</tr>
<tr>
<td>Cronbach’s alpha</td>
<td>Overall</td>
</tr>
<tr>
<td></td>
<td>General</td>
</tr>
<tr>
<td></td>
<td>HIV specific</td>
</tr>
<tr>
<td>Inter-item correlation</td>
<td>Overall</td>
</tr>
<tr>
<td></td>
<td>General</td>
</tr>
<tr>
<td></td>
<td>HIV specific</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Importance (I), performance (P), and quality improvement (Q) scores of 44 HIV patients for different care providers</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
<td>Aspect description</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1</td>
<td>Work efficiently</td>
</tr>
<tr>
<td>2</td>
<td>Explain my medication clearly</td>
</tr>
<tr>
<td>3</td>
<td>Take me seriously in replacement</td>
</tr>
<tr>
<td>4</td>
<td>Is aware of my situation at home and work/school</td>
</tr>
<tr>
<td>5</td>
<td>Cooperate well with other social workers</td>
</tr>
<tr>
<td>6</td>
<td>Can easily be reached by telephone</td>
</tr>
<tr>
<td>7</td>
<td>Never keep me wait in the waiting room longer than 15 minutes</td>
</tr>
<tr>
<td>8</td>
<td>Prescribe drugs which are free of charge</td>
</tr>
<tr>
<td>9</td>
<td>Keep his appointments</td>
</tr>
<tr>
<td>10</td>
<td>Make sure I have an appointment within 24 hours if necessary</td>
</tr>
<tr>
<td>11</td>
<td>Take my opinion into account</td>
</tr>
<tr>
<td>12</td>
<td>Allow me to ask a second opinion</td>
</tr>
<tr>
<td>13</td>
<td>Allow me to check my personal file</td>
</tr>
<tr>
<td>14</td>
<td>Inform me about the pros and cons of a treatment</td>
</tr>
<tr>
<td>15</td>
<td>Explain laboratory results</td>
</tr>
<tr>
<td>16</td>
<td>Have special knowledge of HIV</td>
</tr>
<tr>
<td>17</td>
<td>Keep me in shape with preventive methods</td>
</tr>
<tr>
<td>18</td>
<td>Have an open ear for a conversation about euthanasia</td>
</tr>
<tr>
<td>19</td>
<td>Give information about possible side effects of drugs</td>
</tr>
<tr>
<td>20</td>
<td>Give information about the use of my HIV medication</td>
</tr>
<tr>
<td>21</td>
<td>Break news gently</td>
</tr>
<tr>
<td>22</td>
<td>Take enough time to talk with me</td>
</tr>
<tr>
<td>23</td>
<td>Treated by the same person</td>
</tr>
<tr>
<td>24</td>
<td>Has organised his replacement well</td>
</tr>
<tr>
<td>25</td>
<td>No interruptions during a consult</td>
</tr>
<tr>
<td>26</td>
<td>Maintain confidentiality about my HIV status</td>
</tr>
<tr>
<td>27</td>
<td>Is organised in such a way I cannot hear conversations at the desk or in the consulting room</td>
</tr>
</tbody>
</table>

GP=general practitioner; SP=specialist; AC=AIDS nursing consultant.
*Generic aspects selected by focus group ratings.
concerning this aspect is needed for AIDS nursing consultants and (to a somewhat lesser extent) medical specialists.

DISCUSSION

Focus group discussions resulted in 27 items (13 generic and 14 HIV specific) being used in the QUOTE-HIV questionnaire. Factor analysis showed that the questionnaire could be divided into two parts: a generic part and an HIV specific part. The internal consistency of the total questionnaire was good (Cronbach’s alpha ≥0.80). Feasibility of the questionnaire was shown by the diversity of importance and performance scores for GPs, HIV specialists, and AIDS nursing consultants. The QUOTE-HIV therefore seems to be a useful instrument for measuring the quality of care from the perspective of patients infected with HIV.

Working with focus groups has its limitations. Firstly, the result of the discussions may be biased by the fact that only a small number of patients were involved, all of whom were members of the Dutch HIV Patient Society. To find out whether the focus groups were successful in determining the relevant quality of care items, an open question was included in the questionnaire. Some patients used this question to comment on the list of items but no new items were mentioned, so we can assume that the questionnaire contains the most relevant items of quality of care for people infected with HIV. Secondly, working with focus group panels includes a subjective element by the moderator and observer(s) who are part of the group dynamics. This influence is minimised by using more than one panel and by recording the meeting on audio or videotape.

We did not subdivide each sub-dimension (continuity of care, autonomy, etc) into different scales because the questionnaire has to be easy to handle and can therefore contain only a limited number of aspects.

With respect to the quality of care aspects reported in table 3, in those such as “take enough time to talk with me”, “treated by the same person”, and “in organisation in such way I cannot hear conversations at the desk or in the consulting room” it is hard to distinguish parts of the items from those used in many other satisfaction with care studies such as those reported by Hall and Dornan and Wensing. However, this is only true for the part of the items which looks at their content and at the way they were formulated. Most items are more informative and more practical than those in existing patient satisfaction scales and better reflect the needs of specific categories of patients. For instance, items concerning medication can be expected to be important to everyone in general, but for HIV infected persons who have to take a lot of drugs in complicated dosing schedules such items are especially important. The patients in our study mentioned three medication aspects (“give information about possible side effects of drugs”, “inform me about the pros and cons of a treatment”, and “give information about the use of my HIV medication”) among the five most important aspects of care for HIV infected patients as well as known patient satisfaction items.

In which an interaction between performance and importance scores gives a high quality improvement score. The QUOTE-HIV questionnaire is therefore exceedingly useful for identifying aspects that really need improvement. Secondly, patients played a crucial role in the development of this instrument, which ensures that the items important to them are included. As a result, an equal number of items encountered specifically by HIV infected patients as well as known patient satisfaction items were generated.

In table 4 a quality improvement score of 3.5 for GPs is shown for the aspect “give information about my HIV medication”. The QUOTE-HIV questionnaire therefore indicates the need for an education course for GPs in which more attention is paid to the use of HIV medication.

To limit the number of items in the questionnaire it could be restricted to a single time rating of the importance of an item in general. We wanted to compare importance scores as well as performance scores for GPs and HIV specialists and therefore asked for an importance rating for each healthcare provider separately.

Further research is needed for specific patient groups. We did not look at different disease categories or medication use: future studies should validate the instrument for specific categories.

We consider that this instrument can be used widely although some adaptation to the local situation in other parts of Europe might be warranted. Healthcare systems in Europe differ between countries and sometimes between regions in the same country, so the instrument has been provided with a method to adapt it to specific characteristics of health care. For such cross-cultural adaptation, additional focus group discussions with patients will play an important role. The scope of the instrument is not limited to providers but can be used for functions of care such as the provision of care for non-institutionalised patients with HIV.

Key messages

- The QUOTE-HIV is a questionnaire developed from the patients’ perspective to judge the quality of care for patients infected with HIV.
- Unlike most satisfaction instruments, QUOTE-HIV has an importance as well as a performance component.
- The QUOTE-HIV instrument allows a quality improvement score to be calculated which indicates where quality improvement might be most useful.

Authors’ affiliations

C F Hekkink, PJ E Bindels, Division of Clinical Methods & Public Health, Department of General Practice, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

J T M van der Meer, Division of Infectious Diseases, Tropical Medicine and AIDS, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands

H J Sixma, CJ J Ysermans, Netherlands Institute for Health Services Research (NIVEL), Utrecht, The Netherlands

L Wigersma, Royal Dutch Medical Association, Utrecht, The Netherlands

K Brinkman, Department of Internal Medicine, Onze Lieve Vrouwe Gasthuis, Amsterdam, The Netherlands

S A Donner, Department of Internal Medicine, VU Medical Centre, Amsterdam, The Netherlands

REFERENCES


14 Kitzinger J. The methodology of focus groups: the importance of interaction between research participants. Soc Health Illness 1994;16:103–21.


QUOTE-HIV: an instrument for assessing quality of HIV care from the patients' perspective


Qual Saf Health Care 2003 12: 188-193
doi: 10.1136/qhc.12.3.188

Updated information and services can be found at:
http://qualitysafety.bmj.com/content/12/3/188.full.html

These include:

References
This article cites 15 articles, 5 of which can be accessed free at:
http://qualitysafety.bmj.com/content/12/3/188.full.html#ref-list-1

Article cited in:
http://qualitysafety.bmj.com/content/12/3/188.full.html#related-urls

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/
Health IT Success and Failure: Recommendations from Literature and an AMIA Workshop

BONNIE KAPLAN, PHD, KIMBERLY D. HARRIS-SALAMONE, PHD

Abstract

With the United States joining other countries in national efforts to reap the many benefits that use of health information technology can bring for health care quality and savings, sobering reports recall the complexity and difficulties of implementing even smaller-scale systems. Despite best practice research that identified success factors for health information technology projects, a majority, in some sense, still fail. Similar problems plague a variety of different kinds of applications, and have done so for many years. Ten AMIA working groups sponsored a workshop at the AMIA Fall 2006 Symposium. It was entitled “Avoiding The F-Word: IT Project Morbidity, Mortality, and Immortality” and focused on this under-addressed problem. Participants discussed communication, workflow, and quality; the complexity of information technology undertakings; the need to integrate all aspects of projects, work environments, and regulatory and policy requirements; and the difficulty of getting all the parts and participants in harmony. While recognizing that there still are technical issues related to functionality and interoperability, discussion affirmed the emerging consensus that problems are due to sociological, cultural, and financial issues, and hence are more managerial than technical. Participants drew on lessons from experience and research in identifying important issues, action items, and recommendations to address the following: what “success” and “failure” mean, what contributes to making successful or unsuccessful systems, how to use failure as an enhanced learning opportunity for continued improvement, how system successes or failures should be studied, and what AMIA should do to enhance opportunities for successes. The workshop laid out a research agenda and recommended action items, reflecting the conviction that AMIA members and AMIA as an organization can take a leadership role to make projects more practical and likely to succeed in health care settings.

Introduction

With the United States Congress appropriating more than U.S. $20 billion for health information technology (IT) as part of the Feb 2009 economic stimulus package, the United States joined other countries in national efforts to reap benefits that such technology can bring to health care quality and savings. Moreover, Medicare and private and commercial health plans are implementing a new paradigm for paying for health care services in the United States, known as Value-Based Purchasing (VBP), or pay for performance initiatives (P4P). Those initiatives rely heavily on the use of electronic health records to document the value of clinical services delivered. Tempering the fervor, though, are sobering reports that raise concerns about how the technology is designed and deployed. In Jan 2009, The United States National Research Council advised that nationwide deployment of health information technology would not achieve its goals unless it provided health care workers and patients with support for decision-making and problem-solving, thereby making health IT adoption all the more complex and daunting.1 In the few weeks before passage of the United States stimulus package, troubles with Britain’s National Health Service’s move towards a nationwide electronic health records system were investigated by Parliamentary inquiries,2 the Dutch minister for health announced that their national electronic health record deployment would be postponed despite announcing the roll out almost three months earlier, and smart card introduction in Germany was seriously delayed.3 In addition, the United States Joint Commission on Accreditation of Healthcare Organizations issued a Sentinel Alert in Dec 2008, which warned of technology-related adverse events.4 These are reminders of the complexity and difficulties of implementing even smaller-scale health IT systems.

Despite an accumulation of best practices research identifying success factors, IT implementation projects are often not successful. Across industry sectors, at least 40% of such generic IT projects either are abandoned or fail to meet business requirements, while fewer than 40% of large systems purchased from vendors meet their goals.5,6 Some sources report 70% failure rates.7 Other studies show that as
few as one in eight information technology projects is consid-
ered truly successful, with more than half overshoot-
ing budgets and timetables and still not delivering what was
promised. According to the 2006 CHAOS Report by The
Standish Group, only 35% of IT projects were completed on
time, on budget, and met user requirements. Although that
is more than double the 16.2% reported in the 1994 CHAOS
Report, it still amounts to about two-thirds of projects with
significant problems, including 19% that “failed outright”
down from 31.1% in 1994.9

The range of systems involved and variations in outcomes
raise questions on how to define project “failure.”10,11 A
common definition in health care is that “[s]ignificant bud-
get and timeline overruns, underdelivery of value, and the
outright termination of a project before completion are all
forms of failure.”12 Regardless of definition and other meth-
ological differences, the studies share a common finding:
over half of IT projects do not deliver as they should, are
over budget, or are late.13 Since the 1990s, organizations
such as The Standish Group International Inc (Boston);
KPMG (Toronto); Gartner, Inc (Stamford, CT); and the
Aberdeen Group (Boston) all repeatedly have pronounced
IT project failure a serious problem.13

Similar failure rates have been reported specifically for
health IT.14,15 Hospitals are among those organizations
where delays and cancellations of software projects are
demic.16 For years, problems have plagued the implemen-
tation of health IT applications, whether for ancillary ser-
dices, for whole institutions, for regional or national sys-
tems, or for consumers. Today’s problems are reminiscent
of those analyzed since at least the 1970s in classic studies
of hospital information and patient record systems.17–19 In
1990, Dowling estimated that staff interfere with or sab-
otage “nearly half” of projects,20 while Heeks noted in 2006
that it is his “best estimate that most HIS [health informa-
tion systems] fail in some way.”15

Recent studies and newspaper accounts cite difficulties in
a variety of health information technology applications. Over
the years, in many countries, patterns of severe problems
repeatedly have beset a variety of efforts: hospital informa-
tion systems and electronic records;21–26 ambulance services;27,28
community, regional, and National Health Information net-
works;29–33 public health systems;34,35 patient education;36 and
physician order entry.18,19,37–41 The situation is even more
disturbing when high-profile failures, partial successes, and
unsustainable IT undertakings are coupled with accumulat-
ing evidence of negative unintended consequences, in-
creased error rates accompanying IT use, and the need for
workarounds.42–49

Much is known about ways to reduce these difficulties, as
evidenced by literature on project and change management,
success factors, and ways to identify and address problem-
atic issues in IT implementation in health care. As in other
application areas in different sectors, problems have been
longstanding, with researchers and practitioners addressing
issues of project success since there were projects.16,50–52 In
health care, lessons learned and prescriptions for success
have been available at least since the 1970s.53 More recent
papers include compilations of evaluation research findings,
implementation and project management advice, and sys-
tem success and failure stories in health care.15,53–61

Management wisdom also has been encapsulated in writ-
ings by well-known health care IT executives and govern-
ment bodies12,62–68 and the advice offered is much like that
in other sectors. A 2007 study of 214 projects in a variety
of sectors that included 18 health care projects identified inade-
quate management practices as accounting for 65% of the
factors associated with project failure. The remaining 35% of
the failed projects were classified by the authors as due to
technical factors, including poor or inappropriate require-
ments, design, development tools, user documentation, test
planning, and technical support,8 all arguably management
issues as well. According to the IT executive managers
surveyed for the 1994 CHAOS Report, the three major
reasons for project success are user involvement, executive
management support, and a clear requirements statement,
while lack of these constituted the main reasons for project
challenges, impairments, and cancelations.69 Their recipe for
project success remained much the same in 2001: executive
support, user involvement, experienced project manager,
clear business objectives, and minimized scope.70

However, despite important similarities, health care differs
in significant ways from other sectors. In healthcare IT imple-
mentation, systems need to have well-defined standards for
 interoperability and terminologies and comply with legal
requirements. Health IT systems must generate quality reports
for a variety of different health plans. In addition, such
systems must be flexible enough to support organizations
ranging from solo practitioner offices to national integrated
delivery networks. Ideally these systems also improve
workflow, reduce cost, and improve quality of care, all the
while maintaining long-standing beneficial patterns of com-
munication, collaboration, and healthcare delivery.71 While
recognizing that there still are technical issues related to
functionality and interoperability, a consensus is emerging
that problems with health care IT projects, as in other
sectors,13,14 are due to sociological, cultural, and financial
issues, and hence, are more managerial in nature than
technical. For some years, it has been recognized that system
success requires a mix of organizational, behavioral, cogni-
tive, and social factors. There must be well-developed meth-
ods for design and dissemination; and early determination
of who defines “success,” and when the determination of
“success” is made.53

There have been some published research reports of health-
care IT failures. There have been a few systematic and
thoughtful publications describing lessons learned from IT
interventions that had null, negative, or disappointing out-
comes.27,53 Despite calls for increased research, there are still
too few published research reports of health care IT failures,
removals, sabotage of systems, or how failures became
successes or were otherwise redefined. As in other sectors,69
IT-related failures in health care often are covered up,
ignored, or rationalized, so mistakes are repeated. The same
barriers and problems to health IT have been identified over
the years.72 They parallel those in other sectors in attributing
problems to actors and circumstances outside of manage-
ment’s or informaticians’ control.21,73 One result is alarming
headlines when high-profile health IT failures that ad-
versely affect patient care or when well known institutions
suspend their systems or halt their development due to physician protest, extreme overspending, errors, and delays.\textsuperscript{26,32,37,41,74,75} Less sensational, but certainly serious, are studies of health care computer applications that cause errors through poor design and management.\textsuperscript{43,66,76,77} Significantly, the US's Joint Commission on Accreditation of Healthcare Organizations recognized this problem and issued a Sentinel Alert recommending good management practices to help prevent patient harm through technology-related adverse events.\textsuperscript{4}

Sensational headlines and studies of systems causing errors have both surprised and dismayed the medical informatics community. The many success stories over the years make sometimes less-than-informed mass media reporting of project failures all the more disappointing and problematic. Such reports produce reactions that are as costly both financially and in terms of the benefits that information technology could bring for improving health care. Health care informatics projects are extremely complex, yet their benefits are manifold if the risks of failure are minimized. Multiple stakeholders share an interest in supporting the implementation of health information technology. The United States Congress has passed incentive packages, the Centers for Medicare and Medicaid Services (CMS) have put considerable effort into Pay for Performance initiatives, and Centers for Medicare and Medicaid Services (CMS) have put considerable effort into Pay for Performance initiatives, and providers all are interested parties. With the Obama administration's emphasis on rapid implementation of health IT, issues of failure are all the more acute.

**Workshop Development**

With years of practical experience and research, and with increasing national and international pressure for health IT, the continued prevalence of project failure leads to questions of how to increase the success rate of IT systems implementations. The topic inspired a lively listserv discussion among members of AMIA working groups. The first author, who at the time chaired the IMIA Working Group (WG) on Organizational and Social Issues, realized that with widespread interest in the topic and considerable experience and wisdom in the AMIA membership, a meeting could continue the discussion and enable participants to learn from each other. As a result, ten working groups cosponsored a workshop at the AMIA Fall 2006 Symposium to examine why health IT implementations and applications do not meet the expectations held for them and what might be done to improve the situation. Entitled “Avoiding The F-Word: IT Project Morbidity, Mortality, and Immortality”, the session was devoted to better defining and characterizing reasons for “success” and “failure.”

Presenters representing the sponsoring WGs are listed in Table 1. In addition, J. Michael Fitzmaurice of the Agency for Healthcare Research and Quality (AHRQ) also spoke at the workshop. Table 2 lists the issues framing their comments. After their remarks, over fifty participants broke into smaller groups to continue the discussion, and to develop sets of important issues, action items, and recommendations.

**Table 1** Workshop Presenters (affiliations at time of workshop)

<table>
<thead>
<tr>
<th>Moderator</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bonnie Kaplan, PhD, Yale Center for Medical Informatics, Yale University, New Haven, CT</td>
</tr>
<tr>
<td>Department of Biomedical and Health Information Sciences, University of Illinois—Chicago, Chicago, IL, and Kaplan Associates, Hamden, CT—Chair, International Medical Informatics Association Working Group on Organizational and Social Issues</td>
</tr>
<tr>
<td>Recorder</td>
</tr>
<tr>
<td>Kimberly D. Harris-Salamone, PhD, Director, Physician Office Quality, Health Services Advisory Group, Phoenix, AZ—Chair, AMIA People and Organizational Issues Working Group</td>
</tr>
<tr>
<td>Clinical Information Systems Working Group</td>
</tr>
<tr>
<td>Scot M. Silverstein, MD, Director, Institute for Healthcare Informatics, College of Information Science and Technology, Drexel University, Philadelphia, PA</td>
</tr>
<tr>
<td>Consumer Health Informatics Working Group</td>
</tr>
<tr>
<td>Rita Kukafka, DrPH, MA, Track Director, Public Health Informatics Specialization, Departments of Biomedical Informatics and Sociomedical Sciences, Mailman School of Public Health, Columbia University, New York, NY</td>
</tr>
<tr>
<td>Ethical, Legal and Social Issues Working Group</td>
</tr>
<tr>
<td>Robert Hsiung, MD, Department of Psychiatry, University of Chicago, Chicago, IL</td>
</tr>
<tr>
<td>Evaluation Working Group</td>
</tr>
<tr>
<td>Nicolette de Keizer, PhD, Department of Medical Informatics, Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands</td>
</tr>
<tr>
<td>Medical Imaging Systems Working Group</td>
</tr>
<tr>
<td>Melvyn Greberman, MD, MS, MPH, FACP, President, Public Health Resources, LLC, helped in organizing the workshop and reviewing the draft paper. Unfortunately, the representative from this working group was unable to attend</td>
</tr>
<tr>
<td>Nursing Informatics Working Group</td>
</tr>
<tr>
<td>Linda Dietrich, MSN, RN, Dearborn Advisors, LLC, Chicago, IL</td>
</tr>
<tr>
<td>Pharmacoinformatics Working Group</td>
</tr>
<tr>
<td>Sandi Mitchell, RPh, MSIS, Program Director, Pharmacy Informatics Residency, Medication Use Team, The Johns Hopkins Hospital, Baltimore, MD</td>
</tr>
<tr>
<td>People and Organizational Issues Working Group</td>
</tr>
<tr>
<td>Jos Aarts, PhD, Institute of Health Policy and Management, Erasmus University Medical Center, Rotterdam, The Netherlands</td>
</tr>
<tr>
<td>Prevention and Public Health Working Group</td>
</tr>
<tr>
<td>Paul Fu, Jr, MD, MPH, Department of Pediatrics and Health Services, David Geffen School of Medicine at UCLA and, UCLA School of Public Health; and ocio/CMIO, Los Angeles County Department of Health Services, Los Angeles, CA</td>
</tr>
<tr>
<td>Agency for Healthcare Research and Quality</td>
</tr>
<tr>
<td>J. Michael Fitzmaurice, PhD, FACMI, Senior Science Advisor for Information Technology, Office of the Director, AHRQ, Rockville, MD</td>
</tr>
</tbody>
</table>
The following workshop report is based on notes the second author, then chair of the AMIA People and Organizational Issues WG, kept during the entire workshop and displayed via a projector for all participants to see in real-time. The first author analyzed the second author’s notes together with additional notes taken by one of the attendees, reviewed the literature, identified themes, and produced a draft of what was to become this paper. This draft was sent to all presenters and their comments were incorporated into this paper. What follows reflects what was said at the workshop.

**Workshop Themes**

Three themes characterized the workshop discussion, as summarized in Table 3:

**Table 3 • Workshop Themes**

<table>
<thead>
<tr>
<th>Theme</th>
<th>What “success” is</th>
<th>What makes it so hard—communication, workflow, and quality</th>
<th>What we know—lessons from experience</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>There are different ideas and definitions of success. We need more understanding of different stakeholder views and more longitudinal and qualitative studies of failure.</td>
<td>Difficulties of communicating across different groups makes it harder to identify requirements and understand workflow.</td>
<td>Provide incentives, remove disincentives; identify and mitigate risks; allow resources and time for training, exposure, and learning to input data; learn from the past and from each other.</td>
</tr>
<tr>
<td></td>
<td>What makes it so hard—communication, workflow, and quality</td>
<td>Difficulties of communicating across different groups makes it harder to identify requirements and understand workflow.</td>
<td>Provide incentives, remove disincentives; identify and mitigate risks; allow resources and time for training, exposure, and learning to input data; learn from the past and from each other.</td>
</tr>
<tr>
<td></td>
<td>What we know—lessons from experience</td>
<td>What we know—lessons from experience</td>
<td>Provide incentives, remove disincentives; identify and mitigate risks; allow resources and time for training, exposure, and learning to input data; learn from the past and from each other.</td>
</tr>
</tbody>
</table>

Common threads cross-cutting these three themes were:

- the complexity of IT undertakings,
- the need to integrate all aspects of projects, work environments, and regulatory and policy requirements; and
- the difficulty of getting all the parts and participants in harmony.

**What “Success” Is**

Many comments concerned the complexity of both large-scale projects and the clinical environment. Participants indicated that this made implementation very difficult because it is not only a technical process, but also a social one replete with interprofessional collaboration, the need for top management understanding, and professional and terminological differences. Success may be defined as simply getting the application or system turned on, getting people to use it, and getting at least grudging acceptance, with the caveat that grudging acceptance can turn to non-acceptance. It might entail only offering even “small successes” to users. Problems are compounded in that what works for one group, such as pharmacists, may not work for another group, such as nurses, and those who gain may not be those who actually do the work. For these reasons, there is little agreement about what “success” or “failure” is. As an audience member put it, “failure is in the eye of the stakeholder.”

Participants spoke about the need for more teaching and research—especially longitudinal and qualitative studies—addressing what failed, why it failed, how an institution
tries to turn itself around after failure, and what to do differently the next time.

**What Makes It So Hard—Communication, Workflow, and Quality**
Participants emphasized that communication and workflow issues add to project complexity. Health care requires collaboration, as does system implementation, yet there is difficulty in translating among specialties, stakeholders, clinicians, and implementers, sometimes to the point of a seeming “culture clash.” Related to these communication challenges is the difficulty of identifying requirements for the various groups involved. Individuals gathering requirements may not include all the necessary people within an organization, or these individuals may not know how to effectively communicate their requirements. Some projects are undertaken for reasons other than need for the project: because requirements come down from the top, or because the project was simple to do, or because developers like the people who want the project. Participants described the difficulty in fully understanding workflow, as evidenced by workflow changes resulting in endless workarounds. Sometimes this was due to the inability of those doing the work to articulate what they do or need; sometimes to senior management or IT not understanding the clinical environment or workflow, or not agreeing on what needs to be done; sometimes to not providing sufficient or meaningful incentives to change. By contrast, participants also described projects that went well because they made the workflow easier. Others emphasized that quality issues also need to be considered, especially in light of the importance of administrative and clinical data reporting for Pay for Performance initiatives. Administrative and quality indicators related to workflow, therefore, also need to be incorporated into policies and procedures, thereby further adding to project complexity.

**What We Know—Lessons from Experience**
Participants drew lessons from their research and experiences on how management might improve project success. These included:

- provide incentives, remove disincentives
- identify and mitigate risks
- allow resources and time for training, exposure, and learning to input data

Participants described systems where clinicians had never used a keyboard or had exposure to computers, yet training was very limited. Sufficient training and time to learn need to be part of the implementation, and need to be on-going afterward.

- learn from the past and from others

Participants spoke of the need for studies of successes, failures, and how failing situations were turned around. They suggested longitudinal studies, qualitative studies, more focus on health care teams as a whole, and incorporating insights from change management, diffusion of innovation and technology, social science and sociotechnical theory, and multilevel frameworks. Although participants suggested drawing on existing theories and knowledge and also incorporating project management and methodology issues, they advised caution when doing so because of differences between health care and the business settings where models were developed. There also were calls for measurable evidence, including evidence of publication bias concerning project failure, and for various databases to be created (see below).

**AMIA Action Recommendations**
The workshop concluded with reports from break-out groups charged with discussing ideas for how AMIA could address health informatics failure. Break-out groups made suggestions concerning:

- research and publication
- best practices
- advocacy
- education
- certification
- databases and knowledge integration

These are summarized in Table 4 and described more fully here.

**Research and Publication**
Participants recognized that the questions framing the workshop, listed in Table 2, constitute a research agenda. As indicated above, they recommended addressing these issues through more qualitative and longitudinal evaluations, including examining teams and also different views of “success.” In addition, groups called for further studying underlying processes throughout the life cycle, interface and workflow.
issues, and how organizations turned around after “failure.” Participants recognized the importance of these issues both for large systems and organizations as well as for office practices. They also called for a JAMIA feature or an AMIA blog, with ideas for how to integrate this with existing databases (see below).

Best Practices
Participants recommended identifying best practices and suggested that AMIA produce a White Paper on best practices for health information technology projects. The scope should cover system design, development (including development models and iterative practices), implementation, change management, intuitive interfaces appropriate for clinical settings, help systems, how to identify all stakeholders and insure a common vision among them, workflow and process redesign, and providing benefits (or, as one breakout group put it, getting “the most bang for the buck” and addressing “the pain points”).

However, it also was noted that while AMIA could make recommendations, much already is known about these areas from health informatics research, as well as from research in other domains. Nevertheless, they noted that it can be “hard to translate general principles into practice in actual organizational settings . . . [because] the context. can be very different across organizations.” Therefore, we need more translational research studies that explicitly explore the effects of context on implementation of IT innovations. Further, studies, databases, and examples are important not only for identifying general principles, but also for how they work in practice in particular settings. Such information would help people gain familiarity with how to pull together regulations, workflow, policies, and IT practices in comprehensive ways that make them easier to apply in particular health care settings.

Advocacy
Participants suggested that AMIA not only advocate for best practices, but also participate in the regulatory process. For example, AMIA could point out difficulties related to privacy issues concerning access to on-line patient information (such as getting lists of patients’ room locations) and ways in which the HIPAA privacy legislation impedes workflow. AMIA also could work with such agencies as Office of the National Coordinator for Health Information Technology (ONCHIT) or the Food and Drug Administration in institutionalizing best practices. Standards for alerting and for interoperability also are possible areas for advocacy.

Education
Participants called for developing informatics curricula for both students and professionals. They pointed to the need for core curriculum in medical informatics that would include project management, implementation, and other topics addressed by the workshop. Another idea was to design curricula around actual projects. Participants suggested more training in executive leadership. They further suggested that AMIA partner with professional organizations, the Centers for Medicare and Medicaid Services (CMS), and other entities pushing for Pay for Performance, to develop and promote curricula on best practices and lessons learned. In addition, AMIA working groups might work with the Education Working Group to help develop curricula especially relevant to implementation issues.

Certification
AMIA could work with various certification agencies, such as the Certification Commission for Health Information Technology (CCHIT), to develop guidelines that promote better development, implementation, and use. This might be aided by a new AMIA working group on software development and certification processes.

Databases and Knowledge Integration
Underlying these ideas was the belief that existing knowledge and experience should be collected, integrated, and available for analysis. Participants called for databases of best practices, vendor implementation services, and a case study repository. This repository would be similar to the Healthcare Information Management Systems Society (HIMSS) database of United States hospitals and the systems implemented, but would also include what workflow adjustments were needed or how problems and potential difficulties were addressed. Participants thought that even less formally structured repositories, such as an AMIA blog, would be useful, for sharing both “success” and “failure” examples.

Conclusions
Much has been learned about success and failure in IT implementation, but we need to understand more. There are legal issues when a system “fails”, including just what constitutes “failure.” There are social issues, ranging from how such failures affect various groups and health informatics as a whole (including possible policy and regulatory reactions), to the social aspects of what makes for a “successful” implementation. Finally there are ethical issues involved in evaluating system “success” or not sufficiently attending to previously identified success factors and best practices.24 Most “failures” are failures to properly apply managerial wisdom that has been substantiated by research and experience. Perhaps the worst aspect of failure is failure to learn from past experiences, so the same issues and problems are perpetuated.

Ten AMIA working groups (Table 1), together with the IMIA WG on Organizational and Social Issues, and over fifty other individuals contributed to the workshop. Participants discussed communication, workflow, and quality; the complexity of IT undertakings; the need to integrate all aspects of projects, work environments, and regulatory and policy requirements; and the difficulty of getting all the parts and participants in harmony (Table 3). They addressed what “success” and “failure” mean, what contributes to making successful or unsuccessful systems, how to use failure as an enhanced learning opportunity for continued improvement, how system successes or failures should be studied, and what AMIA should do to enhance opportunities for successes.

The proposed research agenda (Table 2) and recommended action items (Table 4) reflect the conviction that AMIA members and AMIA as an organization can take a leadership role to make projects more likely to succeed in health care settings. Action items address curriculum development, advocacy in the regulatory process, and documenting and disseminating best practices based on both research and learning from experience.
AMIA has been active in some of these areas, but more could be done. Though the workshop affirmed accumulated wisdom concerning best practices, its call for more research and repositories of lessons learned recognized that tools and prescriptions for success need empiric validation and that failures need to be studied to appropriately change practice.27 Participants joined with others in challenging dominant approaches to project management and evaluation. The alternative approach favors more nuanced and broader views of project leadership that include complex intertwined relationships, multi-faceted analyses, political and stakeholder issues, institutional and cultural realities, sensitivity to who benefits and who does not, and different views of what constitutes “success.”18,10,11,15,27,28 In a time of increasing Pay for Performance, pressure for electronic health records, integration across systems, massive expenditures for national health IT programs, and flux in the health care system, workshop participants urged that AMIA and its members take a proactive and applied perspective that addresses the complexities of health informatics projects to realize the potentials of informatics for improving health.

References

17. Lundsgaarde HP, Fischer PJ, Steele DJ. Human Problems in Computerized Medicine, Lawrence, KS: University of Kansas, 1981.
29. Hagland M. From struggles to success. Part technology, part cooperation and part good old fashioned trial and error are what it take to build—or break—A RHIO. Healthc Inform 2007;34:36–7.
33. Another government IT scheme in trouble. Telegraph December 20, 2008. Available at: http://www.telegraph.co.uk/comment/


37. Williams JS. Microchips versus stethoscopes—Calgary Hospital, MDs face off over controversial computer system. Can Med Assoc J 1992;147(10):1534–47.


47. Silverstein S. Sociotechnologic issues in clinical computing: Common examples of healthcare IT difficulties. Available at: http://www.ischool.drexel.edu/faculty/silverstein/failurecases/3mSlWnrSFdrrsYY2ZnNCrozgTIR1h&hl=en&sa=X&oi=book_result&resnum=1&ct=result&sig=652pE9ONQCG&sig=3mslWnrSFdrrsYY2ZnNCrozgTIR1h&hl=en&sa=X&oi=book_result&resnum=1&ct=result&sig=6PZxEN9QC&sig=3mslWnrSFdrrsYY2ZnNCrozgTIR1h&hl=en&sa=X&oi=book_result&resnum=1&ct=result&sig=652pE9ONQCG&sig=3mslWnrSFdrrsYY2ZnNCrozgTIR1h&hl=en&sa=X&oi=book_result&resnum=1&ct=result&sig=6PZxEN9QC.


The epidemiology of ocular trauma in rural Nepal

S K Khatri, A E Lewis, O D Schein, M D Thapa, E K Pradhan, J Katz

Aims: To estimate the incidence of ocular injury in rural Nepal and identify details about these injuries that predict poor visual outcome.

Methods: Reports of ocular trauma were collected from 1995 through 2000 from patients presenting to the only eye care clinic in Sarlahi district, Nepal. Patients were given a standard free eye examination and interviewed about the context of their injury. Follow up examination was performed 2–4 months after the initial injury.

Results: 525 cases of incident ocular injury were reported, with a mean age of 28 years. Using census data, the incidence was 0.65 per 1000 males per year, and 0.38 per 1000 females per year. The most common types of injury were lacerating and blunt, with the majority occurring at home or in the fields. Upon presentation to the clinic, 26.4% of patients had a best corrected visual acuity worse than 20/60 in the injured eye, while 9.6% had visual acuity worse than 20/400. 82% were examined at follow up: 11.2% of patients had visual acuity worse than 20/60 and 4.6% had vision worse than 20/400. A poor visual outcome was associated with increased age, care sought at a site other than the eye clinic, and severe injury. 3% of patients were referred for further care at an eye hospital at the initial visit; 7% had sought additional care in the interim between visits, with this subset representing a more severe spectrum of injuries.

Conclusions: The detrimental effects of delayed care or care outside of the specialty eye clinic may reflect geographic or economic barriers to care. For optimal visual outcomes, patients who are injured in a rural setting should recognise the injury and seek early care at a specialty eye care facility. Findings from our study suggest that trained non-ophthalmologists may be able to clinically manage many eye injuries encountered in a rural setting in the “developing” world, reducing the demand for acute services of ophthalmologists in remote locations of this highly agricultural country.

Ocular trauma is a major cause of monocular blindness and visual impairment throughout the world, although little is known about its epidemiology or associated visual outcome in developing countries. A review suggested that at least half a million people are monocularly blind from ocular trauma worldwide.

A national population based survey of blindness in Nepal (1981) found a blindness prevalence rate of 0.84%, with trauma responsible for 7.9% of monocular blindness. In Nepal, “corneal trauma and ulceration” is the second most common cause of monocular blindness after cataract. Superficial corneal trauma sustained specifically in agricultural societies such as Nepal often leads to rapidly progressing corneal ulceration and visual loss. An ideal data collection system for ocular injury includes: (1) population based comparisons using a known denominator; (2) a record of demographic data and details of injury at time of clinical presentation; (3) physician’s diagnosis and visual acuity at presentation; and (4) the final outcome for this injury with appropriate follow up care. We were able to apply much of this model to our data.

This study aimed to estimate incidence of ocular trauma in a rural area of Nepal, identify the risk factors for trauma, describe the visual acuity in the injured eye at time of clinical presentation, and assess visual outcomes 2–4 months after the injury occurred.

METHODS
This study collected prospective case series data from all patients presenting with ocular trauma between November 1995 through May 2000 to the primary eye care centre run by Nepal Nutrition Intervention Project–Sarlahi (NNIPS) under the auspices of the Nepal National Society for Comprehensive Eye Care and Sushil Kedia Seva Mandhir. This clinic was originally established in 1991 as part of a large USAID sponsored vitamin A supplementation trial (NNIPS) conducted collaboratively by Johns Hopkins University Bloomberg School of Public Health and the Nepal National Society for Comprehensive Eye Care. Over the past decade, the group has been responsible for a number of population based nutrition, health, and blindness prevention research projects in the east central terai (plains) district of Sarlahi, along the Indian border, revealing public health interventions that reliably lower maternal and child mortality, morbidity, and nutritional blindness. Incidence of trauma was calculated for those individuals who lived within the 30 village development communities (VDCs) of the NNIPS study area, using recent census data from those 30 VDCs. The primary eye care centre provided the only free allopathic eye care offered in the district at the time of the study. Owing to the longstanding relationship that exists between this community and the NNIPS Project, as well as the 4–6 hour distance by bus to other specialty eye care, it is likely that the clinic evaluated almost all severe cases of ocular trauma in the 30 VDCs.

Trauma was defined by patient or parental report of recent eye injury for which care was sought. Patient interviews and eye examinations were conducted by a single senior ophthalmic assistant with 15 years of experience in this district. Information collected included the type, cause, and location of injury, whether alternative forms of treatment had been sought, time between injury and presentation, and demographic data.

The eye examination consisted of visual acuity using a (tumbling E) Early Treatment of Diabetic Retinopathy Study (ETDRS) chart, anterior segment examination by slit lamp...
including fluorescein stain, intraocular pressure measured with a Shiotz tonometer, and, if possible, given the condition of the eye, a posterior segment examination after pupil dilatation.

Clinical features of the injury were recorded and categorised as “severe” or “mild.” Severe injuries were corneal ulcer, cataract, penetrating foreign body, corneal rupture, iridodialysis, iris prolapse, dislocated lens, scleral rupture, corneal blood stain, macular/retinal damage and hyphaema. Mild injuries were lid laceration, superficial foreign body, conjunctival tear, subconjunctival haemorrhage, and traumatic uveitis. If a patient complained of a recent injury but the ophthalmic assistant did not note a particular diagnosis, these “injuries” were categorised as mild. For injuries with multiple diagnoses, those with any “severe” component were categorised as severe, and those without a severe component were considered “mild.” Trauma cases in need of treatment beyond the scope of the clinic were referred to eye hospitals in neighboring districts, or to an eye hospital in Kathmandu.

The patients’ corrected visual acuities were categorised as having corrected vision worse than 20/400 (WHO definition of blindness) or vision worse than 20/60 (WHO definition of visual impairment) in the injured eye. Because information on bilateral eye injuries did not appear to differ by eye, the right eye was selected for analysis.

Two to four months following the injury, the ophthalmic assistant attempted to visit the homes of all enrolled patients to measure visual outcomes in the injured eye. Patients who were at least 16 years of age, or who were younger than 16 but were married and no longer living with parents, provided verbal informed consent to participate. If a patient was younger than 16 years old, consent was obtained from at least one parent. Ethical approval for this study was given by the Committee on Human Research of the Johns Hopkins Bloomberg School of Public Health, and by the Nepal Health Research Council.

Statistical methods
In order to predict visual acuity at 2–4 month follow up, two regression models were fitted, one predicting follow up visual acuity worse than 20/60 as the outcome in the affected eye, and the other with acuity worse than 20/400. An initial model for visual acuity was made based on the bivariate associations significant at the 0.05 level. Predictors included in the regression models were demographics (age, sex, literacy), care seeking behaviour (time and location of first treatment), and “severe” versus “mild” injury. Stepwise logistic regression methods were used to determine a “parsimonious” model.

RESULTS
Of the 751 patients who presented with ocular trauma over the 4 year study, 650 (86.6%) were from Sarlahi district. A total of 525 (69.4%) patients from the NNIPS area, who presented with their first eye injury during the study period, comprised the cases for this analysis.

Population based incidence of ocular trauma
Using the number of new cases presenting at the clinic over 4 years, along with census data collected door to door by NNIPS staff in 2001, annual incidence was calculated by age and sex (table 1). The crude incidence of ocular trauma was 0.51 per 1000 population at risk per year (95% CI 0.47 to 0.56). Incidence peaked at 40–49 years for both sexes; 0.65 per 1000 males per year (95% CI 0.58 to 0.72), and 0.38 per 1000 females per year (95% CI 0.32 to 0.42). The relative risk of injury for males compared to females was 1.74 (95% CI 1.45 to 2.09).

Of the 525 patients, 342 (65.1%) were male and 183 (34.9%) were female. Age ranged from newborn to 87 years, with a mean of 28.1 years, although the distribution of age differed between males and females (see table 1).

The most common occupations were farming, domestic work, and student. Literacy was reported among 51% of men, and 31% of women (table 2). Of the 525 incident cases, 239 (45.5%) were right eyes, 280 (53.3%) were left eyes, and six (1.1%) were bilateral injuries. The most common types of injuries were lacerations (73.3%), and the places where injuries most commonly occurred were at home (32%), or in the field (27%). The first place that care was sought was more likely to be at the primary eye care clinic (48.5%) than at the local pharmacy, “doctor,” or health post (42%), or at home with a shaman or local medicine person (10%). For individuals whose injury was categorised as “severe,” the distributions of demographic and injury characteristics were similar to that of the larger population (table 2).

If treatment was first sought outside of this clinic, at locations such as a local pharmacy, health post, or medicine shop (as it was for 51.5% of patients), the mean time to seeking treatment was 2.5 days. The mean time between occurrence of injury and seeking eye care at this eye clinic was 7.8 days. The mean time between the injury and any form of eye care, regardless of location, was 7 days.

Diagnosis at presentation
In all, 355 of the 525 cases (68%) were categorised as “severe” injuries; over 75% of these severe cases had lacerating injuries, with 6% blind and 13% visually impaired at the 2–4 month follow up (table 2).

Topical treatment was offered to 99.4%, and systemic treatment to 12.8% of patients. Surgery was performed on about 19% of patients at initial presentation. Only 15 (2.9%) of the 525 were referred by the ophthalmic assistant for more advanced treatment, and more than 50% of those were referred to Birganj Eye Hospital, the shortest distance for most individuals to travel. Of the injuries referred, more than half were blunt injuries that occurred at home.

Visual acuity could not be measured in 56 patients at clinical presentation. Of these, 73% were age 5 or younger, and visual acuity measurement was not attempted. Of the 469 (89.3%) patients with measurable visual acuity, 45 (9.6%) had corrected visual acuity<20/400 in the injured eye at initial presentation for care, while 124 (26.4%) had corrected visual acuity<20/60 (tables 2 and 3).

Visual acuity at 2–4 month examination
Of the 525 cases of ocular trauma, 431 (82.1%) were seen at follow up; 82 were known to be alive but had permanently moved, and two had died. Of the 94 individuals who were not seen again, 73% had initially been categorised as 20/20 or better. There was no statistically significant difference in visual acuity between those lost and those seen at follow up (p = 0.97 for VA<20/60, p = 0.48 for VA<20/400). Of the 431 who were followed, 392 (91.0%) had a measurable visual acuity; 4.6% of individuals had a corrected visual acuity<20/400 in the injured eye, and 11.2% had visual acuity<20/60 in that eye. The majority of patients demonstrated an improved visual acuity, and this was seen across all age groups. At the time of follow up, 29 people (6.7%) had sought additional incitement for their injury. These injuries were most commonly of the lacerating type, occurred in the fields, and were caused by a plant or stick. Sixty four per cent of those 29 had gone to the Nepal Eye Hospital in Kathmandu, a bus ride of 8–12 hours. Of these 29, 22 were self referred and seven had been referred by the ophthalmic assistant. Characteristics for these subgroups are compared in table 4. The percentage of injuries categorised as “severe” did not differ between males and females (see table 1).
not differ significantly (p = 0.50) between groups, or when compared with the 68% of larger group (p = 0.13). The visual acuities for these subgroups (both initially and at the 2–4 month examination) appear to be worse than those who did not receive additional clinical care for their injury. Severity of injury and clinical diagnoses differed by subgroup.

Statistical modelling of visual acuity

For follow up acuity <20/60 or <20/400 in the injured eye, risk factors included in the parsimonious regression model were age, sex, literacy, paramedical treatment versus eye clinic treatment, and injury categorised as severe versus mild (table 5). While the odds of blindness in the injured eye increased with age (OR 1.03, 95% CI 1.00 to 1.06), severe injury appeared to contribute the greatest risk, with an odds ratio of 8.04 (95% CI 1.02 to 63.42), followed by treatment at a paramedical facility (OR 4.15, 95% CI 1.42 to 12.10). Literacy offered a significant protective effect in all models. Similar risk factors predicted visual acuity worse than 20/60 (table 5).

Likelihood ratio tests indicated that the parsimonious models both for blindness and for visual impairment were similar at predicting visual acuity when compared to the more comprehensive models (p = 0.92 for blindness, p = 0.60 for visually impaired).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Annual ocular trauma incidence by sex and age (n = 525), between November 1995 and May 2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>Age group (years)</td>
<td>Population* (No)</td>
</tr>
<tr>
<td>&lt;15</td>
<td>51 483</td>
</tr>
<tr>
<td>15–19</td>
<td>14 919</td>
</tr>
<tr>
<td>20–29</td>
<td>24 705</td>
</tr>
<tr>
<td>30–39</td>
<td>18 176</td>
</tr>
<tr>
<td>40–49</td>
<td>10 871</td>
</tr>
<tr>
<td>50+</td>
<td>11 232</td>
</tr>
<tr>
<td>Total</td>
<td>131 386</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Distribution of characteristics of cases (n = 525)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No (%)</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
</tr>
<tr>
<td>Farmer</td>
<td>136 (25.9)</td>
</tr>
<tr>
<td>Domestic work</td>
<td>113 (21.5)</td>
</tr>
<tr>
<td>Student</td>
<td>85 (16.2)</td>
</tr>
<tr>
<td>Day/unskilled labour</td>
<td>55 (10.5)</td>
</tr>
<tr>
<td>Private service (skilled, salaried, factory)</td>
<td>31 (5.9)</td>
</tr>
<tr>
<td>Business (shopkeeper, vendor)</td>
<td>30 (5.7)</td>
</tr>
<tr>
<td>Literate (yes)</td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td>174 (50.9)</td>
</tr>
<tr>
<td>Females</td>
<td>57 (31.2)</td>
</tr>
<tr>
<td>Type of injury</td>
<td></td>
</tr>
<tr>
<td>Lacerating</td>
<td>385 (73.3)</td>
</tr>
<tr>
<td>Blunt</td>
<td>98 (18.7)</td>
</tr>
<tr>
<td>Small penetrating (&lt;1/3 corneal diameter)</td>
<td>18 (3.4)</td>
</tr>
<tr>
<td>Large penetrating (&gt;1/3 corneal diameter)</td>
<td>8 (1.5)</td>
</tr>
<tr>
<td>Chemical burn</td>
<td>7 (1.3)</td>
</tr>
<tr>
<td>Location of injury occurrence</td>
<td></td>
</tr>
<tr>
<td>Home</td>
<td>168 (32.0)</td>
</tr>
<tr>
<td>Fields/agriculture</td>
<td>144 (27.4)</td>
</tr>
<tr>
<td>Road</td>
<td>72 (13.7)</td>
</tr>
<tr>
<td>Factory</td>
<td>44 (8.4)</td>
</tr>
<tr>
<td>Construction site</td>
<td>14 (2.7)</td>
</tr>
<tr>
<td>Place of treatment first sought</td>
<td></td>
</tr>
<tr>
<td>Self/family/neighbour/shaman</td>
<td>52 (9.9)</td>
</tr>
<tr>
<td>Paramedical</td>
<td>218 (41.6)</td>
</tr>
<tr>
<td>NNIPS primary eye care centre</td>
<td>254 (48.5)</td>
</tr>
<tr>
<td>VA distribution, initial (n = 469)</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>300 (64.0)</td>
</tr>
<tr>
<td>VA&lt;20/60 (n = 328)</td>
<td>124 (24.6)</td>
</tr>
<tr>
<td>VA&lt;20/400</td>
<td>45 (9.6)</td>
</tr>
<tr>
<td>VA distribution, follow up (n = 392)</td>
<td></td>
</tr>
<tr>
<td>Normal</td>
<td>330 (64.2)</td>
</tr>
<tr>
<td>VA&lt;20/60</td>
<td>44 (11.2)</td>
</tr>
<tr>
<td>VA&lt;20/400</td>
<td>18 (4.6)</td>
</tr>
</tbody>
</table>

*Severe injury (n = 355) was defined as deep (>1/3 corneal thickness) foreign body, corneal ulcer, corneal rupture, corneal blood stain, cataract, dislocated lens, iridodialysis, iridal prolapse, scleral rupture, hyphaema, macula/retina damage.
†Paramedical includes care sought at local pharmacy, local medicine shops or clinics with medical or paramedical workers, health posts, and non-eye care hospital.

DISCUSSION

The primary eye care centre in Sarlahi serves the entire district of approximately 607 823 people (Khatry S, personal communication, 25 April 2003), although cases came from outside the district, and even from India. The population of the 30 NNIPS VDCs was 253 667 in 2001, or a third of the district. The 525 trauma cases that lived within the 30 differed in literacy when compared to that reported in past studies of the same population. A previous NNIPS study reported that 13–16% of the female population was literate,7 while we found much higher rates (31%). This suggests that the population seeking care for ocular trauma was more literate (suggestive of a higher socioeconomic status) than those in the general population of the NNIPS study area. The differences may be that cases of ocular trauma differ from the general population, or because groups differ in care seeking behaviour (Katz J, personal communication, 28 April 2003).7

The apparent association between poor visual acuity and increasing age may be due to several phenomena. Age alone may be a risk factor for a new eye injury, but it may also be responsible for worsening vision unrelated to ocular injury (that is, cataract). Socioeconomic factors may also have a role, as access to care may change with age. A visit to a clinic or hospital also requires taking time from work, and a more socioeconomically stable individual is more capable of this. The eye care provided at this clinic was free for all individuals. We cannot say, based on these data, that the differences in incidence by varied risk factors is related to the injuries themselves, or to a difference in access to care.

The lag between injury occurrence and care seeking was associated with a worse visual outcome in our models, and may reflect not only the type of injury, but also economic constraints beyond a lack of awareness about treatment options. Those individuals with more severe injuries (that is, lacerating or penetrating) may seek care earlier, although there may be little that can be done to save sight. Those with less severe injuries who do not seek care may suffer...
secondary infections and lose vision through progressive corneal scarring. This lack of care seeking is not restricted to rural environments, however. In a study of lifetime prevalence of ocular trauma in Baltimore, Katz and Tielsch found that care was not sought for 18% of ocular injuries (22% for black and 14% for white people), even though 15% of white and 8% of black people monocularly blind as a result of trauma said they had not sought treatment for the injury. Educating patients about the proper place and timing of treatment is fundamental to an improved visual outcome.

The most common locations for injury were similar to findings from the Nepal Blindness Survey where 70% of trauma cases involved either agriculture or domestic work. Agriculture accounts for about one half of the country’s gross national product, and is thought to employ more than nine tenths of the workforce. One study estimated the incidence of eye injuries in agriculture to be 3.5 per 10 000, compared with construction (5.3 per 10 000), and industry (1.9 per 10 000). This suggests the need for better monitoring of rural occupational injuries.

The low rate of referral of patients for further treatment indicated to us that this experienced ophthalmic assistant was comfortable diagnosing and treating most of the eye injuries that presented in the clinic. Those referred and treated had more severe injuries and poorer visual outcomes, suggesting the more difficult cases were those referred. Injury types and locations differed among those who were referred (blunt, in the home) and those who sought additional care without a referral (lacerating, in the field). We can speculate that more people sought additional treatment than expected because those particular individuals felt that the ophthalmic assistant at the Hariau clinic could not properly treat their injuries. These individuals also appeared to have more severe injuries than the case series, but less severe injuries than those referred by the ophthalmic assistant. They may have

### Table 3 Clinical presentation and associated visual acuity

<table>
<thead>
<tr>
<th>Clinical presentation</th>
<th>Distribution: Number (%)</th>
<th>VA at 20/60 at presentation (%)</th>
<th>VA=20/400 at presentation (%)</th>
<th>VA=20/60 at follow up (%)</th>
<th>VA=20/400 at follow up (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Corneal ulcer†</td>
<td>264 (50.3)</td>
<td>32.1***</td>
<td>13.3**</td>
<td>12.5</td>
<td>5.8</td>
</tr>
<tr>
<td>Foreign body†</td>
<td>177 (33.7)</td>
<td>5.6***</td>
<td>1.9**</td>
<td>1.9**</td>
<td>0.8**</td>
</tr>
<tr>
<td>Subconjunctival</td>
<td>70 (13.3)</td>
<td>10.2**</td>
<td>3.4</td>
<td>3.9</td>
<td>0</td>
</tr>
<tr>
<td>Haemorrhage</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Traumatic uveitis</td>
<td>68 (13.0)</td>
<td>75.8***</td>
<td>41.9***</td>
<td>45.3***</td>
<td>26.4***</td>
</tr>
<tr>
<td>“Other” irid damage†</td>
<td>47 (9.0)</td>
<td>78.8***</td>
<td>38.1***</td>
<td>27.8***</td>
<td>22.2***</td>
</tr>
<tr>
<td>Cataract</td>
<td>28 (5.3)</td>
<td>78.6***</td>
<td>35.7***</td>
<td>63.2***</td>
<td>26.3***</td>
</tr>
<tr>
<td>Conjunctival tear</td>
<td>12 (2.3)</td>
<td>10</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Corneal rupture†</td>
<td>7 (1.3)</td>
<td>80.0**</td>
<td>40.0***</td>
<td>40.0**</td>
<td>20.0**</td>
</tr>
<tr>
<td>Iridodialysis†</td>
<td>7 (1.3)</td>
<td>100***</td>
<td>66.7***</td>
<td>66.7**</td>
<td>50.0**</td>
</tr>
<tr>
<td>Dislocated lens†</td>
<td>7 (1.3)</td>
<td>85.7***</td>
<td>57.1***</td>
<td>57.1**</td>
<td>57.1**</td>
</tr>
<tr>
<td>Scleral rupture†</td>
<td>6 (1.1)</td>
<td>100***</td>
<td>25</td>
<td>33.3</td>
<td>33.3</td>
</tr>
<tr>
<td>Iris prolapse†</td>
<td>6 (1.1)</td>
<td>100***</td>
<td>66.7***</td>
<td>40.0</td>
<td>40.0*</td>
</tr>
<tr>
<td>Hyphaema‡</td>
<td>4 (0.8)</td>
<td>100**</td>
<td>50</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Corneal blood staining</td>
<td>2 (0.4)</td>
<td>100</td>
<td>50</td>
<td>100*</td>
<td>0</td>
</tr>
<tr>
<td>Macula/retinal damage</td>
<td>2 (0.4)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

†Variable categorised as “severe.”  
* p<0.05; ** p<0.01; *** p<0.001.

### Table 4 Characteristics of individuals receiving additional eye care after initial assessment by ophthalmic assistant (OA)

<table>
<thead>
<tr>
<th>Age (years) [SD]*</th>
<th>15 initially referred for more advanced care</th>
<th>22 not initially referred by OA, but self referred in interim</th>
<th>Comparison to larger group of 525 ocular trauma patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>% “severe” injuries†</td>
<td>14 (93.3%)</td>
<td>16 (76.2%)</td>
<td>35 (68%)</td>
</tr>
<tr>
<td>Median number of days to treatment at Hariau clinic (mean)</td>
<td>10 (35.7)</td>
<td>4 (6.3)</td>
<td>4 (7.8)</td>
</tr>
<tr>
<td>Initial VA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VA&lt;20/60</td>
<td>25.7 (24.0)</td>
<td>41.2 (13.1)</td>
<td>28.1 (16.4)</td>
</tr>
<tr>
<td>VA&lt;20/600</td>
<td>19 (91.7%)</td>
<td>13 (61.9%)</td>
<td>124 (26.4%)</td>
</tr>
<tr>
<td>VA at 2–4 month follow up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VA&lt;20/60</td>
<td>9 (75%)</td>
<td>7 (33.3%)</td>
<td>45 (9.6%)</td>
</tr>
<tr>
<td>VA&lt;20/600</td>
<td>10 (76.9%)</td>
<td>10 (47.6%)</td>
<td>44 (11.2%)</td>
</tr>
<tr>
<td>Most common injuries</td>
<td>Any iris damage{11}</td>
<td>Corneal ulcer (13)</td>
<td>Corneal ulcer (264)</td>
</tr>
<tr>
<td>Traumatic uveitis {10}</td>
<td></td>
<td>Corneal ulcer (9)</td>
<td>Subconjunctival</td>
</tr>
<tr>
<td>Conical ulcer (5)</td>
<td></td>
<td>Foreign body (177)</td>
<td>Hemothorax (70)</td>
</tr>
<tr>
<td>Dislocated lens (3)</td>
<td></td>
<td>Any iris damage (8)</td>
<td>Traumatic uveitis (68)</td>
</tr>
<tr>
<td>Corneal rupture (2)</td>
<td></td>
<td>Subconjunctival</td>
<td>Cataract (28)</td>
</tr>
<tr>
<td>Retinal damage (2)</td>
<td></td>
<td>Hemorrhage (70)</td>
<td>Cataract (4)</td>
</tr>
</tbody>
</table>
| *Group of 15 and group of 22 were found to differ significantly at 0.05 level.  
†Group of 15 differed significantly from larger group of 525 at 0.05 level.  
‡Both the group of 15 and the group of 22 had a significantly worse visual acuity the group of 525 at the 0.001 level, but did not significantly differ from each other.  
⊥Injuries included iris prolapse, iridodialysis, or other iris damage.
been better off economically and therefore more able to travel far distances and leave work to seek care. This study has a role in estimating a rural population’s needs for acute services of ophthalmologists, as it appears to support the notion that a trained non-ophthalmologist can provide most of what is needed. It also demonstrates the potential impact of education about injury prevention and treatment.

One strength of this study was the use of a locally undertaken door to door census to estimate incidence of ocular trauma in a defined area, which has not been done in many studies developing countries. The incidence in men was greater than in women, similar to what has been reported previously. It is possible that the lower was greater than in women, similar to what has been reported previously. It is possible that the lower incidence among women may be due to their poorer access to care. However, this difference between men and women is seen in so many populations including the United States, so it is likely that at least some part of this finding is due to the difference in exposure risk between men and women.

Another strength was that 82% of cases were seen 2–4 months after initial presentation to the clinic, allowing a comparison of initial with final visual acuity. Since most patients will never have access to corneal transplants or corneal repair, the vision at this time is likely to be their ultimate visual acuity. In addition, a single clinician examined patients using a slit lamp, eliminating intraobserver error, and ETDRS charts were used to determine acuity.

The census data used to calculate incidence were from 2001, a year after the study ended. Any increase in population size during that time would likely lead to an underestimate of the incidence. The incidence may also be underestimated if patients had gone elsewhere for treatment, if they did not seek treatment for minor injuries, or if access to or utilisation of care was problematic.

Ocular trauma in developing countries has not been studied extensively. This study indicates that it is a significant cause of monocular vision loss in all age groups, and an experienced non-ophthalmologist can treat a large proportion of these successfully. Many injuries and their resulting vision loss may be prevented through education about prompt and appropriate care seeking.

### Table 5 Logistic regression models for prediction of visual acuity (VA) at follow up

<table>
<thead>
<tr>
<th>Variable</th>
<th>Comprehensive model OR (95% CI) for VA &lt; 20/60</th>
<th>Parsimonious model OR (95% CI) for VA &lt; 20/60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Self/neighbour/shaman</td>
<td>0.90 (0.25 to 3.26)</td>
<td>0.57 (0.06 to 6.60)</td>
</tr>
<tr>
<td>Sex</td>
<td>1.05 (0.54 to 2.22)</td>
<td>0.59 (0.19 to 1.65)</td>
</tr>
<tr>
<td>Age (years)</td>
<td>1.00 (1.00 to 1.06)</td>
<td>1.03 (1.00 to 1.06)</td>
</tr>
<tr>
<td>Literacy</td>
<td>0.02 (0.01 to 0.41)</td>
<td>0.05 (0.01 to 0.42)</td>
</tr>
<tr>
<td>Place of treatment first sought</td>
<td>2.47 (1.17 to 5.20)</td>
<td>3.75 (1.19 to 11.8)</td>
</tr>
<tr>
<td>Self/neighbour/shaman</td>
<td>0.90 (0.25 to 3.26)</td>
<td>0.57 (0.06 to 6.60)</td>
</tr>
<tr>
<td>Time to treatment first sought</td>
<td>1.11 (0.47 to 2.61)</td>
<td>1.28 (0.36 to 4.53)</td>
</tr>
<tr>
<td>Place of treatment first sought</td>
<td>1.34 (0.36 to 4.67)</td>
<td>1.14 (0.17 to 7.72)</td>
</tr>
<tr>
<td>Worse vision</td>
<td>2.39 (1.21 to 4.73)</td>
<td>4.15 (1.42 to 12.10)</td>
</tr>
<tr>
<td>Clinical presentation</td>
<td>1.89 (0.81 to 4.43)</td>
<td>8.26 (1.04 to 65.43)</td>
</tr>
<tr>
<td>Severe injury</td>
<td>0.05 (0.01 to 0.42)</td>
<td>8.04 (1.02 to 63.42)</td>
</tr>
</tbody>
</table>

* Determined using forward, backward, and stepwise regression methods with inclusion criteria of p < 0.05, fixing age and sex into the model.
† Severe injury (n = 355) was defined as deep (>1/3 corneal thickness) foreign body, corneal ulcer, corneal rupture, corneal blood stain, cataract, dislocated lens, iridodialysis, iris prolapse, “other” injury damage, scleral rupture, hyphaema, macula/retinal damage.

Values in bold if confidence intervals do not contain 1.0. "Paramedical" includes care sought at local pharmacy, local medicine shops or clinics with medical or paramedical workers, health posts, and non-eye care hospitals.

### ACKNOWLEDGEMENTS

This study was carried out under Cooperative Agreement HRN-A-00-97-00015-00 between Office of Nutrition, US Agency for International Development (USAID), Washington, DC, and the Center for Human Nutrition (CHN), Department of International Health, and the Sight and Life Research Institute, Johns Hopkins University, Bloomberg School of Public Health, Baltimore, MD, USA. It was a joint undertaking between the CHN and the Nepal National Society for Comprehensive Eye Health Care, Kathmandu, Nepal.

### Authors’ affiliations

6. A E Lewis, E K Pradhan, J Katz, Department of International Health, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Baltimore, MD 21205, USA
7. O D Schein, J Katz, Dana Center for Preventive Ophthalmology, Wilmer Eye Institute, Johns Hopkins School of Medicine Health, 600 North Wolfe Street, Baltimore, MD 21205, USA

### REFERENCES

The epidemiology of ocular trauma in rural Nepal


doi: 10.1136/bjo.2003.030700

Updated information and services can be found at:
http://bjo.bmj.com/content/88/4/456.full.html

These include:

References
This article cites 10 articles, 2 of which can be accessed free at:
http://bjo.bmj.com/content/88/4/456.full.html#ref-list-1

Article cited in:
http://bjo.bmj.com/content/88/4/456.full.html#related-urls

Email alerting service
Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes

To request permissions go to:
http://group.bmj.com/group/rights-licensing/permissions

To order reprints go to:
http://journals.bmj.com/cgi/reprintform

To subscribe to BMJ go to:
http://group.bmj.com/subscribe/
Published criteria for evaluating health related web sites: review

Paul Kim, Thomas R Eng, Mary Jo Deering, Andrew Maxfield

Abstract

Objective To review published criteria for specifically evaluating health related information on the world wide web, and to identify areas of consensus.

Design Search of world wide web sites and peer reviewed medical journals for explicit criteria for evaluating health related information on the web, using Medline and Lexis-Nexis databases, and the following internet search engines: Yahoo!, Excite, Altavista, Webcrawler, HotBot, Infoseek, Magellan Internet Guide, and Lycos. Criteria were extracted and grouped into categories.

Results 29 published rating tools and journal articles were identified that had explicit criteria for assessing health related web sites. Of the 165 criteria extracted from these tools and articles, 132 (80%) were grouped under one of 12 specific categories and 33 (20%) were grouped as miscellaneous because they lacked specificity or were unique. The most frequently cited criteria were those dealing with content, design and aesthetics of site, disclosure of authors, sponsors, or developers, currency of information (includes frequency of update, freshness, maintenance of site), authority of source, ease of use, and accessibility and availability.

Conclusions Results suggest that many authors agree on key criteria for evaluating health related web sites, and that efforts to develop consensus criteria may be helpful. The next step is to identify and assess a clear, simple set of consensus criteria that the general public can understand and use.

Introduction

The large volume of health information resources available on the internet has great potential to improve health, but it is increasingly difficult to discern which resources are accurate or appropriate for users. Because of the potential for harm from misleading and inaccurate health information, many organisations and individuals have published or implemented criteria for evaluating the appropriate-ness or quality of these resources. Two published reviews of evaluation criteria for health related web sites did not present information on the range of criteria proposed by various authors, and included rating tools that were not developed exclusively for health related sites. Our study reviews criteria currently proposed or employed specifically to evaluate health related web sites.

Methods

Databases and search engines


Criteria

We included criteria when they were explicit, specifically used for evaluating health related web sites, and published in a peer reviewed journal or publicly accessible web site. We also considered peer reviewed journals not indexed by Medline. We included resources framed as “guidelines” because there was little difference between them and other criteria, and the intent of the authors was similar. When subcriteria provided details about main criteria, we included only the main criteria to prevent overrepresenting that author’s perspective. Criteria were extracted and sorted into similar groups according to their wording and description. When a criterion seemed to combine several concepts and could fit in multiple groups, we considered the first mentioned concept. To examine the reproducibility of the criteria groupings, four independent, naive coders assigned 40 randomly selected criteria to the 13 criteria groups. Overall, the coders’ assignment of criteria agreed with us 76% of the time. The agreement coefficient, indicating “per cent agreement above chance” was 0.74 or 74%.

www.bmj.com
Information in practice

Table 1 Rating tools and journal articles with explicit criteria for evaluation of health related web sites

<table>
<thead>
<tr>
<th>Source of rating tool or article</th>
<th>URL address*</th>
</tr>
</thead>
<tbody>
<tr>
<td>AllMed/PharmaLINKS</td>
<td><a href="http://www.allmed.com/links/ratings.html">www.allmed.com/links/ratings.html</a></td>
</tr>
<tr>
<td>American Medical Association</td>
<td><a href="http://www.ama-assn.org/med_link/med_link.htm">www.ama-assn.org/med_link/med_link.htm</a></td>
</tr>
<tr>
<td>Biosites, Pacific Southwest Regional Medical Library</td>
<td><a href="http://www.library.ucsf.edu/biosites/help/guidelines.html">www.library.ucsf.edu/biosites/help/guidelines.html</a></td>
</tr>
<tr>
<td>Growth House</td>
<td><a href="http://www.growthhouse.org/award.html">www.growthhouse.org/award.html</a></td>
</tr>
<tr>
<td>Health A to Z</td>
<td><a href="http://www.healthatoz.com/aboutus.htm">www.healthatoz.com/aboutus.htm</a></td>
</tr>
<tr>
<td>Health Information Institute’s Aesculapius Awards</td>
<td><a href="http://www.hi.org/judging.htm">www.hi.org/judging.htm</a></td>
</tr>
<tr>
<td>Health On the Net Foundation</td>
<td><a href="http://www.hon.ch/HONcode/Conduct.html">www.hon.ch/HONcode/Conduct.html</a></td>
</tr>
<tr>
<td>Health Summit I Mtg/Mitretek Systems</td>
<td><a href="http://www.mitretek.org/hi/showcase/documents/criteria.html">www.mitretek.org/hi/showcase/documents/criteria.html</a></td>
</tr>
<tr>
<td>Health Web</td>
<td>healthweb.org/wp/content/papers/guidelines.html</td>
</tr>
<tr>
<td>HealthFinder</td>
<td><a href="http://www.healthfinder.gov/aboutus/selectionpolicy.htm">www.healthfinder.gov/aboutus/selectionpolicy.htm</a></td>
</tr>
<tr>
<td>McGill University Health Sciences and Other Libraries, Canada</td>
<td><a href="http://www.health.library.mcgill.ca/resource/criteria.htm">www.health.library.mcgill.ca/resource/criteria.htm</a></td>
</tr>
<tr>
<td>Medical Matrix</td>
<td><a href="http://www.medmatrix.org/info/sitesurvey.html">www.medmatrix.org/info/sitesurvey.html</a></td>
</tr>
<tr>
<td>Medsite Navigator, Guide to Digital Science and Medicine</td>
<td><a href="http://www.medsitenavigator.com/mail/submit.html">www.medsitenavigator.com/mail/submit.html</a></td>
</tr>
<tr>
<td>Mental Health Net</td>
<td><a href="http://www.cmhc.com/help/ratings.htm">www.cmhc.com/help/ratings.htm</a></td>
</tr>
<tr>
<td>Mountain and Plains Partnership</td>
<td><a href="http://www.uchsc.edu/csa/areas/health/mapp/5aWWW.htm#public">www.uchsc.edu/csa/areas/health/mapp/5aWWW.htm#public</a></td>
</tr>
<tr>
<td>Nutrition Navigator</td>
<td>navigator.tufts.edu/ratings.html</td>
</tr>
<tr>
<td>Organising Medical Networked Information</td>
<td>omni.ac.uk/agec/invalguid.html</td>
</tr>
<tr>
<td>Reference 16</td>
<td>Not available online</td>
</tr>
<tr>
<td>Physician’s Choice</td>
<td><a href="http://www.midechoice.instruct.htm">www.midechoice.instruct.htm</a></td>
</tr>
<tr>
<td>Psych Central: Best of the Web in Mental Health</td>
<td><a href="http://www.grohot.com/ragistug.htm">www.grohot.com/ragistug.htm</a></td>
</tr>
<tr>
<td>Reference 8</td>
<td><a href="http://www.ama-assn.org/sci-pubs/journals/archive/jama/vol_277/no_15/ed7016x.htm">www.ama-assn.org/sci-pubs/journals/archive/jama/vol_277/no_15/ed7016x.htm</a></td>
</tr>
<tr>
<td>Sympatico’s HealthyWay Health Links</td>
<td>www1.sympatico.ca/healthylife/GENERAL/info_2.html</td>
</tr>
<tr>
<td>The Six Senses Review</td>
<td><a href="http://www.sixsenses.com/FAQ.htm#rating">www.sixsenses.com/FAQ.htm#rating</a></td>
</tr>
<tr>
<td>The Virtual Hospital</td>
<td>indy.radiology.uiowa.edu/Beyond/PeerReviews/01Introduction.html</td>
</tr>
<tr>
<td>The Wilton Library</td>
<td>w3.mai.ni/-/wta/erval.htm</td>
</tr>
<tr>
<td>Reference 8</td>
<td><a href="http://www.bmj.com/archive/708Bp2.htm">www.bmj.com/archive/708Bp2.htm</a></td>
</tr>
</tbody>
</table>

*Because of dynamic nature of web, some URLs may have changed. URLs prefixed with http://

Results

Twenty nine rating tools and articles—24 web sites and five journal articles—had explicit criteria for assessing health related web sites (table 1). Of the 165 criteria identified, 132 (80%) were grouped under 12 specific categories (table 2). Thirty three (20%) criteria that lacked specificity or were unique were categorised as “miscellaneous.” Frequently cited criteria included those dealing with content, design and aesthetics of site, and disclosure of authors, sponsors, or developers.

Discussion

Not surprisingly, “content” of the site, which includes concepts of information quality and accuracy, was the most commonly cited criterion group. Design and aesthetics of the site and ease of use were the second and sixth most frequently cited groups respectively, indicating that authors highly value good quality application design and user interfaces. Disclosure of authors, sponsors, or developers had the third highest frequency, highlighting the need for users to be able to consider a site’s content in the context of who created or financed the site. It was somewhat surprising that disclosure was not more commonly cited given recent reports about misleading health information and fraud on the internet. Most rating tools discriminated between content and the fourth most common criterion group, currency of information (includes frequency of update, freshness, maintenance of site), suggesting that

Table 2 Frequency of explicit criteria for evaluation of health related web sites by criteria groups*

<table>
<thead>
<tr>
<th>Criteria groups</th>
<th>Frequency (%) (n=165)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content of site (includes quality, reliability, accuracy, scope, depth)</td>
<td>30 (18)</td>
</tr>
<tr>
<td>Design and aesthetics (includes layout, interactivity, presentation, appeal, graphics, use of media)</td>
<td>22 (13)</td>
</tr>
<tr>
<td>Disclosure of authors, sponsors, developers (includes identification of purpose, nature of organisation, sources of support, authorship, origin)</td>
<td>20 (12)</td>
</tr>
<tr>
<td>Currency of information (includes frequency of update, freshness, maintenance of site)</td>
<td>14 (8)</td>
</tr>
<tr>
<td>Authority of source (includes reputation of source, credibility, trustworthiness)</td>
<td>11 (7)</td>
</tr>
<tr>
<td>Ease of use (includes usability, navigability, functionality)</td>
<td>9 (5)</td>
</tr>
<tr>
<td>Accessibility and availability (includes ease of access, fee for access, stability)</td>
<td>9 (5)</td>
</tr>
<tr>
<td>Links (includes quality of links, links to other sources)</td>
<td>5 (3)</td>
</tr>
<tr>
<td>Attribution and documentation (includes presentation of clear references, balanced evidence)</td>
<td>5 (3)</td>
</tr>
<tr>
<td>Intended audience (includes nature of intended users, appropriateness for intended users)</td>
<td>3 (2)</td>
</tr>
<tr>
<td>Contact addresses or feedback mechanism (includes availability of contact information, contact address)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>User support (includes availability of support, documentation for users)</td>
<td>2 (1)</td>
</tr>
<tr>
<td>Miscellaneous (includes criterion that lacked specificity or were unique)</td>
<td>33 (20)</td>
</tr>
</tbody>
</table>

*Off five authors who assigned weights or priorities to their proposed criteria, four cited content and one cited peer review (categorised as miscellaneous) as most important criterion. Percentage total does not equal 100 owing to rounding-off.
currency of information is nearly as important as the information itself.

Criteria related to confidentiality and privacy of information were only cited by one author despite widespread interest in this issue.6 Some health related web sites are already collecting personal health information to "tailor" content, and as sites begin to integrate healthcare services and information, confidentiality and privacy safeguards will become increasingly important.20-21

Study limitations
Study limitations include the subjective variables around the scope of the criteria categories used. Testing of the category groupings, however, showed that they were reproducible by others. It is also possible that some authors used the same criteria terms to describe different concepts. Because subcriteria were not included, some concepts may not have been represented. Inherent limitations of web search engines and the dynamic nature of the web also prevented us from locating all existing published criteria.22 Nevertheless, our review located more sources of criteria specifically for health related sites than did previous reviews.15 17

Conclusion
Given the evolving state of the internet, it may be difficult or even inappropriate to develop a static tool or system for assessing health related web sites. Our results suggest that many authors agree on key criteria, and that efforts to develop consensus criteria may be helpful.10 20-25 The next step is to identify and assess a clear, simple set of consensus criteria that the general public can understand and use. Tools that integrate them need to be developed and validated, and their ultimate impact and effectiveness in assisting the public with health related decisions should be monitored to ensure that they remain useful.

We thank Farrokh Alemi and Anne Restino for their assistance and advice on this study. The views expressed in this paper are solely those of the authors and do not necessarily reflect those of the US Department of Health and Human Services.

Contributors: PK participated in data collection, analysis, and interpretation, and writing the paper. TRE formulated the study design, developed the core ideas, and participated in data analysis and interpretation, and writing the paper. MJF participated in the study design and interpretation, and edited the paper. AM participated in data analysis and interpretation, and edited the paper. PK and TRE will act as guarantors for the paper.

Funding: Internal funds of the US Department of Health and Human Services. Competing interests: None declared.

Key messages
- Many organisations and individuals have published criteria to evaluate health related information on the world wide web
- A literature and world wide web search found that the most frequently cited criteria were those dealing with content, design and aesthetics of site, disclosure of authors, sponsors, or developers, currency of information, authority of source, and ease of use
- Criteria related to confidentiality and privacy were only cited by one author
- Consensus regarding critical criteria for evaluation of web based health information seems to be emerging
- Our results indicate that many authors agree on key criteria for evaluating health related web sites, and that efforts to develop a set of key criteria may be helpful

References

Endpiece
How to start
The last thing we find out when writing a book is what we must put first.
Blaise Pascal, Pensées

Pardeep Kumar ¹, Sang-Gon Lee ² and Hoon-Jae Lee ²,*

¹ Department of Ubiquitous-IT, Graduate School of Design & IT, Dongseo University, Sasang-Gu, Busan 617-716, Korea; E-Mail: pradeepkhl@gmail.com
² Division of Computer & Information Engineering, Dongseo University, San 69-1, Jurye-2-Dong, Sasang-Gu, Busan 617-716, Korea; E-Mail: nok60@dongseo.ac.kr (S.-G.L.)

* Author to whom correspondence should be addressed; E-Mail: hjlee@dongseo.ac.kr; Tel.: +82-51-320-1730; Fax: +82-51-327-8955.

Received: 29 November 2011; in revised form: 13 January 2012 / Accepted: 2 February 2012 / Published: 7 February 2012

Abstract: A wireless medical sensor network (WMSN) can sense humans’ physiological signs without sacrificing patient comfort and transmit patient vital signs to health professionals’ hand-held devices. The patient physiological data are highly sensitive and WMSNs are extremely vulnerable to many attacks. Therefore, it must be ensured that patients’ medical signs are not exposed to unauthorized users. Consequently, strong user authentication is the main concern for the success and large scale deployment of WMSNs. In this regard, this paper presents an efficient, strong authentication protocol, named E-SAP, for healthcare application using WMSNs. The proposed E-SAP includes: (1) a two-factor (i.e., password and smartcard) professional authentication; (2) mutual authentication between the professional and the medical sensor; (3) symmetric encryption/decryption for providing message confidentiality; (4) establishment of a secure session key at the end of authentication; and (5) professionals can change their password. Further, the proposed protocol requires three message exchanges between the professional, medical sensor node and gateway node, and achieves efficiency (i.e., low computation and communication cost). Through the formal analysis, security analysis and performance analysis, we demonstrate that E-SAP is more secure against many practical attacks, and allows a tradeoff between the security and the performance cost for healthcare application using WMSNs.
Keywords: medical sensor network; secure healthcare; user authentication; mutual authentication; session key establishment; smart card

1. Introduction

During the last few years, we have seen the great emergence of wireless medical sensor networks (WMSNs) in the healthcare industry. Wireless medical sensors are the cutting edge components for healthcare application and provide drastically improved quality-of-care without sacrificing patient comfort.

A wireless medical sensor network is a network that consists of lightweight devices with limited memory, low computation processing, low-battery power and low bandwidth [1]. These medical sensors (e.g., ECG electrodes, pulse oximeter, blood pressure, and temperature sensors) are deployed on patient’s body and collect the individual’s physiological data and sends the collected data via a wireless channel to health professionals’ hand-held devices (i.e., PDA, iPhone, laptop, etc.). A physician can use these medical sensor readings to gain a broader assessment of patient’s health status. The patient’s physiological data may include heartbeat rates, temperature, blood pressure, blood oxygen level, etc. A typical patient monitoring in hospital environment is shown in Figure 1.

Figure 1. Patient monitoring using a wireless medical sensor network in a hospital environment.

Several research groups and projects are working in health monitoring using wireless sensor networks, for example, CodeBlue [2], LiveNet [3], MobiHealth [4], UbiMon [5], Alarm-Net [6], ReMoteCare [7], SPINE [8], etc. Thus, healthcare systems are the applications that most benefit from using wireless medical sensor technology that can perform patient care within hospitals, clinics and homecare.
Wireless medical sensor technology has offered tremendous advantages to healthcare applications, such as continuous patient monitoring, mass-causality disaster monitoring, large-scale in-field medical monitoring, emergency response, etc. Further, these WMSNs provide many new ways for acute disease analysis (e.g., motion analysis for Parkinson’s disease) [9,10].

However, wireless healthcare development has many challenges, such as reliable data transmission, fast event detection, timely delivery of data, power management, node computation and middleware [8,11–17]. Further, patients’ security and privacy is one of the big concerns for healthcare applications, especially when it comes to adopting a wireless healthcare system (i.e., wireless medical sensors, wireless gateways, mobile devices, etc. [18]). Although wireless healthcare offers many advantages to patient monitoring, the physiological data of an individual are highly vulnerable. Further, due to the wireless nature of devices (i.e., medical sensors, iPhone, PDA, etc.), the patients’ vital signs are much easier to query and monitor (i.e., in an ad hoc manner) within the hospital ward rooms using smart phones, iPhones, PDAs, and laptops, so any adversary can be eavesdropping on patients locally in the ward room using their hand-devices that could cause of patient privacy breaches. More importantly, the patient vitals are very sensitive; so they (i.e., the patient’s vitals) must be kept secure from unauthorized users and security threats [19–28]. Moreover, government laws (e.g., the Health Insurance Portability and Accountability Act of 1996 (HIPAA)) also regulated stringent rules for healthcare providers, such as; individuals’ vital signs are only revealed to authorized professionals (i.e., doctors, caregivers and nurses) and family members [29,30]. A healthcare provider is subject to strict civil and criminal penalties (i.e., either fine or imprisonment) if HIPAA rules are not followed properly [29,30]. Furthermore, as wireless medical sensor nodes themselves provide services to users (doctors, nurses, and technicians, are a few examples) it is necessary to control who is accessing their (the medical sensors’) information and whether they are authenticated to do so. Therefore, strong user authentication is a core requirement to protect from illegal access to patients’ vital signs, and can attain the highest levels of patients’ privacy.

So far many significant researches have been proposed for healthcare using sensor networks and provide sufficient security, such as data confidentiality, authentication, integrity and preserving patient privacy [31–39]. These schemes do not considere strong user authentication, and hence, lack a security mechanism, according to the HIPAA laws [29,30]. Further, in [40–46] the authors proposed a few user authentication protocol for wireless sensor networks, which are either broken or provide less security at very high computation and communication costs. Consequently, to the best of our knowledge, a strong user authentication (i.e., professional authentication) protocol for wireless healthcare applications has not yet been addressed effectively in order to prevent illegal access to wireless medical sensor data.

In this paper, we discuss: (1) the healthcare architecture and major security requirements for healthcare application using wireless medical sensor networks; and (2) propose an efficient-strong authentication protocol, named E-SAP, for healthcare applications using WMSNs. The proposed scheme uses two-factor (i.e., password and smartcard) user authentication, where each user must prove their authenticity first and then access the patient vital signs. (Note: We used user and professional, interchangeably and user or professional may be a doctor, a nurse, a surgeon or a technician. Furthermore, it is now widely believed that two-factor authentication provides strong and high level of security (i.e., secure access of individual physiological data from wireless sensors) [29,30,47]).
In addition, E-SAP provides secure session key establishment between the users and the medical sensor nodes, and allow users to change their password. Furthermore, we demonstrate the formal verification of the proposed protocol by the Burrows, Abadi and Needham (BAN) logic model [48], where two main security properties are verified: authenticity and secure session key establishment. Moreover, the proposed scheme resists many practical attacks (e.g., replay, user and gateway masquerade, smartcard stolen-verifier, gateway secret key guessing, password guessing, and information-leakage). To attain the low computational overheads, our scheme uses one-way hash functions along with XOR operations and symmetric cryptosystem.

The rest of paper is organized as follows: Section 2 discusses the healthcare architecture using wireless medical sensors, adversary attack model, and wireless healthcare security requirements. Section 3 briefly reviews the related literature for secure healthcare monitoring using medical sensor networks. Section 4 introduces and describes a novel E-SAP: efficient-strong authentication protocol for healthcare application using WMSNs. Section 5 describes the brief introduction of BAN logic and provides formal verification of E-SAP using the BAN logic model. Section 6 discusses the security analysis and efficiency evaluation in contrast to exiting schemes and finally, in Section 7 conclusions and future directions are presented.


This section presents healthcare monitoring architecture for hospital environments, adversary attack models and security requirements for healthcare application using WMSNs.

2.1. Healthcare Architecture

A patient healthcare monitoring architecture is depicted in Figure 2, where usual patient monitoring is needed after patient hospitalization (e.g., after cardiac infarction). When a patient is hospitalized, he/she can get some suitable medical sensor devices, deployed strategically on the patient’s body. These sensors sense the health parameters, (e.g., blood pressure, movement, breathing, ECG etc.) and send physiological parameters to the professionals’ mobile devices (such as PDA, smart phone and laptop).

Later, a professional may store patient data on the backend server for further processing, which is currently outside the scope of this paper. It is obvious that a professional can access the patient’s health parameters directly from the medical sensor, in an ad-hoc manner.

As shown in the Figure 2, the healthcare architecture has three active entities, namely, user, medical sensors and base-station/gateway. We assume a real-time scenario, and suppose a professional wants to query the patient’s medical sensors for physiological information, as follows: (a) the user (\(U\)) sends a query to the gateway node (\(GW\)); (b) upon receiving the professional’s request, the gateway node forwards the user’s query to the medical sensor; and (c) thereafter, the medical sensor responds to the user. Here, the gateway node plays an important role between the professional and the medical sensor. Based on the above scenario, the next sub-section describes an adversary attack model for healthcare application using WMSNs.
2.2. Adversary Attack Model

The patient’s physiological information is very sensitive and may attract many attackers, such as insurance companies, corrupt media persons, individual enemies, etc. Furthermore, the patient’s medical sensors and the professionals’ hand-held devices are wireless in nature. So, these wireless devices may attract unauthorized users or thieves, more especially. For example, they (unauthorized users or thieves) can roam to the hospital ward and easily eavesdrop on the patients locally, so we have categorized the attack models as follows:

2.2.1. Eavesdropping on Wireless Medical Data

As the medical sensors sense the patient’s body data, they transmit it over the radio communication channel. The wireless transmission ranges are not confined to hospital wards and these wireless channels are highly susceptible. As a result, an attacker may eavesdrop air messages (i.e., a patient’s physiological information), and can disclose the patient’s physiological information. Hence, the patient privacy is breached.

2.2.2. Active Attack

In an active attack scenario, the capability of an attacker depends on his/her skill (i.e., ability to monitor all the communication). An attacker may inject bogus messages into the wireless channel and may alter the wireless medical sensor data during the communication. Any spurious messages injection into the healthcare network could cause mistreatment. Furthermore, an attacker may replay the old messages again and again, which could cause overtreatment (i.e., medicine overdose). Thus, active attacks endanger and may pose a life-threatening risk to the patients.
2.3. Security Requirements for Healthcare Application Using Wireless Medical Sensor Networks

Based on the above attack model and literature survey [19–28] and [31–39], this sub-section sketches out the paramount security requirements for healthcare application in WMSNs, as follows:

2.3.1. Strong User Authentication

The major problem in wireless healthcare environments is the vulnerability of wireless messages to access by unauthorized users, so it is desirable that strong user authentication be considered, where each user must prove their authenticity before accessing the patient’s physiological information. Furthermore, strong user authentication, also known as two-factor authentication, provides greater security for healthcare application using wireless medical sensor networks [47].

2.3.2. Mutual Authentication

In real-time healthcare applications, the user and the medical sensor must authenticate each other; hence, they can ensure the communication is established between the authenticated user and the medical sensors.

2.3.3. Confidentiality

The patient health data are highly sensitive and medical sensors are wireless in nature, therefore patient physiological data should remain confidential from passive attacks such as eavesdropping or traffic analysis. Thus, patient’s health data is only accessed or used by authorized professionals.

2.3.4. Session Key Establishment

A session key should be established between a user/professional and a medical sensor node, so that subsequent communication could take place securely.

2.3.5. Low Communication and Computational Cost

Since wireless medical sensors are resource constrained devices, and the healthcare application’s functions also need room for executing their tasks, the protocol must be efficient in terms of communication and computational cost.

2.3.6. Data Freshness

Generally, professionals need patient physiological data at regular intervals, so there must be guarantee that patient health data is recent or fresh. Furthermore, it (data freshness) also ensures that an adversary cannot replay the old messages.

2.3.7. Secure Against Popular Attacks

In real-time healthcare environments the protocol should be defensive against different popular attacks, such as replay attack, impersonation attack, stolen-verifier attack, password guessing attack,
and information-leakage attack. As a result, the protocol can be easily applicable to the real-time wireless healthcare applications.

2.3.8. User-Friendliness

The healthcare architecture should be easy to deploy as well as user-friendly; such as, a user can update his/her password securely, whenever he/she needs to.

3. Related Work

This section discusses the literature reviewed for secure healthcare monitoring using wireless sensor networks and general user authentication protocols for wireless sensor networks that have been recently proposed.

Malasri et al. [31] designed and implemented a secure wireless mote-base medical sensor network for healthcare applications. The main components of their scheme are: (i) two-tier architecture is designed for the patient data authentication; (ii) a secure key exchange protocol (i.e., elliptic curve cryptography (ECC)) is used to establish secret shared keys between the sensor nodes and the base station; and (iii) a symmetric encryption/decryption algorithm provides confidentiality and integrity to patient data. Moreover, in their architecture each sensor mote has incorporated a fingerprint scanner; by doing so, the patient’s identity is verified with the aid of a base station. Although, their scheme provides adequate security to patients, it does not care about the strong professional authentication (i.e., who is accessing the patients’ vital signs), whereas user authentication is a prime concern under various laws [29].

Hu et al. [32] have designed and proposed a software and hardware based real-time cardiac patient healthcare monitoring system named ‘tele-cardiology sensor network’ (TSN). TSN is particularly intended for the U.S. healthcare society. It enables real-time healthcare data collection for elderly patients in a large nursing home. In this architecture, a patient’s ECG signals are automatically collected and processed by an ECG sensor and transmitted in a timely way through a wireless channel to an ECG server for further analysis. TSN integrates with large number of wireless ECG communication units; each unit being called a mobile platform. A block cipher algorithm (i.e., skipjack) is used for securing ECG data transmission, and protecting patient privacy. Although their proposal provides privacy in term of confidentiality and achieves integrity, strong user authentication is not addressed effectively.

Huang et al. [18] proposed a secure hierarchical sensor-based healthcare monitoring architecture. The proposed architecture has three network tiers (i.e., sensor network, mobile network, and back-end network), and has considered three real-time healthcare applications (i.e., in-hospital, in-home, and nursing-house) scenarios. The authors used wearable sensor systems (WSS) and wireless sensor motes (WSM) at the sensor network tier. The WSS are Bluetooth enabled and integrated with biomedical sensors; and the WSS are strategically placed on the patient’s body, whereas, the WSMs are deployed within the building, and are used to collect the environmental parameters and transmit through the Zig-bee wireless network standard. WSS and WSM broadcast data securely to the upper layer. Here, WSS uses an advance encryption standard (AES)-based authentication and encryption, while WSM uses a polynomial-based encryption scheme to establish secure point-to-point communication between
two WSMs. In the mobile network tier, mobile computing devices (MCDs) such as PDAs are organized as an \textit{ad-hoc} network and connected to the local station. MCD has the more computational capabilities to analyze the WSS and WSM data. The back-end tier is structured with a fixed station as a server, that provides application level services for lower tiers and process various sensed data from MCDs. Even though Huang \textit{et al.} proposed a secure pervasive hierarchical sensor-based healthcare monitoring, they did not consider the need for strong user authentication, which is an imperative security for healthcare applications according to laws (\textit{i.e.}, HIPAA [29]).

Very recently, Le \textit{et al.} [34] suggested a mutual authentication and access control protocol (MAACE) where legitimate professionals can access their patient’s data. The MAACE facilitates mutual authentication and access control, which is based on elliptic curve cryptography (ECC). Furthermore, these authors argue that their scheme is secure enough in practical attacks, \textit{e.g.}, replay attack, and denial-of-service attacks. Their architecture (\textit{i.e.}, MAACE) consists of three layers: (i) sensor network layer (SN); (ii) coordination network layer (CN); and (iii) data access layer (DA). In their architecture, the SN transmits data to the CN (\textit{i.e.}, PDA, laptop or cell phone), later, the data is forwarded to the DA for future record. Although, Le \textit{et al.}’s protocol facilitates sufficient security against practical attacks, but their scheme susceptible to information-leakage attacks, which could be risky for the patient’s privacy. As a result, patient vital signs could exposed to illegal users (\textit{e.g.}, insurance agents, media persons, \textit{etc.}), which is not acceptable for real-time healthcare applications. Thus, a strong user authentication is required for the healthcare application using sensor networks.

In 2009, Das [42] has proposed two-factor user authentication protocol for wireless sensor networks. Das claimed that his protocol is safe against many attacks (\textit{i.e.}, replay attack, password-guessing attack, user impersonation attack, node compromise attack, and stolen-verifier attack). Later, others [44,46] have pointed out that Das protocol is susceptible to the gateway bypass attack, user impersonation attack, insider attack, \textit{etc}. Furthermore, Das’ protocol does not provide message confidentiality, and mutual authentication between the sensor and the user. Consequently, this protocol is not applicable to healthcare applications using sensor networks.

In [49], Kumar-Lee has shown that some authentication protocols [44,46] have security weaknesses and the computation costs of their protocols are very expensive. Thus, the protocols in [44] and [46] are not suitable for such wireless healthcare applications.

As we can notice from the above literatures, strong user authentication for healthcare application using wireless medical sensor networks has not yet been addressed adequately. Hence, a significant research effort is still required to explore the user authentication for WSN healthcare application. So, next section proposes an efficient-strong authentication protocol, named E-SAP, for healthcare applications using WMSNs.

\textbf{4. The Proposed E-SAP Protocol}

This section presents the proposed efficient-strong authentication protocol (E-SAP) where only legitimate professionals can access the patient’s body data in an authentic manner. The proposed protocol can be applicable to hospitals, homes and clinical environments. The basic idea of E-SAP is quite simple: professionals need to register with the gateway node at hospital registration center. Upon successful registration, the professional receives a smart card from the registration center. Then,
professionals can access the patient physiological information’s from the patient body area sensor network, whenever demanded. In order to prove the professional legitimacy, a professional sends his/her password and smart card based login request to the gateway node. Upon receiving the professional requests the gateway node first authenticates him/her, and then forwards the professional’s request to the dedicated medical sensor, whose data the user is demanding. Thereafter, the medical sensor checks the authenticity of the gateway node and establishes a secure session key between the medical sensor and the professional and responds to the professional. In order to execute the proposed protocol, we have considered the following assumptions:

1. We assumed that the hospital registration center is a trusted authority.
2. The gateway node has three long master keys (i.e., J, K and Q (256 bits long each)).
3. Initially, it is assumed that the gateway and the medical sensor nodes share a long-term secret key \( SK_{gw} = h(Q||ID_g) \) using any key agreement method [50,51].

Table 1 gives a list of notations with descriptions which are used throughout in the paper.

The proposed E-SAP consists of four phases, namely, the professional registration phase, patient registration phase, login and authentication phase, and password change phase.

### Table 1. Notation and Description.

<table>
<thead>
<tr>
<th>Notations</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( U_i )</td>
<td>User’s ( i^{th} ) want to login</td>
</tr>
<tr>
<td>( ID_i )</td>
<td>ID of user ( U_i )</td>
</tr>
<tr>
<td>( PW_i )</td>
<td>Password of user ( U_i )</td>
</tr>
<tr>
<td>( ID_{pt} )</td>
<td>Patient ID</td>
</tr>
<tr>
<td>( GW )</td>
<td>Gateway node</td>
</tr>
<tr>
<td>( ID_g )</td>
<td>Gateway ID</td>
</tr>
<tr>
<td>( Sn )</td>
<td>Sensor node</td>
</tr>
<tr>
<td>( J, K and Q )</td>
<td>Gateway secrets</td>
</tr>
<tr>
<td>( E_{Kol}[\cdot] )</td>
<td>Symmetric encryption using shared key.</td>
</tr>
<tr>
<td>( D_{Kol}[\cdot] )</td>
<td>Symmetric decryption using shared key.</td>
</tr>
<tr>
<td>( M )</td>
<td>User’s generated nonce</td>
</tr>
<tr>
<td>( h(.) )</td>
<td>One-way cryptographic hash function</td>
</tr>
<tr>
<td>( \oplus )</td>
<td>XOR operation</td>
</tr>
<tr>
<td>( | )</td>
<td>Concatenation operation</td>
</tr>
</tbody>
</table>

4.1. Professional Registration Phase

In this phase, the professional initially needs to register with the gateway node at the registration center, as follows:

- User chooses \( ID_i \) and \( PW_i \) and submits to \( GW \) node using secure channel.
- Upon receiving user’s \( ID_i \) and \( PW_i \), the \( GW \) node computes the following:
  a. \( C_{ig} = E_J[ID_i||ID_g] \)
  b. \( N_i = h(ID_i \oplus PW_i \oplus K) \)

Thereafter, the \( GW \) node issues a smart card to the professional with the following \{\( h(.) \), \( C_{ig} \), \( N_i \), \( K \)\}. Here, \( K \) is a long-term \( GW \) node secret, which is securely stored in the smart card.
4.2. Patient Registration Phase

In order to execute the proposed E-SAP, a patient needs to register at the hospital registration center [38], as follows:

- Patient passes his/her name to the registration center.
- After patient registration, registration center choose the suitable sensor kit (i.e., medical sensor and gateway) and designate professionals/users.
- Later, registration center sends patient $ID_{pt}$ and medical sensors kit information (i.e., gateway, sensor etc.) to the designated professionals/users.

Now, the technician deploys wireless medical sensors on the patient body area, strategically, as shown in Figure 2.

4.3. Login and Authentication Phase

This phase is invoked when a professional roams into the patients’ ward and wants to perform a query or to access the patients’ physiological information from the body network. This phase is further divided into login phase and authentication phase.

4.3.1. Login Phase

The professional inserts his/her smart card into the terminal and inputs keys, $ID_i$ and $PW_i$. Upon receiving the login request, the smart card verifies the user locally with pre-stored values and performs operations, as follows:

- $N_i^* = h(ID_i \oplus PW_i \oplus K)$ and compare $N_i^* = N_i$, if yes, then go to the next step, otherwise, terminates the request.
- Compute: $h(ID_i)$ and $CID_i = E_K[h(ID_i)||M||Sn||Cig||T']$. Here, $M$ is a random nonce that is generated by professional system, which is used to establish the secure session key.

Then professional’s system sends message $<CID_i, T'>$ to GW node. Here, $T'$ is the current time stamp of professional system.

4.3.2. Authentication Phase

This phase is invoked when the GW node receives a login request from a professional. Upon receiving the login request at time $T''$, the GW node performs the following and authenticates him/her, as:

- Validate the time $T$: check, if $(T'' - T') \geq \Delta T$, if yes, then rejects the request and aborts any further process. Otherwise, it performs the next steps. Here, $T''$ is the current time of GW node and $\Delta T$ is the time interval for the transmission delay.
- Decrypt sub-message $CID_i$ using key $K$ (i.e., $D_K [CID_i]$) and obtain $h(ID_i)^\delta$, $Sn$, $M$ and $T'^\delta$. Similarly, decrypt sub-message $Cig$ using the shared key $J$ (i.e., $D_J [Cig]$) and obtain $ID_i^*$ and $ID_g^*$.
- Compute $h(ID_i)^*$, and compare $h(ID_i)^* = h(ID_i)^\delta$, $ID_g^* = ID_g$ and $T'' = T'^\delta$, if yes, then the request is authentic; otherwise, terminate any further processes.
• Compute: \( A_i = E_{SKgs}[ID_i||Sn||M||T''']||T''] \), here \( T''' \) is the current time stamp of \( GW \) node. Thereafter, the \( GW \) node sends a message \(<A_i, T'''>\) to the medical sensor that the professional wants to access. Furthermore, \( A_i \) ensures to the medical sensor that the request has come from the legal gateway node.

Upon receiving the gateway node message, the medical sensor node performs the following steps:

• Validate the time \( T \): check, if \((T''' - T''') \geq \Delta T\), if yes, then it rejects the request and aborts any further process. Otherwise, it performs the next steps. Here, \( T''' \) is the current time of the medical sensor node and \( \Delta T \) is the time interval for the transmission delay.

• The medical sensor \((Sn)\) decrypts the sub-message \( A_i \) using shared key \( SKgs \) (i.e., \( D_{SKgs}[A_i] \)), and obtains \( ID_i^*, Sn^*, M^*, T''''=T''' \), and \( T' \).

• Now, \( Sn \) compares \( Sn^* = Sn \) and \( T'''' = T''' \), if not, then it aborts the request; otherwise it continues with the next steps.

• Compute session key \( SK = h(ID_i^*||Sn^*||M^*||T'') \), and message \( L = E_{SK}[Sn^*||M^*||T'''] \), here, \( T''' \) is the current time stamp of the medical sensor node. After that, the medical sensor node sends a response message \(<L, T'''>\) to the professional.

Upon receiving the medical sensor node response, the professional validates the time as follows:

• Validate the time \( T''' \): check, if \((T''' - T'''') \geq \Delta T\), if yes, then it rejects the request and terminates. Otherwise, it continues with the further process. Here, \( T''' \) is the current time of the professional system and \( \Delta T \) is the time interval for the transmission delay.

• The professional system computes \( SK = h(ID_i||Sn||M||T') \).

• Decrypt the message \( L \) using \( SK \), and obtain \( Sn^* \) and \( M^* \). Thereafter, compare \( Sn^* = Sn \) and \( M^* = M \), if yes, then a secure session key has been established; otherwise not.

The flow of the login and authentication phases is shown in Figure 3.

4.4. Password-Change Phase

The password-change phase is invoked when \( U_i \) wants to change/update the password, when he/she requires. The password change procedure is as follows:

• The user inserts his/her smart card into the terminal and enter keys (i.e., \( ID_i \) and \( PW_i \)).

• Smart card performs the operations:

  a. \( N_i^* = h(ID_i \oplus PW_i \oplus K) \)

  b. Compare \( N_i^* = N_i \), if yes, then perform the next step; otherwise abort the operation.

• Enter new password \( PW_{inew} \).

• Compute \( N_{inew} = h(ID_i \oplus PW_{inew} \oplus K) \).

• Replace \( N_i \) with \( N_{inew} \) from the smart card.
5. Formal Analysis of E-SAP Using BAN Logic

Formal analysis ensures that the protocol functions are correctly modeled, and needs to be verified, (i.e., error free) before their real-time implementation [48]. In this regards, this section describes the formal verification of E-SAP using BAN logic, which is popular for formal verification of authentication protocols. The section is divided into: (A) brief overview of the BAN logic, which was introduced by Burrows, Abadi and Needham [48]; and (B) a demonstration of the formal execution and validity proofs of the proposed E-SAP using the BAN authentication logic model.
5.1. BAN Logic

The BAN logic is a popular authentication protocols analysis model, and it is useful to prove the validity of authentication and key establishment protocols, for more details the readers may refer to [48]. The notations used in BAN logic are defined as follows:

- **P believes X**: The main construct of logic is ‘P believes X’ (i.e., the principal P believes on X) or P would be entitled to believe X.
- **P sees X**: Only ‘P sees X’, i.e., suppose someone has sent a confidential message (i.e., encrypted message) containing X to P, then P can read X (i.e., after performing some decryption).
- **P said X**: The principal ‘P once said X’; means, at some time the principal P sent a message including X.
- **P controls X**: The principal ‘P has controls over X’; means, the principal P is an authority on X and should be trusted (e.g., a server is often assumed trusted and generate secret keys properly).
- **Fresh(X)**: Fresh(X) means, X has not been sent recent in a message during the protocol execution. Furthermore, Fresh(X) protects from replay attack.
- **PleftrightarrowQ**: The principal P and Q may use secret shared key K for secure communication. The keys K will never be disclosed to others except for the designated principals (i.e., P and Q).
- **{X}K**: Means the formula X is encrypted using the key K.
- **<X>Y**: The formula X is combined with secret parameter Y.

Now, we have defines some logical rules that we use in proofs, and which are directly adopted from [49], as follows:

- **Message-meaning rule**

  \[
  \begin{align*}
  P \text{ believes } Q \leftrightarrow P, & \text{ P sees } \{X\}K \\
  P \text{ believes } Q \text{ said } X
  \end{align*}
  \]

- **Nonce-verification rule**

  \[
  \begin{align*}
  P \text{ believes } \text{fresh} (X), & \text{ P believes } Q \text{ said } X \\
  P \text{ believes } Q \text{ believes } X
  \end{align*}
  \]

- **Controls rule**

  \[
  \begin{align*}
  P \text{ believes } Q \text{ controls } X, & \text{ P believes } Q \text{ believes } X \\
  P \text{ believes } X
  \end{align*}
  \]

- **Fresh rule**

  \[
  \begin{align*}
  P \text{ believes } \text{fresh} (X) \\
  P \text{ believes } \text{fresh} (X,Y)
  \end{align*}
  \]
5.2. Formal Verification of the Proposed E-SAP

This sub-section demonstrates the formal verification of our proposed protocol using the BAN logic analysis model [48]. The main principals of E-SAP are: user ($U_i$), gateway ($GW$) and medical sensor node ($Sn$). The following symbols are used: (a) the secret keys are $J$, $K$, $SK_{gs}$ and $SK$; (b) the time-stamps are $T'$, $T''$ and $T^*$. The main goal of formal verification is to establish a secure session key between the user and the medical sensor node. To perform the formal verification of E-SAP, we use the following logical postulates:

- $U_i$ believes $U_i \leftrightarrow Sn$, $SK$
- $U_i$ believes $Sn$ believes $U_i \leftrightarrow Sn$
- $Sn$ believes $U_i$ believes $U_i \leftrightarrow Sn$

The protocol messages (as shown in Figure 3) are needs to be transform into the idealized form, as shown in Table 2:

<table>
<thead>
<tr>
<th>E-SAP messages:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message1: $U_i \rightarrow GW: &lt;CID_i, T'&gt; (i.e., E_K[h(ID_i)</td>
</tr>
<tr>
<td>Message2: $GW \rightarrow Sn: &lt;A_i, T''&gt; (i.e., E_{SK_{gs}}[ID_i</td>
</tr>
<tr>
<td>Message 3: $Sn \rightarrow U_i: &lt;L, T*&gt; (i.e., E_K[Sn</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Idealized form:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message 1: $U_i \rightarrow GW: {h(ID_i)</td>
</tr>
<tr>
<td>Message 2: $GW \rightarrow Sn: {ID_i</td>
</tr>
<tr>
<td>Message 3: $Sn \rightarrow U_i: {Sn</td>
</tr>
</tbody>
</table>

Session key $SK = h(ID_i, *||Sn||M*||T')$

E-SAP formal analysis using BAN logic required further assumptions, as follows:

A1) $U_i$ believes $GW \leftrightarrow U_i$
A2) $GW$ believes $U_i \leftrightarrow GW$
A3) $GW$ believes $U_i \leftrightarrow GW$
A4) $GW$ believes $Sn \leftrightarrow GW$
A5) $Sn$ believes $GW \leftrightarrow Sn$
A6) $Sn$ believes $U_i \leftrightarrow Sn$
A7) $U_i$ believes $Sn \leftrightarrow U_i$
A8) $Sn$ believes (U_i controls $U_i \leftrightarrow Sn$)
A9) $GW$ believes $GW \leftrightarrow GW$
A10) $GW$ believes (U_i controls ID_i)
A11) $U_i$ believes fresh $(M)$
A12) \textit{U, believes Sn fresh (T*)}
A13) \textit{GW believes Ui fresh (T*)}
A14) \textit{Sn believes GW fresh (T'*)}
A15) \textit{Sn believes GW fresh (M)}
A16) \textit{Sn believes (GW controls IDi)}

Based on the above assumptions and BAN logic rules, we perform the verification of the proposed E-SAP, as shown in Table 3.

\textbf{Table 3.} Formal verification of E-SAP using BAN logic model.

| Message 1: \textit{U→GW:} | \textit{h(IDi)||M||Sn||IDg||T'∗}K, T' |
|--------------------------|------------------------------------------|
| S1) GW sees \{h(IDi)||M||Sn||IDg||T'∗\}K, T' // by seeing rule |
| S2) GW believes \textit{U,} \leftrightarrow \textit{GW} // by A1, A2, S1, message-meaning rule |
| S3) GW believes \textit{U,} \textit{controls IDi} // by A10, controls rule |
| S4) GW believes \textit{U,} \textit{fresh (M)} // by message-meaning and fresh rule |
| S5) GW believes \textit{GW} \leftrightarrow \textit{GW} // by A9, S1, message-meaning rule |
| S6) GW believes \textit{U,} \textit{fresh (T')} // by S4, A13, fresh rule |
| S7) GW believes \textit{Sn said} \{IDi||Sn||M||T''||T'}SKgs , T'' // by message-meaning rule |

| Message 2: \textit{GW→Sn:} | \{IDi||Sn||M||T''||T'}SKgs , T'' |
|--------------------------|------------------------------------------|
| S8) Sn sees \{IDi||Sn||M||T''||T'}SKgs , T'' // by seeing rule |
| S9) Sn believes \textit{GW fresh}(T'') // by A14, fresh rule |
| S10) Sn believes \textit{GW} \leftrightarrow \textit{Sn} // by A5, S8, message-meaning rule |
| S11) Sn believes (GW \textit{controls IDi}) // by S8, A16, controls rule |
| S12) Sn believes GW\textit{ fresh(M)} // by A15, fresh rule |
| S13) Sn believes \textit{U,} \leftrightarrow \textit{Sn} // by A6, S8, message-meaning rule |
| S14 Sn believes \textit{U,} \textit{said} \{Sn||M*||T*/SK , T* |

| Message 3: \textit{Sn→U:} | \{Sn||M*||T*/SK , T* |
|--------------------------|------------------------------------------|
| S15) \textit{Ui sees} \{Sn||M*||T*/SK , T* // by seeing rule |
| S16) \textit{Ui believes Sn fresh (T*)} // by A12, fresh rule |
| S17) \textit{Ui believes Sn} \leftrightarrow \textit{U,} // by A7, S15, message-meaning rule |
| S18)\textit{Ui believes Sn fresh(M)} // by fresh rule |
| S19) \textit{Ui believes Sn believes U,} \leftrightarrow \textit{Sn} // by S 15, message-meaning rule |
| S20) \textit{Ui believes U,} \leftrightarrow \textit{Sn} |

As we can see from the above verification, \textit{A7, S13, S19} and \textit{S20} establish the secure session key between the user and the medical sensor. Furthermore, \textit{A3, A10, S5, S11, S19} and \textit{S20} verify mutual authentication between the user and medical sensor using the gateway. Hence, the goal of E-SAP is
achieved (i.e., secure session key has established and only authentic users can access an individual’s body information from the wireless medical sensor networks).

6. E-SAP Evaluation

This section discusses the security analysis and functionality analysis of the proposed E-SAP for healthcare application using medical sensor networks. Further, we present a performance analysis of E-SAP. The following assumptions are considered before evaluating the proposed protocol, which is based on a smart card and password (i.e., two-factor):

- The adversary has total control of wireless communication; he/she may intercept, delete or alter any message in the communication (recall the discussion of attack model in Section 2).
- The attacker either obtains a user’s password, or extracts the secrets from the smart card through [52,53], but not both (i.e., password and smart card) at the same time [50].
- Assumed that, extracting secrets from smart card is quite complex and some smart card manufacturer provide countermeasures against side channel attacks [42,50]. In [54] authors proposed some software countermeasures against power analysis attack.
- We assumed that the symmetric cryptosystem are secure enough to protect patient physiological information from cracking, and any encrypted text cannot be decrypted without having the secret keys, which is known only to the trusted entities (i.e., user, gateway, medical sensor and hospital registration center).

6.1. Security Analysis

This sub-section shows the proposed protocol is secure against many practical attacks. In additions, the proposed E-SAP facilitates: confidentiality, mutual authentication between the user and the medical sensor, a secure session key establishment between the medical sensor node and the professionals, and professionals can change their password, securely.

Replay attack: The proposed protocol is resistant to replay attacks. Assume that an adversary re-play the old captured messages to the gateway (i.e., \(<CID_i,T'\>\), the medical sensor (i.e., \(<A_i,T''\>\), and the user (i.e., \(<L,T*\>\)). However, he/she (attacker) cannot pass the old messages, because all messages are validated by the fresh time stamps, which are contained in the protocol messages (i.e., \((T'-T) ≥ \Delta T, (T'''-T'') ≥ \Delta T\) and \((T**-T*) ≥ \Delta T\).

Masquerading user attack: An attacker cannot masquerade as the professional \((U_i)\). Suppose an adversary were able to forge a login message \(<CID_i,T'>\). Now the adversary will try to login into the WMSN with a modified message \(<CID_i*,T'>\). He/she cannot pass the fake message because the forged \(CID_i^*\) will not be verified at the gateway node and the gateway node cannot get the original message (i.e., \((h(ID_i)||S_n||M||C_{ig}||T')\)) by decrypting the fake \(CID_i^*\).

Masquerading gateway attack: An attacker cannot impersonate a gateway, since he/she does not have any idea how to get \(J, K\) and \(SK_{gs}\) from the protocol messages. So, masquerading as the gateway is not applicable to the E-SAP.
**Gateway secret guessing attack:** The proposed scheme is secure against the gateway secret guessing attack. The gateway has three master keys (i.e., $J$, $K$ and $Q$), which are not transmitted as plaintext. Hence, E-SAP is secure against gateway secret guessing.

**Stolen verifier attack:** In [43], a user table (i.e., $ID_i$ and $PW_i$) is stored on the gateway node, which may be a high risk to breach the security of protocols. In contrast, the E-SAP protocol does not use any $ID_i$ table and password table. So any stolen-verifier attack will not applicable on the proposed protocol.

**Password guessing attack:** An attacker cannot guess the password in our scheme. In the proposed protocol password is not passing as plaintext, instead $N_i = h(ID_i \oplus PW_i \oplus K)$, so password guessing is not possible.

**Mutual authentication:** The proposed E-SAP provides mutual authentication between the user and the medical sensor. As shown in the Figure 3, the gateway sends message $<A_i, T'''>$ to the medical sensor. Here, $A_i = ES_{Kgs}[ID_i||Sn||M||T'''||T]$ and it ensures to the medical sensor that the message has come from the legitimate gateway node. Thus, the medical sensor believes that the user is a legitimate user. Furthermore, when the user receives a medical sensor message $<L, T*>$, then he/she verifies the medical sensor (i.e., whether real or not). Hence, the proposed protocol achieves mutual authentication between the user and the medical sensor.

**Information-leakage attack:** The protocol information-leakage gives room to the attackers, which could be harmful for the patient privacy. In E-SAP, suppose an adversary eavesdrops the protocol messages (i.e., $<CID_i, T'>$ $<A_i, T'''>$ and $<L, T*>$). Here, the sub-message $CID_i$ is encrypted using shared secret $K$, the message $A_i$ is encrypted using shared $SK_{gs}$, and the sub-message $L$ is encrypted using $SK$. Therefore, E-SAP messages information’s are not leaked during communication. As a result, information-leakage attacks not applicable to our protocol.

**Secure session key:** The proposed E-SAP establishes a secure session key between the user and the medical sensor node after the authentication phase taken place. As we can see in Figure 3, a session key ($SK = h(ID_i^*||Sn||M^*||T^)$) is setup between the medical sensor node and the user. Furthermore, the established session key provides confidentiality for subsequent communication; and for each session the session key will fresh.

**Confidentiality:** Confidentiality is a paramount requirement for healthcare application using wireless medical sensor networks. In the proposed E-SAP, the session key could be used for further secure subsequent communication between both (i.e., user and medical sensor node may encrypt patient physiological information’s using the session key ($SK$)). Furthermore, the proposed protocol provides air message confidentiality to their messages ($CID_i=ES_K[h(ID_i)||M||Sn||C_{ig}||T]$), $A_i=ES_{Kgs}[ID_i||Sn||M||T''||T]$and $L = ES_K[Sn||M^*||T^]$).

**Secure password change:** In the password-change phase, the proposed protocol first verifies the user’s old password and identity, and only then requests a new password. Otherwise it rejects the password change request. Thus, the proposed scheme is secure against changed passwords.

6.2. E-SAP Functionality Analysis

This subsection shows the E-SAP functionality and makes a comparison with related schemes (i.e., Le et al. [34], Das [42], Vaidya et al. [43] and He et al. [46]). As shown in Table 4, the proposed protocol provides more functionality such as strong user authentication, mutual authentication between
the user and the medical sensor node, it establishes a secure session key for the user and the medical sensor node, message confidentiality and professionals are able to change their password, whereas in [34,42,43] and [46] the schemes provides less security functionality, which are paramount requirements (recall section 2-C) for wireless healthcare applications. Further, it can be seen from Table 4 that the proposed E-SAP is robust against many popular types of attacks (e.g., replay attack, masquerade attack, gateway secret guessing attack, and information-leakage attack) as compared to other schemes. It is worth notice that our protocol provides indispensable security features, whereas, the schemes in [34,42,43,46] provide less security functionality for real-time healthcare applications.

<table>
<thead>
<tr>
<th>Functionalities</th>
<th>[34]</th>
<th>[42]</th>
<th>[43]</th>
<th>[46]</th>
<th>Proposed E-SAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strong user authentication</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Mutual authentication between (U_i) and (Sn)</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Session key establishment</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Secure password change</td>
<td>NA</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Message confidentiality</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Protection to replay message</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Secure against (GW) secret key guessing attack</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Secure against user masquerading attack</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Secure against gateway masquerading attack</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Secure against Information-leakage attack</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Protocol formal verification</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

NA: Not applicable.

6.3. E-SAP Performance Evaluation

This subsection evaluates the performance of proposed protocol in term of computation cost, communication cost and compares the results with [34,42,43,46].

The performance evaluation parameters are:

\[ T_{pu}: \] public-key computation, \( T_{pr}: \) private-key computation,

H (performing one hash function), S (symmetric-cryptosystem), and M (performing one message authentication code).

**Computation cost:** The medical sensor devices (i.e., gateway node and sensor node) have limited power resources and computation capability. Therefore, the computation cost is a prime factor for resource constrained devices. The user registration computation cost is a one-time task and it is not a main concern, whereas the login and authentication computation cost are a prime concern due to the resource constrained nature of the gateway node and the medical sensors nodes. Table 5 shows the computation cost of the proposed E-SAP and related schemes, i.e., Das [42], Vaidya et al. [43], and He et al. [46]. It is easy to see from Table 5, in registration phase the proposed E-SAP needs only 1H and 1S at GW node, whereas [42,43,46] require, 3H, 4H and 5H, respectively, which a is high computation cost at GW node.

Further, the Le et al. [34] scheme requires modular exponentiation to compute the public and private keys, so their scheme is computationally expensive and time-consuming, and it also needs to
generate and verify digital certificates. In the login and authentication phase, E-SAP requires 6H and 7S, and provides more security. In contrast [34,42,43,46] require 4H+4S+6M, 7H, 9H and 7H, respectively, and provide less security services. This is due to the fact that the proposed E-SAP incurred more computation cost and provides paramount security functionality to healthcare applications as compared to [42,43,46]. Thus, the computation cost of E-SAP is well-suited to the healthcare applications using wireless medical sensor networks.

**Table 5. Performance comparison of E-SAP with existing schemes.**

<table>
<thead>
<tr>
<th>Schemes</th>
<th>Registration</th>
<th>Login and authentication</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>User</td>
<td>GW</td>
</tr>
<tr>
<td>Le et al.’s [32]</td>
<td>$T_{pu}+T_{pr}$</td>
<td>$T_{pr}$</td>
</tr>
<tr>
<td>Das’s [40]</td>
<td>–</td>
<td>3H</td>
</tr>
<tr>
<td>Vaidya et al.’s [41]</td>
<td>2H</td>
<td>2H</td>
</tr>
<tr>
<td>He et al.’s [44]</td>
<td>1H</td>
<td>5H</td>
</tr>
<tr>
<td><strong>Proposed E-SAP</strong></td>
<td>–</td>
<td>1H+1S</td>
</tr>
</tbody>
</table>

**Communication cost:** The communication cost is an important issue in wireless communication, (i.e., more message exchanges consume more power). From Figure 3, it is easy to visualize that the proposed E-SAP requires three message exchanges between the user, the gateway and the medical sensor, whereas the schemes in [42] and [46] require three message exchanges, and [34] and [43] require four exchanges. Hence, the proposed protocol is well-suited and quite simple in enhancing the wireless communication security for healthcare application.

Considering the functionality, computation cost, and communication cost of E-SAP, it is clear that our protocol is more efficient for healthcare applications using medical sensor networks as compared to others [34,42,43,46].

7. Conclusions

Wireless medical sensors offer services to professionals; but what do we do to verify the professionals (i.e., authentic or not). That poses a question to researchers, how to protect medical sensor data from illegal users?

In order to solve the above questions, this paper proposed E-SAP, an efficient-strong user authentication protocol for healthcare application using wireless medical sensor networks. E-SAP utilizes two-factor security features and provides strong user authentication, confidentiality and session key establishment for healthcare application using WMSNs. It is noteworthy that E-SAP is more capable in terms of security services, computation and communication cost, as compared to other existing protocols. Furthermore, through intensive analysis (i.e., BAN logic authentication model) we have shown that E-SAP achieves its stated security goals and is defensive against many popular types of attacks. It is a well-suited protocol for hospital, homecare, and clinic healthcare applications using wireless medical sensors.

The future directions for this study are: (1) to develop a real-time heterogeneous biomedical sensor network for healthcare monitoring, (2) implement E-SAP on a real-time test-bed for healthcare application, and (3) more focus on access control in patient mobility scenarios and strong patient privacy.
Acknowledgments

This research work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (Grant Number: 2011-0004713 and 2011-0023076).

References


© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).
Focus on People and Organization Issues

Antecedents of the People and Organizational Aspects of Medical Informatics: Review of the Literature

NANCY M. LORENZI, PHD, ROBERT T. RILEY, PHD, ANDREW J. C. BLYTH, PHD, GRAY SOUTHON, PHD, BRADLEY J. DIXON, MBA

Abstract People and organizational issues are critical in both implementing medical informatics systems and in dealing with the altered organizations that new systems often create. The people and organizational issues area—like medical informatics itself—is a blend of many disciplines. The academic disciplines of psychology, sociology, social psychology, social anthropology, organizational behavior and organizational development, management, and cognitive sciences are rich with research with significant potential to ease the introduction and on-going use of information technology in today’s complex health systems. These academic areas contribute research data and core information for better understanding of such issues as the importance of and processes for creating future direction; managing a complex change process; effective strategies for involving individuals and groups in the informatics effort; and effectively managing the altered organization. This article reviews the behavioral and business referent disciplines that can potentially contribute to improved implementations and on-going management of change in the medical informatics arena.


In the first stages of the information revolution in health care, technical hardware and software issues understandably received far more attention than people and organizational issues. Many of the early implementations were in the “business” areas of health care, which are characterized by hierarchical management structures that are similar to other businesses. When early clinical implementations did occur, they were often of limited scope, affected relatively few people, and had strong local champions.

Today’s informatics implementations often involve far larger systems with wide-ranging effects and requiring the cooperation of many people. These systems are also being introduced into organizations that are more complex and that are often traumatized by other changes, such as downsizing or mergers. While some implementation failures do make it into print, others are only discussed in private at conferences. The complexity of the environment in which the technology operates does make failure analysis difficult and controversial. Today’s informatics implementations—and especially the larger scale ones—are becoming increasingly dependent upon how well the people and organizational issues are managed.
Link Between Medical Informatics and Change

The relationship between technological change and organizational change is always an interesting one in a classic chicken-or-egg sense. Do the technological advances drive the change, as in the case of developments in medical imaging? Or does the technology merely enable changes that are largely driven by non-technical forces, as in the case of information systems developed to support cost reductions forced by economic pressures?

People can easily be overwhelmed by change, especially within large organizations where they may perceive they have little or no voice in or control over the changes they believe are descending upon them. The typical response is fight or flight, not cooperation. Managers often interpret such human resistance to change as "stubbornness" or as "not being on the team." This reaction solves nothing in terms of reducing resistance to change or gaining acceptance of it. Many managers do not accept that they are regarded as imposing "life-threatening" changes and establishing "no-win" adversary relationships between management and those below in the organization. Sometime managers try to disguise the impending change with what they consider innocuous names, (e.g., organizational effectiveness) or they adopt a metaphor (e.g., architecture for the future). In this case, we are referring to the parent organization, but the same analogy applies to the medical informatics area as well. The end result of people feeling disenfranchised is inevitable, as the following examples from around the world attest.

- In 1988, the University of Virginia Medical Center began implementing a medical information system based on mandatory physician order entry. The implementation process was much more difficult than expected. The program experienced considerable delays and cost much more than was originally estimated. Although there were some legitimate questions concerning the user-friendliness of the new technology, these were less significant than the cultural and individual behavioral problems encountered. The new system challenged basic institutional assumptions; it disturbed traditional patterns of conduct; and it forced people to modify established practice routines. Real progress toward the integration of the system into the center's operational culture occurred only after a senior management team representing important sectors of the hospital staff and administration began meeting regularly to address the institution-wide issues that had been raised.

- In 1990, at Calgary's Foothills Hospital, a war began between the users of the new information system called OSCAR and the people responsible for delivering the system. At the heart of the problem was the perception that OSCAR was prohibiting the medical staff at the hospital from performing their functions. The medical staff perceived that management was attempting to impose its will over them, restricting and redefining their work roles and pat-
terns. The medical staff consequently refused to work with OSCAR.  

In 1992, the Computer Aided Dispatch system for the London Ambulance Service failed, primarily because the information system supported the values and norms of senior management, not the values and norms of the ambulance crews and the command and control staff.

Each implementation was made with the best intentions and in response to the perceived changes needed to function in today’s health care environment. However, it appears that the implementors did not consider an organizational change model or the people and organizational issues associated with these implementations.

Figure 1 shows a basic four-stage model of organizational change that applies at the general level as well as for the implementation of change through informatics systems. The initial steady state (1) is affected by some impetus for change—whether technical or non-technical. This impetus might be a quite visionary perception or it might be a rather tardy reaction to environmental change. The organization then conceptualizes the desired outcome (2) and implements the change (3). If the change is not trivial, the organization is itself altered in various ways by the change (4). Over time, the altered organization adjusts and becomes the “new” initial state for the next change. In a complex organization functioning in a volatile environment, various portions of the organization are passing through various stages of the process at varying rates. There is a continuous cycle of change that organizations are constantly managing.

The people and organizational issues area—like medical informatics itself—is a blend of many disciplines. These referent disciplines include psychology, sociology, social psychology, social anthropology, organizational behavior and organizational development, management, and cognitive sciences. They are displayed as flowing through the arrow toward Stages 2–4 in Figure 1. The research from these disciplines focuses on individuals, groups, and organizations and contributes to understanding and designing effective strategies for the non-technical side of change, including informatics changes. Each of the disciplines contains knowledge and skills that can contribute to more effective passage through all the latter three change stages.

Those not educated in these areas have widely varying reactions to social science knowledge. Some regard it as “just common sense.” To others, it represents a mysterious “black box” that they may appreciate but do not pretend to understand. In reality, there has been much research and knowledge accumulated within the black box that is available to increase the quality of the change outcomes. Some of the more relevant topics researched are outlined in Table 1. To illustrate the breadth of the research, the topics are listed and then categorized according to their area of focus: e.g., individuals (I), Groups (G), and Organizations (O). Some of the topics listed have research that involves all of the broad areas of focus. Experts in the various areas might argue that an “X” could be placed in each box. However, we have elected to focus on a portion of the research that we believe affects medical informatics the most.

While these topical areas may not seem difficult or challenging to the informatics change leader, failing to address them well usually leads more to the failure of a systems effort than hardware or software deficiencies do.

### Table 1

A Listing of Some of the Topics Researched In the Referent Disciplines According to Individual, Group, and/or Organization Area of Focus

<table>
<thead>
<tr>
<th>Research Topics</th>
<th>Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>User cordial information system design</td>
<td>X</td>
</tr>
<tr>
<td>Empowering end users</td>
<td>X X</td>
</tr>
<tr>
<td>Behaviors lead to system success/failure</td>
<td>X X</td>
</tr>
<tr>
<td>Role of professional values/cultures</td>
<td>X X</td>
</tr>
<tr>
<td>Why and how people accept or resist change</td>
<td>X X</td>
</tr>
<tr>
<td>User involvement and participation</td>
<td>X X</td>
</tr>
<tr>
<td>User satisfaction/acceptance and attitudes</td>
<td>X X</td>
</tr>
<tr>
<td>Strategies for managing complex change</td>
<td>X</td>
</tr>
<tr>
<td>Organizational structures</td>
<td>X</td>
</tr>
<tr>
<td>Role of information in organizations</td>
<td>X</td>
</tr>
<tr>
<td>Total quality management in the organization</td>
<td>X</td>
</tr>
<tr>
<td>How ideas/technology are diffused through an organization</td>
<td>X</td>
</tr>
<tr>
<td>Organizational political process</td>
<td>X</td>
</tr>
<tr>
<td>Types of Organizations</td>
<td>X</td>
</tr>
<tr>
<td>Transforming health care</td>
<td>X</td>
</tr>
<tr>
<td>Reengineering</td>
<td>X X</td>
</tr>
<tr>
<td>The role of social control systems</td>
<td>X X</td>
</tr>
<tr>
<td>Organizational culture</td>
<td>X X</td>
</tr>
<tr>
<td>Strategies to motivate people</td>
<td>X X</td>
</tr>
<tr>
<td>Effective leadership strategies</td>
<td>X X X</td>
</tr>
<tr>
<td>Decision making</td>
<td>X X X</td>
</tr>
<tr>
<td>Implementation strategies</td>
<td>X X X</td>
</tr>
</tbody>
</table>

(81)
Knowledge from the Core Referent Disciplines

The antecedents for learning about and understanding the processes that might have led to greater success are known and have been known for some time. We will briefly review the extensive research that pertains to medical informatics according to (1) business and organizational issues; (2) individuals and groups; and (3) management of information issues.

Business and Organizational Issues

It has always been difficult to develop a clear and concise concept of what an organization is and how it behaves. The classic concepts developed early this century were dominated by the somewhat mechanistic models of scientific management and bureaucracy. These models emphasized formal regulations, clear lines of accountability, and compliance with authority. In the middle of the century, there was a move to more "organic" concepts to take better account of the more human characteristics of organizations.22 As understanding matured, a wider range of perspectives developed.23 Some of the more important of these are:

- **Ecology**: The organization is seen as a discrete entity that has a primary objective of surviving in its environment. Issues of financial viability, strategic goals, market niches, and competition are paramount.
- **Structure**: The organization is seen as a set of control and accountability systems that must be attuned to the particular needs of that organization. Structure may be formal and hierarchical or quite loose and informal.24
- **Politics**: The organization is composed of and affected by a wide range of often conflicting interest groups that make demands on it. It is the interaction among these various forces that determines the organization’s behavior.25
- **Culture**: There are common factors such as symbols, customs, values, and assumptions that characterize and influence the way people in an organization think and act. This "culture" is an important determinant of the way that the organization functions.26
- **Psychology**: Organizations have sets of behaviors that mimic those of human behaviors.27
- **Human Resource**: Staff are seen as the organization’s principal asset, and the most important task is the development of staff skills, commitment, and initiative.28
- **Functional**: The nature of the organization arises from its function, the technology and systems that it uses, and the skills that it requires.
- **Intelligence**: The core assets of many organizations are "soft" and lie in a combination of data, programs, procedures, skills, knowledge, culture, and values that make the organization function.29
- **Decision Making**: This is seen as the basis of the organization.30

This diverse set of perspectives illustrates the complexity involved in developing a coherent concept of an organization. Each perspective deals with very different sets of characteristics and demands quite different ways of thinking. Yet they are intimately related and must be integrated to provide a comprehensive view of the organization. However, there is no consensus of how this integration is to be done.31

Organizational Types

Not only are there many ways of looking at organizations, but their basic structures can change dramatically. Mintzberg identified five basic types of organization:59

- **Simple.** Typically, this organization was started by one person and is still dominated by that person. Its style is idiosyncratic, and a person’s position and power are very much dependent on his or her relationship to the leader.
- **Machine Bureaucracy.** This organization is controlled from the top by formal processes. There are strong reporting and accountability lines emanating from the chief executive, and a high degree of compliance is required. This type of structure can manage large organizations, but it is not effective in managing the unpredictable or in enabling staff to take initiative.
- **Professional Bureaucracy.** This organization is built around professional activities. Because of the complexity of professional work, it cannot be formally directed. The organization is primarily a supportive framework in which professionals can operate largely autonomously. They are controlled more by professional values and culture than by formal means.
- **Adhocracy.** This is a very adaptive organization set up around a specific project. The demands of the task dictate the structure, which may change quite frequently. It is driven very much by the skills and the commitment of the people involved.
- **Divisionalized.** When large organizations take on diverse activities, they may become divisionalized. Each division is a semi-autonomous unit with over-
sight from the center that provides a base for effective management in each division.

Two other types of organization can be added to Mintzberg’s five:

- **Federal.** These are large, fairly homogeneous organizations that may be broken down into regional units for effective management (e.g., some of today’s health care enterprise systems). These sub-units may be put into competition with each other, or they may work together to enhance their performance.

- **Networked.** These organizations are “non” organizations in many respects. They are independent individuals or sub-organizations that link opportunistically to address a particular task. These occur typically in primary health care; a general practitioner may refer a patient to a surgeon, who then engages a hospital in which to do the surgery and subsequently refers the patient to a physiotherapist for rehabilitation.\(^5\)

While these are archetypal organizations, any complex organization may be a combination of types. A large hospital system may have a federal structure, with management having largely a machine bureaucracy form, yet its autonomous medical staff could give it a substantial professional bureaucratic nature. However, some sections may be run by a strong autocrat, similar to the primitive form. In its wider function, the system may be part of a network operation, and many of its internal functions may be driven by network relationships. The types of organizational structures will have an impact on the implementation of an information system.

**Evolution of Organizational Concepts**

While organizations, large and small, have been managed in many different ways for centuries, the formal understanding of how this is done is relatively recent. The formalization began early in this century, dominated by Fayol, Weber, and Taylor. While there were many aspects to their theories, their overall emphasis was on a formal organizational structure, a line of command, and specialization of function.\(^33\) The critical information involved was the command and control structure—providing instructions and ensuring compliance—and the knowledge of optimum production techniques, which was the province of the industrial engineer. The worker’s role was to comply with instructions. These principles were modified to some extent by the human-relations movement. Western industrial development in this century was based on these principles up to and into the sixties.

In the 1970s, however, the Japanese economic recovery started to threaten many Western industries. The principles on which many Japanese industries were based seemed dramatically different from those in the West—principles such as singular commitment to the company, life-time employment, and collective decision making. However, perhaps the most dramatic difference was that of “Total Quality Management” (TQM), a principle that had been developed in the United States by Deming but adopted in Japan. Instead of limiting production staff to observing set procedures, TQM gave production staff the tools and authority to monitor the quality and efficiency of their work and to find better ways of working. The focus of activity moved from the individual worker to the team. The critical information in this environment was what the production workers chose to collect to assist them in addressing the important problems at each stage. Although attempts were made to transfer the statistical tools of TQM to the service environment, it was clear that they required considerable modification.\(^36\) While TQM is now well recognized, there are concerns that there still remain basic conflicts with management theory.\(^37,38\)

Into the nineties, the increasing importance of the knowledge industry became evident from the rapidly advancing capabilities of information technology and the complexities of the services that were being provided. In 1996, it was recognized that the element that had been largely ignored in these organizational concepts was the role of the professional. According to Quinn, this was most surprising because “the professional intellect creates most of the value in the new economy.”\(^39\) This, in turn, requires rather special forms of organization.\(^40\)

Thus, there is a tremendous dynamism in the understanding and restructuring of organizations.\(^41,42\) National culture can also have an important impact on the style of management and the nature of organizations.\(^43\) There is considerable confusion and a proliferation of different theories and techniques that are often marketed as the solution to a particular organization’s problems.\(^44\) Such terms as “excellence,” Management by Objectives, Management By Walking Around, Continuous Quality Improvement, Strategic Planning, Business Process Reengineering, downsizing, right-sizing, and restructuring have all had their advocates and detractors.\(^45\) Many organizations have gone through such repetitive changes that they have lost their organizational memory and hence any resistance to such abuses.\(^46,47\)

Overall, organizations are extraordinarily complex entities, and our understanding of them and their management is still in a state of considerable flux and con-
Health Service Organizations

Health service organizations are particularly complex for a number of reasons.

- Health services are provided by a wide range of institutions, ranging from major specialty hospitals to a complex of community hospitals, small clinics, and individual professionals. These components work in conjunction with each other, forming network organizations that are much larger than each component.
- Public, not-for-profit, and volunteer organizations are often dominant in the health services arena. There are typically strong humanitarian values that may override commercial or financial objectives.
- Professionals dominate in both the definition and the execution of the task. In some organizations, they also dominate in the management and the governance of the organization.
- The definition of the task and its objectives are in many cases very difficult to establish in advance. Often the task is related to such fundamental human questions as: What is life? What can we expect of life? How should we die?
- The health system is undergoing fundamental structural change in most places in the world, with many countries following quite different principles.

The ever-advancing technology has prompted dreams of being able to quantify health service performance and outcomes, thereby enabling managers to manage clinical operations more effectively and enabling consumers to choose the best provider—a step toward a genuine health services market. Previous research about organizational issues provide the antecedent to help us identify the relationship between visions and reality to manage this process.

Individuals and Groups

The coffee machine and the computer have influenced twentieth-century organizational development because both mediate social relationships. While there are many social relationships that we could include, we have selected three—motivation, culture, and leadership.

Motivation

Two theories—Field and Reinforcement—are relevant to understanding individuals and groups within organizations. Field Theory was established by Kurt Lewin, who focused much of his attention on motivation. Field theory is primarily concerned with the purposes that underlie behavior and the goals toward or away from which behavior is directed. Lewin researched a wide variety of topics, including motivation of individuals, decision making both by individuals and in groups, group leadership, resumption of interrupted activities, level of aspiration, and how people interact with each other. Field theory establishes the need for active involvement of individuals within any change process. If people are actively involved, they will be motivated toward making the end effort a success. If people are not involved in the change process, they can and will resist because they might see the changes as detrimental to them personally.

Reinforcement theorists are from the discipline of psychology and focused their efforts on the motivation and conditioning. B. F. Skinner, a major reinforcement theorist, said that behavior can be reinforced either positively or negatively. Some medical informatics implementations have used positive reinforcement techniques. One example is the design of training programs to meet physicians’ needs. Training aids for physicians are prepared on 3 × 5 cards because most physicians are accustomed to keeping pertinent information in this manner. Another reinforcer is to have a clinically knowledgeable person (e.g., nurse, pharmacist, physician) to explain the system. Both of these steps help create positive reinforcers to meet physician needs. The positive training reinforcement will lead to a greater awareness of the system capabilities and acceptance of the system. On the other hand, announcing a very lengthy training session at a time that conflicts with known medical timetables in the organization will lead to distrust of the system and will reinforce any negative perceptions that physicians may have acquired. Once a series of negative reinforcers begins in the system, there is a snowball effect on the new system.

Computers will be increasingly used to support clinical practice. When adopting and utilizing information technology are seen to enable and empower people and groups, people will support the system and its development. The issue is ownership. If the group members perceive that they own the problem and the solution, they will work with the developers to make the system work.

Culture

The culture of an organization will draw upon the culture of the society in which it is set and the other organizations and individuals with which it interacts. Social anthropology is the study of how and why cultures are created and mediated within a society. The
mediation can only take place through the application of a social structure and a communication process.

Every culture supports a political and social value system, and these systems are used to create and define a social control system to support and aid the organization in fulfilling its aims and objectives with regard to the political and social value system. Within a given organization, every individual will perform a given set of roles that are derived from the aims or objectives of the organization and its social control system. Roles are only meaningful when they stand in relationship to, and interact with, other roles. A health care organization has a culture that is unique to that organization. The health care organization will have a given set of social and political power structures, roles, and languages that support that power structure. To successfully develop and utilize a significant information system, the change leader must examine how that information will change the work that is performed by the individuals within that system and the groups to which they belong. When developing information systems, we need to identify the group values that the information system is required to support.

Leadership

A concept that has long excited and baffled the world, leadership remains as elusive today as it was in the sixteenth century, when Niccolo Machiavelli wrote The Prince. Machiavelli saw success and failure for states as stemming directly from the qualities of the leader. Antony Jay states that today “success and failure for corporations also stem directly from the qualities of their leaders. Management techniques are obviously essential, but what matters is leadership.”

Early leadership research tried to discover the traits that differentiate leaders from followers, effective from ineffective leaders, and higher echelon from lower echelon leaders. One researcher reviewed 70 years of trait research covering approximately 280 published and unpublished studies and review articles and found that only five traits consistently correlated positively with leadership: (1) intelligence, (2) dominance, (3) self-confidence, (4) high energy, and (5) task-relevant knowledge.

The debate over the value of leadership traits shifted to studies of what separates effective from ineffective leadership. These studies looked at the behaviors of leaders, the effectiveness of their leadership, and their situational leadership capabilities. The effectiveness of leadership research is based on the assumption that to be effective a specific leadership style should be used in specific situations. These theories typically treat leadership as an independent variable.

Although there is a tremendous amount of research and many suggestions on leadership success traits or characteristics, experts do not agree on the definite traits necessary for success. What constitutes success is very dependent on the individual situation.

Management of Information Issues

Information has traditionally been managed through a combination of people’s memories and a variety of paper-based systems that were limited in their capacity and convenience. The advent of information technology enabled substantial changes in work practices. The first applications were principally labor-saving ones—the replacement of tedious clerical tasks, particularly in areas of well-defined data such as finance. People later found they could use the technology for providing new services, including organizational coordinating and integrating.

Computers interact in complex ways with the organization and can significantly affect peoples’ attitudes. Other important factors are the relationship between the technology and normal information processes within an organization. In particular, there are important differences between vertical and horizontal communications. The vertical communication relates mainly to the management control and accountability processes and was typically what managers sought first. However, performance gains have been most notable when systems perform horizontal communications at the process level: linking activities together, supporting front line decision makers, and enabling the business to flow more efficiently. The types of system and the types of information handled by these two dimensions are typically very different, and inappropriate use of technology can be destructive. The role of information technology creates a very complex picture. Its impact is often immense and has an extensive impact on how organizations are operated and managed. The key is to understand where information technology will be successful and what problems might be anticipated.

Implementation

Few topics in information technology management have attracted more attention from both researchers and practitioners than implementing systems. This is not a simple, straightforward area to examine or discuss, because implementation covers such a broad variety of issues. Knowledge of implementation has evolved throughout the years, driven by less than ideal implementations experienced. Because the field of inquiry is so large, relatively few common themes have emerged.
The role of people and organizational issues was heightened when it appeared that people issues were more responsible for implementation failure than the technical abilities of the system. \(^{44–97}\) Viewed chronologically, implementation was initially identified as an organizational change project, and implementation leaders suggested that an “Unfreeze, Change and Refreeze” approach was important. \(^{48}\) Later, larger organizational issues started to be identified as important. Some of these issues included: management support, \(^{49}\) goals and objectives, \(^{100}\) commitment to change, \(^{101}\) and user acceptance, satisfaction, and attitudes. \(^{102,103}\) Keen focused on the organizational change requirements and suggested that a small project or incremental approach would facilitate communication. \(^{104}\) This has come to be known as the rapid prototype approach to change. Also, studies of packaged software found that organizational issues were as important as the capabilities of the software to complete the ultimate task requirements. \(^{105}\) Users’ perceptions of their organizational situations are also important, whether it is their perceived control \(^1\) or their perceptions of how fairly they have been treated relative to others. \(^{106}\)

Medical informatics researchers have also examined health care system implementations from the viewpoint of both the hospital or clinical systems \(^{107–111}\) and the physician. \(^{112–114}\) There is much room to delve further into implementing medical information systems, building upon the work discussed above and enhancing the methodological approaches for example, using both quantitative and qualitative methods. \(^{115}\)

Social–Technical Implications

Computers are increasingly used within the workplace to support organizational, group, and individual activities. Introducing computers into an organization can be seen either as an enabler and empowerer of groups (with a consequent positive effect on the organization) or as a prohibitor, inhibitor, and disempowerer of individuals and groups (with a corresponding negative effect on the organization). \(^{116–118}\) When the adoption and utilization of information technology is seen as an enabler and empowerer of individuals and groups, people support the system and its development. Again, the issue is ownership. If the group members perceive that they own the problem and the solution, they will cooperate with the developers to make the system work. \(^{119}\)

Social–technological issues refer to the issues surrounding the interaction of the technology with the people using the technology to perform a task. This area has been studied by Management of Information Technology researchers over the past 20 years. \(^{120,121}\) The focus is on how to facilitate the successful implementation, adoption, use, and positive outcomes of information technology in business organizations. Without successful and positive impacts of information technology at the user level, it is difficult to realize positive outcomes at the organizational level. \(^{122}\)

User involvement and participation in system development and adoption were identified early as critical factors of system success. \(^{123}\) The importance of user involvement has been supported empirically, \(^{124}\) but some of the research has not been guided by well-defined theories. \(^{125}\) This research stream has focused on the attitudes, intentions, and beliefs of technology users; these were seen as precursors of system utilization. It was assumed that higher utilization resulted in more and better performance impacts. \(^{125,126}\) A number of important research streams drawn from psychology have guided work into understandings attitudes, intentions, beliefs, and behavior. \(^{127,128}\)

In reviewing over 20 articles, Ives and Olson found that involvement plays a role in better defining user requirements, providing better understanding on how to use the system in the organization, avoiding inappropriate features, and enhancing the user’s knowledge of the system. \(^{129}\) Participation leads to increased user acceptance and use by encouraging realistic expectations, facilitating the user’s system ownership, decreasing resistance to change, and committing users to the system.

The broad topics of user involvement and participation incorporate a variety of factors that can be categorized into three groups:

- **Cognitive.** These factors are the more rational ones and are the easiest to understand and measure. They include knowledge of technology, role in system development, experience, and other task and technology characteristics.

- **Motivational.** These factors try to get at the individual’s motivational tendencies to use technology. They include interest, self-efficacy or confidence, and expectations of, beliefs about, and desire for technology.

- **Situational.** The last category refers to the factors that describe the individual’s social system or environment. These factors include social norms toward technology, facilitation conditions, role of system department, management expectations, how an individual is treated relative to others, \(^1\) and the nature of user involvement. \(^{132}\)

There have been a number of attempts to further refine and develop an integrative theory that can guide future work, \(^{133,134}\) to further refine measurement instruments, \(^{135}\) and to understand the process of user involvement. \(^{136,137}\)
User acceptance of technology research grew out of the user involvement research and some of the criticisms about its theoretical underpinnings. User acceptance research focused on perceived ease of use of the system, the perceived usefulness, and how these perceptions affect the user’s acceptance of technology. Usefulness was found to be a very important predictor of system use intentions. This research targets a sub-set of the variables in the user involvement research. Cyert and March clearly outlined how user involvement works within an organization. They explained how conflicts developed between users and developers and then what processes were used to resolve the conflicts. Medical informatics researchers have adopted these scales to measure physicians’ attitudes toward adopting technology.

User satisfaction is a set of factors that measure the users’ attitudes of satisfaction with their information technology. These factors include system-specific factors (e.g., the quality of the data) and service quality factors. User satisfaction has enjoyed widespread use both in academia and in industry, because user satisfaction is viewed as a proxy for the overall success of information systems and, more tenuously, positive organizational outcomes.

Recent attempts at frameworks to categorize past work and suggest a causal relationship between information technology and organizational success are important first steps in guiding future information systems research, but the frameworks themselves remain largely theoretical or theoretically eclectic. It is not surprising, then, that research into causal understanding of information technology success has been decidedly mixed.

Two popular User Satisfaction surveys include the instrument produced by Ives, Olson, and Baroudi and the Bailey and Pearson instrument that measures a wide variety of constructs theorized to be important. Enhancements to the work by Ives and Bailey have been used in medical informatics research.

Usability of technology refers to the human factors that impact the user’s ability to work with the system. These directly affect the usefulness and ease of use of the system. The increase in systems based on the Graphical User Interface (GUI) provides much more power and flexibility to system developers. How developers present information and control system use has tremendous impacts on the users. Usability research has a long history of studying the human–machine interfaces, but it is a relatively new field for information technology. The GUI interfaces have provided many opportunities to develop systems that are not intuitive and thus are more complex than earlier character-based system interfaces. Researchers have identified a number of fundamental issues that technologists should know. Several books are available that offer practical advice for designing user interfaces.

Implications for Informatics

Effective change requires that the proposed technical plan—plus the information about both the current situation and the historical perspective—be filtered through the black box of knowledge that is derived from the referent discipline areas of sociology, psychology, business, organizational development, etc., as shown in Figure 2. Effectively incorporating the Black Box concepts and tools can improve the design and implementation of systems, the acceptance/use of the systems, the management of the change process, and the management of the altered organization after the technological changes are implemented. This in turn leads to improved informatics outcomes and improved organizational outcomes.

Core Principles

Many past and current organizational gurus have advanced their systems for implementing the black box, and many of these systems have worked rather well for those organizations that have made a true commitment to implementation. However, management is often guilty of seeking “Band-Aid” solutions that can work magic without the painful costs of cultural change in the organization. Also, some of the gurus’ disciples are often not as competent as the gurus are in communicating the complete message of the changes required. Following a poorly communicated or poorly understood prescription often leads to
unwanted side-effects, leaving the organization at a loss as to what to do. By the time the leaders discover that the prescription is not working—and perhaps is even toxic for the organization—the organization may require heroic measures for survival.

Lorenzi, Riley, Ball, and Douglas outline a set of core principles underlying most of the various managerial systems that have been espoused in recent years. These principles, rooted in the social sciences and behavioral research, form the underlying structure or guides to more effectively managing change.

**Vision Oriented**

In any change effort, it is crucial to establish a vision that sets the direction. Without a clear direction, the organization tends to “throw resources” at the issues until its resources are depleted. In a dynamic environment, the vision may well alter over time because of the external imperatives for change and the internal responses. However, a well-defined and communicated vision gives the people in the organization a context that helps in understanding the probable directions of change and the desired outcomes.

Information technology leaders must be vision oriented, and their responsibilities include educating top management as to the opportunities that information technology offers the organization as well as the strategic issues that information technology must address. Medical informatics leaders must clearly define the vision and effectively communicate it to all the stakeholders who may directly or indirectly be affected by the vision.

**Respect for People**

A cornerstone of all interpersonal transactions is treating people with respect through honesty and trust. All the other principles, in turn, follow and enrich this basic respect. This principle includes issues such as how information is presented to the stakeholders, how people are incorporated within the change process, and so forth. Respect must be given to all the people within the organization regardless of their diverse backgrounds, educational levels or disciplines, hierarchical levels, or ethnic groups.

**Involvement**

For people to embrace change, they must be actively involved in the change process, not merely informed of it. A major involvement aspect is the aggressive seeking of inputs at the earliest possible stages of the overall process. Further, there needs to be continuous feedback on the status of the inputs and detailed explanations of why some inputs cannot be utilized or implemented; i.e., the inputs must be treated with respect.

**Empowerment**

People must be empowered if they are to move beyond involvement to the next stage, commitment. True empowerment often enables a significant flattening of the organizational structure, effectively eliminating some traditional middle management positions, which can in the short run contribute to insecurities about what empowerment really means.

**Teamwork**

Another core principle is people working together to make change happen. In today’s complex organizations, changes affect far more than just one job or one area; therefore, high degrees of teamwork are needed. The use in many organizations of self-directed work teams is an integration of the empowerment and teamwork principles. This also means actively investing in people at all organizational levels to develop their skills. The teamwork must be not only among the decision makers, but also among people at all levels of the organization.

**Customer First**

The customer must come first. This principle places the customer in a central position and requires those inside the organization to shift from their traditional internal perspective and view the organization from the external or customer point of view. According to Price Waterhouse, “Serving customers is a powerful common denominator in your organization; customers are the raison d’être of the organization. Their needs, rigorously examined, should dictate change.” Once employees feel comfortable with themselves, it is only natural to refocus the thinking of the organization to determine what its customers need.

**Openness to Change**

All prescriptive programs imply that a system and its people must be open to change. This cannot happen in a closed and highly structured bureaucratic system. In our traditional systems, stability has been viewed as the norm, with change being a temporary deviation from that norm. However, as modern chaos theory tells us, complex systems thrive only close to the edge of chaos. Our organizational cultures must accept that change is the norm and stability is the deviation. Further, this openness to change must be at the emotional or “gut” level, not just at the intellectual level.

**A Memory Aid**

This section has outlined seven core principles for effectively managing change. When planning for an information systems change, it is helpful to think of the word “victory.” However, in this case, the word is
spelled, “victore”—vision, involvement, customers, teamwork, openness, respect, and empowerment!

Strategy

While many health care organizations are facing generally similar issues and situations, no two institutions are precisely alike. The strategy that each institution develops must meet its particular needs, goals, and culture. The knowledge from the various black box disciplines must be selected and applied to maximize the probability of success in the specific organization. Using the guiding principles as a base, the change leaders must develop an effective strategy. Based on the antecedents of the people and organizational issues, the organization’s strategy must encompass five major areas: (1) clarification of organizational direction, (2) design strategy, (3) implementation strategy, (4) evaluation strategy, and (5) diffusion strategy.

Clarification of Organizational Direction

Any information technology change strategy must focus on the specific desired outcomes for the total organization. Therefore, those responsible for information technology changes must make every effort to: (1) ensure that they and the overall organization’s plans and needs are the same, and (2) educate the overall organizational leaders on both the potential opportunities and the potential threats that stem from the information technology efforts within the organization. The best practice is to seek input and information revolving around the seven guiding principles. What is the vision for the change? What are the true needs of the customers? How will all of the stakeholders be actively involved—or at least represented—in the process? How open is the organization to change?

Design Strategy

To meet its current and future information needs, the organization must be willing to invest the appropriate economic resources into the technology, the infrastructure, and its people. A seemingly elementary but quite key issue in systems design is whether the proposed system will indeed enable the organizational changes required. Does it have the necessary features, flexibility, and expandability to support the necessary organizational changes, or will it leave the people feeling that they are trying to extinguish a forest fire with a garden hose? Another key issue is the management of expectations. It is often easier for the person responsible for designing the information system to “promise the moon” either to gain support or to avoid conflict at the design stage. Similarly, it is often tempting to understate the costs to the end users in terms of training time, changes in daily routines, etc. Finally, the organization must actively involve its stakeholders in the overall system design or selection process.

Implementation Strategy

A best practice implementation plan has the traditional technically-oriented financial, training, and project management components. The plan must also include an effective change management plan that involves the relevant people and organizational issues and incorporates the seven core principles into the total process. There must be an economic investment in the people-side of technology transfer. The organization needs to continue to manage the expectations of both the organizational leaders and the end-users as an important component of success.

Evaluation Strategy

Evaluation strategy is concerned with two areas. The first is the implementation process itself and whether it succeeded. What changes could be made in future implementations? The second is the outcome of the new system: namely, will it help the organization accomplish its overall desired outcomes? All of the best practices that we know have included the collection of baseline data about the attitudes and perceptions of the end users to compare with follow-up information. These comparative data can be used to help map the organization’s diffusion strategy and can also be a baseline to view the total organizational changes.

Diffusion Strategy

The complexity of health care organizations often makes it impossible to bring up new information systems simultaneously throughout the organization. Therefore, informatics systems must be implemented or “diffused” throughout the entire organization according to a specific strategy. Those organizations with the best practices have looked beyond the current implementation (evaluation strategy) to determine the phases and strategies for the diffusion of informatics process and technologies throughout the enterprise.

Tactics

Tactics are the processes that implement strategies. How do we achieve the desired outcome? What are the appropriate tactics from the black box that will be right for the current environment? Tactics include a variety of areas. Several people and organizational issues areas are: communication and involvement processes, design processes, change management practices, project management processes, training, and evaluation processes. One goal is a better, more effective, less painful implementation process as well as greater acceptance and use of the new information
Table 2

Sample Tactics and Processes for Implementing Change

<table>
<thead>
<tr>
<th>Tactics/Processes</th>
<th>Positive Impacts</th>
</tr>
</thead>
</table>
| Communication and involvement | More involved staff  
|                    | Better understanding how the changes will impact the organization  
|                    | Better knowledge of the changes  
|                    | Better ability to cope with the changes |
| Design process     | Better systems design  
|                    | More effective work processes |
| Changes management | Less stressful organizational change  
|                    | Smoother implementation  
|                    | Better acceptance of the changes  
|                    | Better management of the altered organization |
| Project management | Better implementation of systems |
| Training           | Better use of the new system  
|                    | Better management of the altered organization |
| Evaluation         | Determination of actual vs. expected systems outcomes  
|                    | Input data for process improvements in future implementations |

Just as different organizations require different strategies, the same is true for tactics. Specific choices of tactics depend heavily upon the particular organization’s needs and culture.

Summary

Positive outcomes come to health care organizations that are doing the right things well; that is, their organizational strategies are aligned with their environments, and they are executing those strategies well. Likewise, the informatics strategies must also be aligned with the organization’s strategies. Without this congruence, informatics does not have the potential to have a substantial positive impact on the overall organizational outcomes.

The execution of the informatics strategies must also be exemplary; however, the change processes required for achieving the desired informatics and organizational outcome goals are demanding and complicated. Implementing them in extremely complex organizations that operate on a 7-day by 24-hour basis is not easy. However, we are constantly learning more about complex change processes and the ways we can better manage them to improve our needed informatics outcomes. The challenge is to build upon the existing research base to move us even further ahead.

References


79. DiGiorgio CJ, Richart CA, Klatt E, Becich MJ. E-mail, the Internet, and information access technology in pathology. Seminars in Diagnostic Pathology. 1994;11:294–304.


Association Between Gout and All-Cause as well as Cardiovascular Mortality: A Systematic Review

Kathrin Lottmann · Xiaoyu Chen · Peter K. Schädlich

Published online: 18 February 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract  Gout affects 1% to 2% of the population, and the prevalence is increasing due to changes in diet and the ageing of the population. Its development and risk factors have been explored frequently, and recommendations for the diagnosis and management of gout implemented. Nevertheless, there is a lack of knowledge regarding the long-term impact on gouty patients. This systematic review therefore evaluates the association between gout and all-cause as well as cardiovascular mortality. A systematic literature search was performed, and seven long-term studies were ultimately analyzed. Six of them used multivariate regressions to assess the adjusted mortality ratio in gouty patients with reference to patients without the disorder. Despite differences in study designs, study populations, and definitions of gout, the results were consistent: There was an independent association between gout and all-cause as well as cardiovascular mortality. Knowing that patients with gout are at risk emphasizes the need for adequate care.

Keywords  Gout · Mortality · All-cause mortality · Cardiovascular mortality · Systematic review · Association · Crystal arthritis

Introduction

Gout is a common disease with increasing prevalence. Its manifestation, progression, adequate therapeutic interventions, and related comorbidities are therefore frequently discussed in the literature. Less attention is paid to the question of whether gout has an impact on mortality. However, this long-term outcome is relevant not only for the patient. New light should also be shed on the significance of this disease and consequently on the treatment and care of gouty patients, as these often remain suboptimal [1].

Gout is a crystal deposition disease in which renal elimination of uric acid is insufficient. The increased uric acid level can lead to the formation of monosodium urate crystals in synovial fluid or soft tissues. This results in painful inflammation of the joints, which is why the disorder is also known as inflammatory arthritis [1, 2]. The typical initial presentation of gout is podagra [3]. Gout is characterized by recurrent acute episodes and can become a chronic condition [3–5].

The prevalence of gout has increased in recent decades. This might be the result of a changing diet and an ageing population [2, 3]. Despite methodologic challenges, the international prevalence is estimated at up to 1% to 2% [1, 6, 7]. Men are more likely to be affected than women, and the prevalence increases with age [1, 8].

In 2006, the European League Against Rheumatism (EULAR) developed evidence-based recommendations for the diagnosis and management of gout [1, 9], according to which the disease may be diagnosed by monosodium urate crystals in synovial fluid, podagra, or tophus, while a raised serum urate level alone is not specific to gout [1]. Risk factors for the development of gout are diet [2, 5]; genetic predisposition [2, 4]; hyperuricemia [4, 10]; and comorbidities such as diabetes, hypertension, obesity, heart failure, and renal insufficiency [2, 6, 11].

The management of gout aims at lowering and maintaining serum urate levels below the saturation point (6.8 mg/dL, or 408 μmol/L). This helps to dissolve existing monosodium
urate crystals and to prevent further crystals from forming [1, 9]. In addition, special diets, weight reduction, and reduced alcohol consumption are among the nonpharmacologic interventions [4].

The objective of this review is to scrutinize in a systematic manner whether there is an association between gout and all-cause or cardiovascular mortality. Despite neither the management of gout nor economic aspects being considered, this systematic review is intended to impact on recurrent discussion about the management of gout. Furthermore, knowledge about the association of gout and mortality may underline the need for adequate care.

Methods

Data Sources

We conducted a systematic literature search of Medline and EMBASE up to April 2011 using a search strategy combining the Medical Subject Headings (MeSH terms) and keywords in the titles/abstracts. The search string consisted of keywords referring to the medical indication gout linked by the Boolean operator “AND” to terms associated with the outcomes of all-cause mortality or cardiovascular mortality. Publications in English, French, and German were included, whereas studies on animals were excluded. In addition, we performed a manual search of references.

Study Selection

Studies were included in this review if they were noninterventional trials investigating the association between gout and all-cause or cardiovascular mortality in patients with gout compared with the population without this disease. Gout could be defined by the respective ICD codes (ICD-9 code 274.x or ICD-10 code M10.x [12, 13]) or diagnosed by a physician in accordance with the evidence-based EULAR recommendations [1]. The patient-relevant outcomes considered were all-cause mortality and cardiovascular mortality (ICD-9 code 390–459, ICD-10 code I00–I99) [12, 13]. Studies analyzing the association between a single cardiovascular disease (CVD) (eg, the number of fatal myocardial infarctions) and gout were therefore excluded because this systematic review focuses on cardiovascular mortality as a whole. In addition, studies of the best available and most feasible evidence were to be included. With regard to the question posed by this review, the best evidence would be a retrospective cohort study with a 2b level of evidence [14].

Validity Assessment and Data Abstraction

Two investigators (Lottmann and Chen) independently scrutinized all identified studies after excluding duplicates. As a first step, the titles and abstracts were investigated with regard to the predefined inclusion criteria, and the full texts of all potentially relevant studies were evaluated next. In case of disagreement, consensus was reached by discussion. The selection process, including the reasons for exclusion, was documented. Due to the lack of a validated quality assessment tool for nonrandomized interventions [15], we developed our own checklist [16]. Our quality checklist was based on the methodologic requirements of the German Agency for Health Technology Assessment [17] and on the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines [18]. Finally, the quality of each study included was assessed and scored. The data from all studies scored with at least “fair” methodologic quality were subsequently collected by one reviewer using a self-developed data abstraction list.

Results

Study Identification and Selection

The process of the systematic literature search is depicted in Fig. 1. The number of articles identified, screened, and excluded, as well as the number of full-texts retrieved and finally included for data abstraction is delineated. We identified seven articles as conforming to our inclusion criteria. All of them analyzed the association between gout and mortality (seven studies) as well as cardiovascular mortality (four studies) as either a primary or secondary outcome.

The study types varied from retrospective to prospective. They consisted of cohort studies, surveys, and data analyses. All these studies, except one [19], analyzed the association between gout and all-cause mortality (seven studies) as well as cardiovascular mortality (four studies) as either a primary or secondary outcome.

The study types varied from retrospective to prospective. They consisted of cohort studies, surveys, and data analyses. All these studies, except one [19], analyzed the association between gout and mortality using multivariate regression (ie, they calculated hazard ratios [HRs] or risk ratios [RRs]). In addition, several studies determined the mortality rate in patients with gout compared with people without this disease. In this review, we focus on the results of the regression analyses, as this is the more appropriate way of obtaining conclusive results.

Patient Characteristics

Patient populations were very heterogeneous in the different studies, which made comparison challenging. Whereas two studies included renal transplant populations [20, 21], another two focused on individuals with coronary heart
disease (CHD) or those who were at risk of this disorder [22**, 23]. One study analyzed data from US veterans [19], while another included only data from health professionals [24]. An additional study evaluating participants in a health-screening program in Taiwan was included. Even if an almost equal sex ratio in the study population was provided, gout was far more common in men than in women (90.4% of the patients with gout were male) [25**]. Furthermore, several studies were conducted in an older population. Despite this heterogeneity, each study population was characteristic of gouty patients. Study populations showed typical comorbidities, age groups, and male dominance in patients with gout.

The definition of gouty patients varied from study to study, as did the prevalence of gout in the study population (range, 2.1%–10.6%). The 1-year mortality rate from all causes also varied (range, 2.6%–14.3%).

Gout and All-Cause Mortality

Abbott et al. [20] used Medicare claims data of primary renal transplant patients reported in the United States Renal Data System (USRDS). The aim of their retrospective cohort study was to evaluate the implications of new-onset gout on survival. They determined an increased risk of mortality (adjusted HR [AHR], 1.26 [95% CI, 1.08–1.47]) in patients developing gout after kidney transplantation compared with renal transplant patients without the disorder. Gout diagnosed prior to transplant was not significantly associated with an increased risk of mortality (AHR, 1.18 [95% CI, 0.98–1.43]). Their study indicates an independent association between new-onset gout after renal transplantation and all-cause mortality.

One more study evaluated the association between incident gout in renal transplant patients and mortality using the USRDS. This study population was treated with dialysis after the transplantation, and gout was associated with an increased risk of mortality (AHR, 1.49 [95% CI, 1.43–1.55]; P<0.001) [21]. The result not only confirmed the independent association with new-onset gout after renal transplantation but even determined an elevated mortality risk in kidney transplant patients treated with dialysis [21] compared with those without dialysis [20].
The association between gout and all-cause mortality in patients with cardiovascular comorbidities was investigated in a time-matched, nested case-control analysis of a retrospective cohort study based on the Quebec Universal Health Insurance database. The study population (age >66 years) had been discharged from the hospital with a primary diagnosis of heart failure. Death was considered to be associated with gout if the patient suffered an acute episode of gout within 60 days before death. Regarding the gout population, a distinction was made between three definitions of gout. The risk of all-cause mortality in patients with any acute gout was significantly increased, even if adjusted for several CVDs and medications (adjusted RR [ARR], 1.76 [95% CI, 1.08–2.86]; P = 0.02). No significant association between acute gout and all-cause mortality was found in patients admitted to the hospital for gout. However, in patients visiting an emergency department for a gout flare, gout was again significantly associated with an increased risk of death from all causes (ARR, 2.10 [95% CI, 1.19–3.70]; P = 0.01) [22••].

Whether there is an association between gout and all-cause mortality in men at above average risk of CHD but without evidence of clinical CVD during the intervention phase was evaluated in the long-term follow-up of the MRFIT (Multiple Risk Factor Intervention Trial) study over 17 years. Again, different definitions of gout were used: self-reported gout and documented hyperuricemia, self-reported gout alone, and prescribed gout medication. Gouty patients based on self-reported gout and documented hyperuricemia had an increased mortality risk (AHR, 1.22; P = 0.009) compared with those patients with neither gout nor hyperuricemia. The mortality risk in gouty patients, defined by use of gout-related medicine (allopurinol, probenecid, or colchicine), was comparable (AHR, 1.23; P < 0.001) [23]. The association between gout and all-cause mortality was revealed once again, although hyperuricemia might have biased the result in one of the analyses.

A further study focused on patients with gout and cardiovascular comorbidities [24]. In the prospective Health Professionals Follow-Up Study, which was limited to male participants, data were acquired via mailed questionnaires. The authors calculated the adjusted mortality risk in patients with gout and with CHD, as well as in gouty patients without CHD comorbidity (in each case compared with patients without gout but with the corresponding CHD status). They furthermore evaluated different definitions of gout: any gout, history of gout, and newly diagnosed gout. Patients with a history of gout without CHD had a slightly higher mortality risk (ARR, 1.28 [95% CI, 1.15–1.41]) than those with a history of gout and CHD comorbidity (ARR, 1.25 [95% CI, 1.09–1.45]). Moreover, an increased mortality risk was determined in patients with newly diagnosed gout (ARR, 1.28 [95% CI, 1.13–1.46]); CHD comorbidities were not reported. In the analysis including any gouty patients, the mortality risk was slightly lower in the group without CHD (ARR, 1.25 [95% CI, 1.13–1.38]) than in the group with CHD (ARR, 1.35 [95% CI, 1.21–1.50]) [24]. This study thus reported an association between gout and all-cause mortality independent of the CHD status and not affected by the gout duration.

While all the studies mentioned above included mainly Caucasians as well as a few black patients, the association between gout and all-cause mortality was also confirmed in an Asian population. Data from a health-screening program in Taiwan were investigated. This analysis revealed an increased risk of death from all causes in patients with gout (AHR, 1.46 [95% CI, 1.12–1.91]; P = 0.005) [25••]. The mortality risk was even higher when a more verifiable definition of gout was used and the analysis included only patients with monosodium urate crystals in the synovial fluid or ICD code 274.x, and excluded those with self-reported gout (AHR, 1.51 [95% CI, 1.14–1.99]; P = 0.004).

One study did not calculate an RR to evaluate the association between gout and all-cause mortality but accounted for a multivariable-adjusted difference in the mortality rate. Based on a mailed survey to US veterans (99% men), similar mortality rates were determined in individuals with and without gout (no gout, 2.22% [95% CI, 1.56%–3.15%]; gout, 2.62% [95% CI, 1.60%–4.28%]; difference, 18.47%; P = 0.230) [19]. Gout therefore did not have an impact on the mortality rate in this study. However, conclusions regarding the association between gout and all-cause mortality should be drawn carefully, as the statistical methods used differ from those in the other studies described previously. Detailed figures are displayed in Table 1.

Gout and Cardiovascular Mortality

Besides the described independent association between gout and all-cause mortality, four of the seven identified studies evaluated the impact of gout on cardiovascular mortality as well. All of them revealed that gout is also associated with an increased risk of cardiovascular mortality; the results are outlined in detail in Table 1. The cardiovascular mortality risk in gouty renal transplant patients treated with dialysis was increased compared with those patients without gout (AHR, 1.47 [95% CI, 1.26–1.59]). However, this mortality ratio was slightly lower than the all-cause mortality risk in the same group [21].

Even in the long-term follow-up study over 17 years, gout was associated with cardiovascular mortality, but the risk changed when data were limited to special definitions of gout. Based on self-reported gout and diagnosed
<table>
<thead>
<tr>
<th>Study (year)</th>
<th>Duration, y</th>
<th>Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbott et al. [20] (2005)</td>
<td>≤3</td>
<td>28,924 patients with renal transplantation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1,583 new-onset gout after transplantation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1,175 new-onset gout prior to transplantation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AHR or ARR (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.26 (1.08–1.47)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.18 (0.98–1.43)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cardiovascular mortality:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AHR or ARR (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 2,773 with gout</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CHD, 1.35 (1.21–1.50)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-reported gout and hyperuricemia, 1.22; P=0.009</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use of gout-related medicine, 1.23; P&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-reported gout and hyperuricemia, 1.30 (1.04–1.61); P=0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use of gout-related medicine, 1.18; P=0.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 24,215 with gout (39.9% men)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AHR or ARR (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.49 (1.43–1.55); P&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.49 (1.43–1.55); P&lt;0.001</td>
</tr>
<tr>
<td>Krishnan et al. [23] (2008)</td>
<td>17</td>
<td>9,105 men at risk of CHD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 655 self-reported gout and hyperuricemia</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use of gout-related medicine, 1.23; P&lt;0.001</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Self-reported gout and hyperuricemia, 1.30 (1.04–1.61); P=0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Use of gout-related medicine, 1.18; P=0.08</td>
</tr>
<tr>
<td>Kuo et al. [25**] (2010)</td>
<td>≤8</td>
<td>61,527 participants in a Taiwanese health-screening program</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1,311 with gout (90.4% men)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AHR or ARR (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.46 (1.12–1.91); P=0.005</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.97 (1.08–3.59); P=0.02</td>
</tr>
<tr>
<td>Thanassoulis et al. [22++] (2010)</td>
<td>≤8</td>
<td>25,090 patients aged ≥66 y discharged from hospital with primary diagnosis of heart failure</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1,053 with gout</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AHR or ARR (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.76 (1.08–2.86); P&lt;0.02</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>Singh and Strand [19] (2008)</td>
<td>1</td>
<td>64,553 US veterans</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 1,581 with gout (99% men)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>AHR or ARR (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>NA</td>
</tr>
</tbody>
</table>

AHR, adjusted hazard ratio; ARR, adjusted risk ratio; CHD, coronary heart disease; CI, confidence interval; NA, not applicable
hyperuricemia, the cardiovascular mortality risk was only significantly higher when the reference group consisted of patients with neither gout nor hyperuricemia (AHR, 1.30; \( P=0.02 \)). The risk was not significantly increased when compared with patients without gout but potentially with hyperuricemia. Two sensitivity analyses with gouty patients defined as 1) using gout medication and 2) having self-reported gout revealed a nonsignificantly increased risk of cardiovascular mortality [23].

The independent association between gout and cardiovascular mortality was also reported in the Health Professionals Follow-Up Study. In this case, the cardiovascular mortality risk in patients with gout was even slightly higher than the all-cause mortality risk for the same group. Again, the analysis of risk was carried out for different gout groups, depending on the duration of illness (history of gout, newly diagnosed gout, and any gout) and on the CHD status (history of gout [no CHD]: ARR, 1.38 [95% CI, 1.15–1.66]; history of gout [CHD]: ARR, 1.26 [95% CI, 1.07–1.50]; newly diagnosed gout [CHD not reported]: ARR, 1.31 [95% CI, 1.08–1.59]; any gout [no CHD]: ARR, 1.32 [95% CI, 1.09–1.60]; any gout [CHD]: ARR, 1.35 [95% CI, 1.19–1.55]) [24].

Compared with the previously reported study results, the cardiovascular mortality risk determined in the study with participants in the Taiwanese health-screening program was higher. The main analysis including gouty patients corresponding to at least one of the three applied gout definitions of this study revealed an almost twofold higher cardiovascular mortality risk in patients with gout compared with those without the disease (AHR, 1.97 [95% CI, 1.08–3.59]; \( P=0.027 \)). When data were limited to patients with monosodium urate crystals in the synovial fluid or corresponding to ICD code 274.x, and excluded those with self-reported gout, the risk was somewhat lower (AHR, 1.86 [95% CI, 1.01–3.44]; \( P=0.048 \)) [25••].

All-Cause and Cardiovascular Mortality Risks
in Gouty Subgroups

Different subgroup analyses revealed an independent association between gout and all-cause mortality as well as cardiovascular mortality not only in the respective study populations as a whole, but also in several subgroups. Results for single subgroups are presented accounting for the impact of comorbidities on the mortality risk.

**Sex**

Male gouty patients after renal transplantation who were treated with dialysis had a lower risk of all-cause mortality than women with the same characteristics (AHR, 1.59 [95% CI, 1.47–1.71] vs AHR, 1.64 [95% CI, 1.54–1.75]) [21].

**Age**

The Health Professionals Follow-Up Study revealed an increased all-cause mortality risk in the population 60 to 69 years of age compared with patients without gout from the same age group. This mortality risk decreased as patients grew older. In the youngest age group (<60 years), the increased mortality risk was not significant (<60 years: AHR, 1.30 [95% CI, 0.96–1.74]; 60–69 years: AHR, 1.35 [95% CI, 1.13–1.61]; ≥70 years: AHR, 1.16 [95% CI, 1.01–1.32]). The cardiovascular mortality risk in patients with gout in the 60- to 69-year subgroup was more than two times higher than in patients without gout from the same age group; the risks in the other age groups were not significantly increased (<60 years: AHR, 1.84 [95% CI, 0.93–3.61]; 60–69 years: AHR, 2.10 [95% CI, 1.38–3.18]; ≥70 years: AHR, 0.98 [95% CI, 0.68–1.41]) [24]. Thus, gouty patients 60 to 69 years of age seem to be at elevated all-cause and cardiovascular mortality risk.

**Race**

The AHR for all-cause mortality was higher in black patients with gout than in other races (AHR, 1.69 [95% CI, 1.59–1.85] vs AHR, 1.59 [95% CI, 1.50–1.69]) [21].

**Gout Duration**

Newly diagnosed gout seems to have a greater impact on all-cause mortality than established gout (<60 years: AHR, 1.30; ≥60 years: AHR, 1.37; 11–15 years: AHR, 1.30; 16–20 years: AHR, 1.37; 21–30 years: AHR, 1.37; 31–40 years: AHR, 1.37; 41–50 years: AHR, 1.37; 51–60 years: AHR, 1.37; >60 years: AHR, 1.37) [24].

**Hypertension**

The mortality risk in gouty patients in the Health Professionals Follow-Up Study both with and without hypertension was almost the same (ARR, 1.24 [95% CI, 1.09–1.41] vs ARR, 1.25 [95% CI, 1.06–1.47]), while subgroup analysis for cardiovascular mortality did not reveal a significant ARR in patients with gout [24]. By contrast, dialysis-treated gouty renal transplant patients with hypertension had a higher mortality risk than those without hypertension (AHR, 1.74 [95% CI, 1.55–1.94] vs AHR, 1.59 [95% CI, 1.51–1.68]) [21].

**Hypercholesterolemia**

Men with gout and hypercholesterolemia have a lower all-cause mortality risk than those without this comorbidity (with hypercholesterolemia: ARR, 1.15 [95% CI, 0.96–1.37] vs without hypercholesterolemia: ARR, 1.30 [95% CI, 1.15–1.47]) [20]. The influence of hypercholesterolemia did
not change with reference to cardiovascular mortality—gouty patients without hypercholesterolemia seem to have a higher mortality risk compared with those with hypercholesterolemia (with hypercholesterolemia: ARR, 1.31 [95% CI, 0.93–1.83] vs without hypercholesterolemia: ARR, 1.55 [95% CI, 1.05–2.30]). Not all results were significant [24].

Diabetes

Diabetes in patients with gout appears to decrease all-cause mortality slightly, as the adjusted risk was lower than in gouty patients without diabetes (AHR, 1.56 [95% CI, 1.44–1.69] vs AHR, 1.68 [95% CI, 1.58–1.78]) [21].

Discussion

Despite the differences among populations in the studies included, all multivariate regressions yielded the same result: There was an independent association between gout and all-cause mortality as well as cardiovascular mortality. Conclusions regarding the causes of this association cannot be drawn, as further research would be needed. This also applies to the question of whether there is an association between cardiovascular mortality and all-cause mortality. However, gout elevates the risk of mortality and thus makes gouty patients an at-risk population.

Patients in whom gout was newly diagnosed after a renal transplantation had an increased all-cause mortality risk, whereas this association was not significant in renal transplant patients with a history of gout [20]. Another study confirmed the impact of gout on all-cause as well as cardiovascular mortality in renal transplant patients treated with dialysis [21]. Furthermore, an independent association between gout and all-cause mortality as well as cardiovascular mortality was also found in patients with any gout or a history of this disorder, irrespective of CHD comorbidities [24]. Another study confirmed the association between gout and all-cause mortality in patients with previous heart failure and acute gout within 60 days before their death [20]. Gouty patients with hyperuricemia also had an increased all-cause and cardiovascular mortality risk [23]. Finally, a study of an Asian population without specific comorbidities obtained the same result (ie, it indicated that gout is associated with all-cause as well as cardiovascular mortality) [25••].

The number of studies evaluating the association between gout and all-cause or cardiovascular mortality is limited. More research has been conducted on the questions regarding whether gout is associated with death from CHD alone or CVD in general, and the impact of gout on cardiovascular events [26–29]. Even these studies indicate corresponding results: Gout is a risk factor for CVD and for death caused by it.

The large number of study participants in each of the studies included strengthens the validity of the results, although the prevalence of gout in the study population was somewhat higher than the general estimated prevalence [2, 6, 7]. Despite the heterogeneous study populations and differences in gout definition, all studies indicated an independent association between gout and all-cause as well as cardiovascular mortality (except for the study that calculated differences in the mortality rates instead of ARR/AHR) [19]). That all multivariate regressions yielded consistent outcomes may be considered another validation of gout being a risk factor for mortality. Moreover, the specific characteristics of the study population reflected the typical characteristics of gouty patients: mainly male, older age, and comorbidities such as CVD or renal failure.

In spite of the consistent outcomes, the comparability of the studies is complicated by use of different definitions of gout. Wijands et al. [30] described the challenge of defining gout in epidemiologic studies. It is difficult to apply diagnostic guidelines (eg, monosodium urate crystals), which are developed for use on an individual level, to large populations. Drawing on ICD codes might offer a solution but implies several methodologic uncertainties [30]. As a consequence, different definitions of a gout population might entail variant outcomes. This is in line with the studies analyzed in this review: The mortality risk varied depending on the gout definition [22••, 23, 25••]. To resolve this challenge, several sensitivity analyses were conducted.

The impact of patient characteristics and comorbidities on the association between gout and mortality was determined in different subgroup analyses. Regarding race, subgroup analysis established a higher mortality risk in blacks [21], which goes hand-in-hand with an increased risk of gout among this ethnic group. Contrary to the elevated risk of gout in men described in the literature, one study indicated a reverse effect, as the risk of all-cause mortality was higher in women than men [21]. It could be concluded that gout is less common, but more serious, in women. Furthermore, results of a subgroup analysis with diabetes, one of the most common comorbidities in gout, were unexpected. Gouty patients with diabetes had a slightly lower mortality risk than those without this comorbidity [21]. One explanation might be that monitoring is better in patients with diabetes, although further research is needed to verify this claim. These results should be interpreted with caution, as they originate from a study of patients with severe comorbidities (renal transplant patients treated with dialysis).

The effect of gout duration on the mortality risk was analyzed in only two studies, both of which reported an
increased mortality risk in patients with new-onset gout and a slightly lower risk over time [20, 24]. This indicates the need for early intervention and treatment of gout.

Finally, it should be noted that four of the seven studies included were explicitly not sponsored by pharmaceutical companies [19, 21, 22••, 25••], two were partially supported by the industry [23, 24], and information on sponsorship or conflict of interest was lacking in one study [20].

Conclusions

Gout is independently associated with all-cause mortality and cardiovascular mortality. This was the message of six studies evaluating the association by multivariate regression. As described, study populations were heterogeneous, but the study outcomes were nevertheless consistent. The elevated mortality risk in patients with gout applies to gouty patients with typical characteristics (black, male, older age) and classical comorbidities (eg, diabetes, hypertension, hyperuricemia). Subgroup analyses suggest that patients with new-onset gout and female sex have a slightly elevated mortality risk. In addition to these two subgroups, gouty patients without comorbidities that require regular monitoring, such as diabetes, might be considered a vulnerable group. Not only these subgroups, but the generally increased mortality risk associated with gout should be borne in mind when treating this disorder. A need exists for adequate treatment of gout to fulfill the needs of these at-risk patients.

Disclosure This project has been supported by an unconditional grant from Berlin-Chemie AG (Berlin, Germany). Mrs. Lottmann, Mrs. Chen, and Dr. Schädlich are employees of IGES Institut GmbH and have worked on projects funded by Berlin-Chemie AG.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

Papers of particular interest, published recently, have been highlighted as:

• Of major importance


17. DAHTA®DIMDI.de: [Methodological requirements for preparing an HTA-report]. Cologne: German Institute for Medical Documentation and Information, n.a.


25. Kuo CF, See LC, Luo SF, et al. Gout: an independent risk factor for all-cause and cardiovascular mortality. Rheumatology. 2010; 49:141–6. This paper was identified in the systematic literature search and used multivariate regression to analyze the association between gout and all-cause as well as cardiovascular mortality.
Abstract  An unresolved issue in the field of implementation research is how to conceptualize and evaluate successful implementation. This paper advances the concept of “implementation outcomes” distinct from service system and clinical treatment outcomes. This paper proposes a heuristic, working “taxonomy” of eight conceptually distinct implementation outcomes—acceptability, adoption, appropriateness, feasibility, fidelity, implementation cost, penetration, and sustainability—along with their nominal definitions. We propose a two-pronged agenda for research on implementation outcomes. Conceptualizing and measuring implementation outcomes will advance understanding of implementation processes, enhance efficiency in implementation research, and pave the way for studies of the comparative effectiveness of implementation strategies.

Keywords  Implementation · Outcomes · Evaluation · Research methods

Background

A critical yet unresolved issue in the field of implementation science is how to conceptualize and evaluate success. Studies of implementation use widely varying approaches to measure how well a new mental health treatment, program, or service is implemented. Some infer implementation success by measuring clinical outcomes at the client or patient level while other studies measure the actual targets of the implementation, quantifying for example the desired provider behaviors associated with delivering the newly implemented treatment. While some studies of implementation strategies assess outcomes in terms of improvement in process of care, Grimshaw et al. (2006) report that meta-analyses of their effectiveness has been thwarted by lack of detailed information about outcomes, use of widely varying constructs, reliance on dichotomous rather than continuous measures, and unit of analysis errors.

This paper advances the concept of “implementation outcomes” distinct from service system outcomes and clinical treatment outcomes (Proctor et al. 2009; Fixsen et al. 2005; Glasgow 2007a). We define implementation outcomes as the effects of deliberate and purposive actions to implement new treatments, practices, and services. Implementation outcomes have three important functions. First, they serve as indicators of the implementation success. Second, they are proximal indicators of implementation processes. And third, they are key intermediate outcomes (Rosen and Proctor 1981) in relation to service system or clinical outcomes in treatment effectiveness and quality of care research. Because an intervention or
treatment will not be effective if it is not implemented well, implementation outcomes serve as necessary preconditions for attaining subsequent desired changes in clinical or service outcomes.

Distinguishing implementation effectiveness from treatment effectiveness is critical for transporting interventions from laboratory settings to community health and mental health venues. When such efforts fail, as they often do, it is important to know if the failure occurred because the intervention was ineffective in the new setting (intervention failure), or if a good intervention was deployed incorrectly (implementation failure). Our current knowledge of implementation is thwarted by lack of theoretical understanding of the processes involved (Michie et al. 2009). Conceptualizing and measuring implementation outcomes will advance understanding of implementation processes, enable studies of the comparative effectiveness of implementation strategies, and enhance efficiency in implementation research.

This paper aims to advance the “vocabulary” of implementation science around implementation outcomes through four specific objectives: (1) to advance conceptualization of implementation outcomes by distinguishing implementation outcomes from service and clinical outcomes; (2) to advance clarity of terminology currently used in implementation science by nominating heuristic definitions of implementation outcomes, yielding a working “taxonomy” of implementation outcomes; (3) to reflect the field’s current language, conceptual definitions, and approaches to operationalizing implementation outcomes; and (4) to propose directions for further research to advance knowledge on these key constructs and their interrelationships.

Our objective of advancing a taxonomy of implementation outcomes is comparable to the work of Michie et al. (2005, 2009), Grimshaw et al. (2006), the Cochrane group, and others who are working to develop taxonomies and common nomenclature for implementation strategies. Our work is complementary to these efforts because implementation outcomes will provide researchers with a framework for evaluating implementation strategies.

Conceptual Framework for Implementation Outcomes

Our understanding of implementation outcomes is lodged within a previously published conceptual framework (Proctor et al. 2009) as shown in Fig. 1. The framework distinguishes between three distinct but interrelated types of outcomes—implementation, service, and client outcomes. Improvements in consumer well-being provide the most important criteria for evaluating both treatment and implementation strategies—for treatment research, improvements are examined at the individual client level whereas improvements at the population-level (within the providing system) are examined in implementation research. However, as we argued above, implementation research requires outcomes that are conceptually and empirically distinct from those of service and clinical effectiveness.

For heuristic purposes, our model positions implementation outcomes as preceding both service outcomes and client outcomes, with the latter sets of outcomes being impacted by the implementation outcomes. As we discuss later in this paper, interrelationships among these outcomes require conceptual mapping and empirical tests. For example, one would expect to see a treatment’s strongest impact on client outcomes as an empirically supported treatment’s (EST) penetration increases in a service setting—but this hypothesis requires testing. Our model derives service outcomes from the six quality improvement aims set out in the reports on crossing the quality chasm: the extent to which services are safe, effective, patient-centered, timely, efficient, and equitable (Institute of Medicine Committee on Crossing the Quality Chasm 2006; Institute of Medicine Committee on Quality of Health Care in America 2001).

Methods

The paper’s methods were shaped around its overall aim: to advance clarity in the language used to describe outcomes of implementation. We convened a working group of implementation researchers to identify concepts for labeling and assessing outcomes of implementation processes. One member of the group was a doctoral student RA who coordinated, conducted, and reported on the literature search and constructed tables reflecting various iterations of the heuristic taxonomy. The RA conducted literature searches using key words and search programs to identify literature on the current state of conceptualization and measurement of these outcomes, primarily in the health and behavioral sciences. We searched in a number of databases with a particular focus on MEDLINE, CINAHL Plus, and PsycINFO. Key search terms included the name
of the implementation outcome (e.g., “acceptability,” “sustainability,” etc.) along with relevant synonyms combined with any of the following: innovation, EBP, evidence based practice, and EST. We scanned the titles and abstracts of the identified sources and read the methods and background sections of the studies that measured or attempted to measure implementation outcomes. We also included information from relevant conceptual articles in the development of nominal definitions. Whereas our primary focus was on the implementation of evidence based practices in the health and behavioral sciences, the keyword “innovation” broadened this scope by also identifying studies that focused on other areas such as physical health that may inform implementation of mental health treatments. Because terminology in this field currently reflects widespread inconsistency, we followed leads beyond what our keyword searches “hit” upon. Thus we read additional articles that we found cited by authors whose work we found through our electronic searches. We also conducted searches of CRISP, TAGG, and NIH reporter and studies to identify funded mental health research studies with “implementation” in their titles or abstracts, to identify examples of outcomes pursued in current research.

We used a narrative review approach (Educational Research Review), which is appropriate for summarizing different primary studies and drawing conclusions and interpretation about “what we know,” informed by reviewers’ experiences and existing theories (McPheeters et al. 2006; Kirkevold 1997). Narrative reviews yield qualitative results, with strengths in capturing diversities and pluralities of understanding (Jones 1997). According to McPheeters et al. (2006), narrative reviews are best conducted by a team. Members of the working group read and reviewed conceptual and theoretical pieces as well as published reports of implementation research. As a team, we convened recurring meetings to discuss the similarities and dissimilarities. We audio-taped and transcribed meeting discussions, and a designated individual took thorough notes. Transcriptions and notes were posted on a shared computer file for member review, revision, and correction.

Group processes included iterative discussion, checking additional literature for clarification, and subsequent discussion. The aim was to collect and portray, from extant literature, the similarities and differences across investigators’ use of various implementation outcomes and definitions for those outcomes. Discussions often led us to preserve distinctions between terms by maintaining in our “nominated” taxonomy two different implementation outcomes because the literature or our own research revealed possible conceptual distinctions. We assembled the identified constructs in the proposed heuristic taxonomy to portray the current state of vocabulary and conceptualization of terms used to assess implementation outcomes.

Taxonomy of Implementation Outcomes

Through our process of iterative reading and discussion of the literature, we worked to nominate definitions that (1) achieve as much consistency as possible with any existing definitions (including multiple definitions we found for a single construct), yet (2) serve to sharpen distinctions between constructs that might be similar. For several of the outcomes, the literature did not offer one clear nominal definition.

Table 1 depicts the resultant working taxonomy of implementation outcomes. For each implementation outcome, the table nominates a level of analysis, identifies the theoretical basis to the construct from implementation literature, shows different terms that are used for the construct in the literature, suggests the point or stage within implementation processes at which the outcome may be most salient, and lists the types of existing measures for the construct that our search identified. The implementation outcomes listed in Table 1 are probably only the “more obvious,” and we expect that other concepts may emerge from further analysis of the literature and from the kind of empirical work we call for in our discussion below. Many of the implementation outcomes can be inferred or measured in terms of expressed attitudes and opinions, intentions, or reported or observed behaviors. We now list and discuss our nominated conceptual definitions for each implementation outcome in our proposed taxonomy. We reference similar definitions from the literature, and also comment on marked differences between our definitions and others proposed for the term.

Acceptability is the perception among implementation stakeholders that a given treatment, service, practice, or innovation is agreeable, palatable, or satisfactory. Lack of acceptability has long been noted as a challenge in implementation (Davis 1993). The referent of the implementation outcome “acceptability” (or the “what” is acceptable) may be a specific intervention, practice, technology, or service within a particular setting of care. Acceptability should be assessed based on the stakeholder’s knowledge of or direct experience with various dimensions of the treatment to be implemented, such as its content, complexity, or comfort. Acceptability is different from the larger construct of service satisfaction, as typically measured through consumer surveys. Acceptability is more specific, referencing a particular treatment or set of treatments, while satisfaction typically references the general service experience, including such features as waiting times, scheduling, and office environment. Acceptability may be measured from the perspective of
<table>
<thead>
<tr>
<th>Implementation outcome</th>
<th>Level of analysis</th>
<th>Theoretical basis</th>
<th>Other terms in literature</th>
<th>Salience by implementation stage</th>
<th>Available measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acceptability</td>
<td>Individual provider</td>
<td>Rogers: “complexity” and to a certain extent “relative advantage”</td>
<td>Satisfaction with various aspects of the innovation (e.g. content, complexity, comfort, delivery, and credibility)</td>
<td>Early for adoption</td>
<td>Survey</td>
</tr>
<tr>
<td></td>
<td>Individual consumer</td>
<td></td>
<td></td>
<td>Ongoing for penetration</td>
<td>Qualitative or semi-structured interviews</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Late for sustainability</td>
<td>Administrative data</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Refused/blank</td>
<td></td>
</tr>
<tr>
<td>Adoption</td>
<td>Individual provider</td>
<td>RE-AIM: “adoption” Rogers: “trialability” (particularly for early adopters)</td>
<td>Uptake; utilization; initial implementation; intention to try</td>
<td>Early to mid</td>
<td>Administrative data</td>
</tr>
<tr>
<td></td>
<td>Organization or setting</td>
<td></td>
<td></td>
<td>Observation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Qualitative or semi-structured interviews</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Survey</td>
<td></td>
</tr>
<tr>
<td>Appropriateness</td>
<td>Individual provider</td>
<td>Rogers: “compatibility”</td>
<td>Perceived fit; relevance; compatibility; suitability; usefulness; practicability</td>
<td>Early (prior to adoption)</td>
<td>Survey</td>
</tr>
<tr>
<td></td>
<td>Individual consumer</td>
<td></td>
<td></td>
<td>Qualitative or semi-structured interviews</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Organization or setting</td>
<td></td>
<td></td>
<td>Focus groups</td>
<td></td>
</tr>
<tr>
<td>Feasibility</td>
<td>Individual provider</td>
<td>Rogers: “compatibility” and “trialability”</td>
<td>Actual fit or utility; suitability for everyday use; practicability</td>
<td>Early (during adoption)</td>
<td>Survey</td>
</tr>
<tr>
<td></td>
<td>Organization or setting</td>
<td></td>
<td></td>
<td>Administrative data</td>
<td></td>
</tr>
<tr>
<td>Fidelity</td>
<td>Individual provider</td>
<td>RE-AIM: part of “implementation”</td>
<td>Delivered as intended; adherence; integrity; quality of program delivery</td>
<td>Early to mid</td>
<td>Observation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Checklists</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Self-report</td>
<td></td>
</tr>
<tr>
<td>Implementation Cost</td>
<td>Provider or providing institution</td>
<td>TCU Program Change Model: “costs” and “resources”</td>
<td>Marginal cost; cost-effectiveness; cost-benefit</td>
<td>Early for adoption and feasibility</td>
<td>Administrative data</td>
</tr>
<tr>
<td>Penetration</td>
<td>Organization or setting</td>
<td>RE-AIM: necessary for “reach”</td>
<td>Level of institutionalization? Spread? Service access?</td>
<td>Mid for penetration</td>
<td>Case audit</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Late for sustainability</td>
<td>Checklists</td>
</tr>
<tr>
<td>Sustainability</td>
<td>Administrators</td>
<td>RE-AIM: “maintenance” Rogers: “confirmation”</td>
<td>Maintenance; continuation; durability; incorporation; integration; institutionalization; sustained use; routinization;</td>
<td>Late</td>
<td>Case audit</td>
</tr>
<tr>
<td></td>
<td>Organization or setting</td>
<td></td>
<td></td>
<td>Semi-structured interviews</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Questionnaires</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Checklists</td>
<td></td>
</tr>
</tbody>
</table>
Adoption is defined as the intention, initial decision, or action to try or employ an innovation or evidence-based practice. Adoption also may be referred to as “uptake.” Our definition is consistent with those proposed by Rabin et al. (2008) and Rye and Kimberly (2007). Adoption could be measured from the perspective of provider or organization. Haug et al. (2008) used pre-post items to capture substance abuse providers’ adoption of evidence-based practices, while Henggeler et al. (2008) report interview techniques to measure therapists’ adoption of contingency management.

Appropriateness is the perceived fit, relevance, or compatibility of the innovation or evidence-based practice for a given practice setting, provider, or consumer; and/or perceived fit of the innovation to address a particular issue or problem. “Appropriateness” is conceptually similar to “acceptability,” and the literature reflects overlapping and sometimes inconsistent terms when discussing these constructs. We preserve a distinction because a given treatment may be perceived as appropriate but not acceptable, and vice versa. For example, a treatment might be considered a good fit for treating a given condition but its features (for example, rigid protocol) may render it unacceptable to the provider. The construct “appropriateness” is deemed important for its potential to capture some “pushback” to implementation efforts, as is seen when providers feel a new program is a “stretch” from the mission of the health care setting, or is not consistent with providers’ skill set, role, or job expectations. For example, providers may vary in their perceptions of the appropriateness of programs that co-locate mental health services within primary medical, social service, or school settings. Again, a variety of stakeholders will likely have perceptions about a new treatment’s or program’s appropriateness to a particular service setting, mission, providers, and clientele. These perceptions may be function of the organization’s culture or climate (Klein and Sorra 1996). Bartholomew et al. (2007) describe a rating scale for capturing appropriateness of training among substance abuse counselors who attended training in dual diagnosis and therapeutic alliance.

Cost (incremental or implementation cost) is defined as the cost impact of an implementation effort. Implementation costs vary according to three components. First, because treatments vary widely in their complexity, the costs of delivering them will also vary. Second, the costs of implementation will vary depending upon the complexity of the particular implementation strategy used. Finally, because treatments are delivered in settings of varying complexity and overheads (ranging from a solo practitioner’s office to a tertiary care facility), the overall costs of delivery will vary by the setting. The true cost of implementing a treatment, therefore, depends upon the costs of the particular intervention, the implementation strategy used, and the location of service delivery.

Much of the work to date has focused on quantifying intervention costs, e.g., identifying the components of a community-based heart health program and attaching costs to these components (Ronckers et al. 2006). These cost estimations are combined with patient outcomes and used in cost-effectiveness studies (McHugh et al. 2007). A review of literature on guideline implementation in professions allied to medicine notes that few studies report anything about the costs of guideline implementation (Callum et al. 2010). Implementing processes that do not require ongoing supervision or consultation, such as computerized medical record systems, may carry lower costs than implementing new psychosocial treatments. Direct measures of implementation cost are essential for studies comparing the costs of implementing alternative treatments and of various implementation strategies.

Feasibility is defined as the extent to which a new treatment, or an innovation, can be successfully used or carried out within a given agency or setting (Karsh 2004). Typically, the concept of feasibility is invoked retrospectively as a potential explanation of an initiative’s success or failure, as reflected in poor recruitment, retention, or participation rates. While feasibility is related to appropriateness, the two constructs are conceptually distinct. For example, a program may be appropriate for a service setting—in that it is compatible with the setting’s mission or service mandate, but may not be feasible due to resource or training requirements. Hides et al. (2007) tapped aspects of feasibility of using a screening tool for co-occurring mental health and substance use disorders.

Fidelity is defined as the degree to which an intervention was implemented as it was prescribed in the original protocol or as it was intended by the program developers (Dusenbury et al. 2003; Rabin et al. 2008). Fidelity has been measured more often than the other implementation...
outcomes, typically by comparing the original evidence-based intervention and the disseminated/implemented intervention in terms of (1) adherence to the program protocol, (2) dose or amount of program delivered, and (3) quality of program delivery. Fidelity has been the overriding concern of treatment researchers who strive to move their treatments from the clinical lab (efficacy studies) to real-world delivery systems. The literature identifies five implementation fidelity dimensions including adherence, quality of delivery, program component differentiation, exposure to the intervention, and participant responsiveness or involvement (Mihalic 2004; Dane and Schneider 2006). Adherence, or the extent to which the therapy occurred as intended, is frequently examined in psychotherapy process and outcomes research and is distinguished from other potentially pertinent implementation factors such as provider skill or competence (Hogue et al. 1996). Fidelity is measured through self-report, ratings, and direct observation and coding of audio- and videotapes of actual encounters, or provider-client/patient interaction. Achieving and measuring fidelity in usual care is beset by a number of challenges (Proctor et al. 2009; Mihalic 2004; Schoenwald et al. 2005). The foremost challenge may be measuring implementation fidelity quickly and efficiently (Hayes 1998).

Schoenwald and colleagues (2005) have developed three 26–45-item measures of adherence at the therapist, supervisor and consultant level of implementation (available from the MST Institute www.mstinstitute.org). Ratings are obtained at regular intervals, enabling examination of the provider, clinical supervisor, and consultant. Other examples from the mental health literature include Bond et al. (2008) 15-item Supported Employment Fidelity Scale (SE Fidelity Scale) and Hogue et al. (2008) Therapist Behavior Rating Scale-Competence (TBRS-C), an observational measure of fidelity in evidence based practices for adolescent substance abuse treatment.

Penetration is defined as the integration of a practice within a service setting and its subsystems. This definition is similar to (Stiles et al. 2002) notion of service penetration and to Rabin et al.’s (2008) notion of niche saturation. Studying services for persons with severe mental illness, Stiles et al. (2002) apply the concept of service penetration to service recipients (the number of eligible persons who use a service, divided by the total number of persons eligible for the service). Penetration also can be calculated in terms of the number of providers who deliver a given service or treatment, divided by the total number of providers trained in or expected to deliver the service. From a service system perspective, the construct is also similar to “reach” in the RE-AIM framework (Glasgow 2007b). We found infrequent use of the term penetration in the implementation literature; though studies seemed to tap into this construct with terms such a given treatment’s level of institutionalization.

Sustainability is defined as the extent to which a newly implemented treatment is maintained or institutionalized within a service setting’s ongoing, stable operations. The literature reflects quite varied uses of the term “sustainability,” but our proposed definition incorporates aspects of those offered by Johnson et al. (2004), Turner and Sanders (2006), Glasgow et al. (1999), Goodman et al. (1993), and Rabin et al. (2008). Rabin et al. (2008) emphasizes the integration of a given program within an organization’s culture through policies and practices, and distinguishes three stages that determine institutionalization: (1) passage (a single event such as transition from temporary to permanent funding), (2) cycle or routine (i.e., repetitive reinforcement of the importance of the evidence-based intervention through including it into organizational or community procedures and behaviors, such as the annual budget and evaluation criteria), and (3) niche saturation (the extent to which an evidence-based intervention is integrated into all subsystems of an organization). Thus the outcomes of “penetration” and “sustainability” may be related conceptually and empirically, in that higher penetration may contribute to long-term sustainability. Such relationships require empirical test, as we elaborate below. Indeed Steckler et al. (1992) emphasize sustainability in terms of attaining long-term viability, as the final stage of the diffusion process during which innovations settle into organizations. To date, the term sustainability appears more frequently in conceptual papers than actual empirical articles measuring sustainability of innovations. As we discuss below, the literature often uses the same term (niche saturation, for example) to reference multiple implementation outcomes, underscoring the need for conceptual clarity as we seek to advance in this paper.

Research Agenda to Advance Implementation Outcomes

Advancing the conceptualization, measurement, and empirical understanding of implementation outcomes requires research on several critical issues. We propose two major themes for this research—(1) conceptualization and measurement, and (2) theory building—and identify important issues within each of these themes.

Research on Conceptualization and Measurement of Implementation Outcomes

Research on several fronts is required to advance the conceptual and measurement properties of implementation outcomes, five of which we identify and discuss.
Consistency of Terminology

For each outcome listed in Table 1, we found literature using different and sometimes inconsistent terminology. Sometimes studies used different labels for what appear to be the same construct. In other cases, studies used one term for a label or nominal definition but a different term for operationalizing or measuring the same construct. This problem was pronounced for three implementation outcomes—acceptability, appropriateness, and feasibility. These constructs were frequently used interchangeably or measured under the common generic label as client or provider perceptions, reactions, and attitudes toward, or satisfaction with various aspects of the innovation, EST, or clinical practice guidelines. For example, Graham et al. (2007) assessed doctors’ attitudes and perceptions toward clinical practice guidelines with a survey that tapped all three of these outcomes, although none of them were explicitly labeled as such: acceptability (e.g., perceived quality of and confidence in guidelines), appropriateness (e.g., perceived usefulness of guidelines), and feasibility (e.g., these guidelines provide recommendations that are implementable). Other studies interchanged the terms for acceptability and feasibility within the same article. For example, Wilkie et al. (2003) begin by describing the measurement of “usability” (of a computerized innovation), including its “acceptability” to clients but later use the findings to conclude that the innovation was feasible.

While language inconsistency is typical in most still-developing fields, implementation research may be particularly susceptible to this problem. No one discipline is “home” to implementation research. Studies are conducted across a broad range of disciplines, published in a scattered set of journals, and consequently are rarely cross-referenced. Beyond mental health, we found articles referencing these implementation outcomes in physical health, smoking cessation, cancer, and substance abuse literatures, addressing a wide variety of topics.

Clearly, the field of implementation science now has only the beginnings of a common language to characterize implementation outcomes, a situation that thwarts the conceptual and empirical advancement of the field but could be overcome by use of a common lexicon. Just as Michie et al. (2009) state the “imperative that there be a consensual, common language” (p. 4) to describe behavior change techniques, so is common language needed for implementation outcomes.

Referent for Rating the Outcome

Several of the proposed implementation outcomes could be used to rate (1) a specific treatment; (2) the implementation strategy used to introduce that treatment into the care setting; or (3) a broad effort to implement several new treatments at once. A lingering issue for the field is whether implementation processes should be tackled and studied specifically (one new treatment) or in a more generalized way (the extent to which a system’s care is evidence-based or guideline congruent). Understanding the optimal specificity of the referent for a given implementation outcome is critical for measurement. As a beginning step, researchers should report the referent for all implementation outcomes measured.

Level of Analysis for Outcomes

Implementation of new treatments is an inherently multi-level enterprise, involving provider behavior, care organization, and policy (Proctor et al. 2009; Raghavan et al. 2008). Implementation outcomes are important at each level of change, but the research has yet to determine which level or unit of analysis is most appropriate for particular implementation outcomes. Certain outcomes, such as acceptability, may be most appropriate for individual level analysis (for example, providers, consumers), while others, such as penetration may be more appropriate for aggregate analysis, at the level of the health care organization. Currently, very few studies reporting implementation outcomes specify the level of measurement, nor do they address issues of aggregation within or across levels.

Construct validity. The constructs reflected in Table 1 and the terms employed in our taxonomy of implementation outcomes derive largely from the research literature. Yet it is important to also understand outcome perceptions and preferences through the voice of those who design and deliver health care. Qualitative data, reflecting language used by various stakeholders as they think and talk about implementation processes, is important for validating implementation outcome constructs. Through in-depth interviews, stakeholders’ cognitive representations and mental models of outcomes can be analyzed through such methods as cultural domain analysis (CDA). A “cultural domain” refers to a set of words, phrases, and/or concepts that link together to form a single conceptual subject (Luke 2004; Bates and Sarkar 2007), and methods for CDA, such as free-listing and pile-sorting, have been used since the 1970s (Bates and Sarkar 2007). While primarily used in anthropology, CDA is aptly suited for health services research that endeavors to understand how stakeholders conceptualize implementation outcomes, informing the generation of definitions of implementation outcomes. The actual words used by stakeholders may or may not reflect the terms used in academic literature and reflected in our proposed taxonomy (acceptability, appropriateness, feasibility, adoption, fidelity, penetration, sustainability and costs).
such research can identify the terms and distinctions that are meaningful to implementation stakeholders.

Measurement Properties of Implementation Outcomes

The literature reflects a wide array of approaches for measuring implementation outcomes, ranging from qualitative, quantitative survey, and record archival. Michie et al. (2007) studied perceived difficulties implementing a mental health guideline, coding respondent descriptions of implementation difficulties as 0, 0.5, or 1. Much measurement has been “home-grown,” with virtually no work on the psychometric properties or measurement rigor. Measurement development is needed to enhance the portability and usefulness of implementation outcomes in real world settings of care. Measures used in efficacy research will likely prove too cumbersome for real-world studies of implementation. For example, detailed assessment of fidelity through coding of encounter videotapes would be too time-intensive for a multi-agency study assessing fidelity of treatment implementation.

Theory-Building Research

Research is also needed to advance our theoretical understanding of the implementation process. Empirical studies of the five issues we list here will inform theory, illuminate the “black box” of implementation processes, and help shape models for developing and testing implementation strategies.

Salience of Implementation Outcomes to Stakeholders

Any effort to implement change in care involves a range of stakeholders, including the treatment developers who design and test the effectiveness of ESTs, policy makers who design and pay for service, administrators who shape program direction, providers and supervisors, patients/clients/consumers and their family members, and interested community members and advocates. The success of efforts to implement evidence-based treatment may rest on their congruence with the preferences and priorities of those who shape, deliver, and participate in care. Implementation outcomes may be differentially salient to various stakeholders, just as the salience of clinical outcomes varies across stakeholders (Shumway et al. 2003). For example, implementation cost may be most important to policy makers and program directors, feasibility may be most important to direct service providers, and fidelity may be most important to treatment developers. To ensure applicability of implementation outcomes across a range of settings and to maximize their external validity, all stakeholder groups and priorities should be represented in this research.

Salience of Implementation Outcomes by Point in the Implementation Process

The implementation of any new treatment or service is widely recognized as a process, involving a sequence of activities, beginning with initial considerations of what and how to change current care. Chamberlain has identified ten steps for the implementation of an evidence-based treatment, Multidimensional Treatment Foster Care (MTFC), beginning with consideration of adopting MTFC and concluding when a service site meets certification criteria for delivering the treatment (Chamberlain et al. 2008). As we suggest in Table 1, certain implementation outcomes may be more important at some phases of implementation process than at other phases. For example, feasibility may be most important once organizations and providers try new treatments. Later, it may be a “moot point,” once the treatment—initially considered novel or unknown—has become part of normal routine.

The literature suggests that studies usually capture fidelity during initial implementation, while adoption is often assessed at 6 (Waldorff et al. 2008), 12 (Adily et al. 2004; Fischer et al. 2008), or 18 months (Cooke et al. 2001) after initial implementation. But most studies fail to specify a timeframe or are inconsistent in choice of a time point in the implementation process for measuring outcomes. Research is needed to explore these issues, particularly longitudinal studies that measure multiple implementation outcomes before, during, and after implementation of a new treatment. Such research may reveal “leading” and “lagging” indicators of implementation success. For example, if acceptability increases for several months, following which penetration increases, then we may view acceptability as a leading indicator of penetration. Leading indicators can be useful for managing the implementation process as they signal future trends.

Where leading indicators may identify future trends, lagging indicators reflect delays between when changes happen and when they can be observed. For example, sustainability may be observed only well into, or even after the implementation process. Being aware of lagging indicators of implementation success may help managers avoid over-reacting to slow change and wait for evidence of what may soon prove to be successful implementation.

Modeling Interrelationships Among Implementation Outcomes

Our team’s observations of implementation suggest that implementation outcomes are themselves interrelated in dynamic and complex ways (Woolf 2008; Repenning 2002; Hovmand and Gillespie 2010; Klein and Knight 2005) and are likely to change throughout an agency’s process to
adopt and implement ESTs. For example, the perceived appropriateness, feasibility, and implementation cost associated with an intervention will likely bear on ratings of the intervention’s acceptability. Acceptability, in turn, will likely affect adoption, penetration, and sustainability. Similarly, consistent with Rogers’ theory of the diffusion of innovation, the ability to adopt or adapt an innovation for local use may increase its acceptability (Rogers 1995). This suggests that when providers believe they do not have to implement a treatment “by the book” (or with precise fidelity), they may rate the treatment as more acceptable.

Modeling the interrelationships between implementation outcomes will also inform their definitional boundaries and thus shape the taxonomy. For example, if two outcomes which we now define as distinct concepts are shown through research to always occur together, the empirical evidence would suggest that the concepts are really the same thing and should be combined. Similarly, if two of the outcomes are shown to have different empirical patterns, evidence would confirm their conceptual distinction.

Modeling Attainment of Implementation Outcomes

Once researchers have advanced consistent, valid, and efficient measures for implementation outcomes, the field will be equipped to conduct important research treating these constructs as dependent variables, in order to identify correlates or predictors of their attainment. Their measurement will enable research to determine which features of a treatment itself or which implementation strategies help make new treatments acceptable, feasible to implement, or sustainable over time. The diffusion of innovation literature posits that the implementation outcome, adoption of an EST, is a function of such factors as perceived need to do things differently (Rogers 1995) perception of the new treatment’s comparative advantage (Frambach and Schillewaert 2002; Henggeler et al. 2002) and as easy to understand (Berwick 2003). Such suppositions require empirical test using measures of implementation outcomes.

Using Implementation Outcomes to Model Implementation Success

Reliable, valid measures of implementation outcomes will enable empirical testing of the success of efforts to implement new treatments, and pave the way for comparative effectiveness research on implementation strategies. In most current initiatives to move evidence-based treatments into community care settings, the success of the implementation is assumed and evaluated from data on clinical outcomes. We believe that an exclusive focus on clinical outcomes thwarts understanding the process of implementation, as well as the effects of contextual factors that must be addressed and that are captured in implementation outcomes.

Established evidence for a “proven” treatment does not ensure successful implementation. Implementation also requires addressing a number of important contextual factors, such as provider attitudes, professional behavior, and the service system. Constructs in the proposed taxonomy of implementation outcomes have potential to capture those provider attitudes (acceptability) and behaviors (adoption, uptake) as well as contextual factors (system penetration, appropriateness, implementation cost).

For purposes of stimulating debate and future research, we suggest that successful implementation be considered in light of a “portfolio” of factors, including the effectiveness of the treatment to be implemented and implementation outcomes such as included in our taxonomy. For example, implementation success (I, in the equation below) could be modeled to reflect (1) the effectiveness (E) of the treatment being implemented, plus (2) implementation factors (IO’s), which heretofore have been insufficiently conceptualized, distinguished, and measured and rarely used to guide implementation decisions.

\[ I = fE + IO’s \]

For example, in situation “A”, an evidence-based treatment may be highly effective but given its high cost, only mildly acceptable to key stakeholders and low in sustainability. The overall potential success of implementation in this case might be modeled as follows:

Implementation success = \( f \) of effectiveness (\( = \) high) + acceptability (\( = \) moderate) + sustainability (low).

In situation “B”, a given treatment might be only moderately effective but highly acceptable to stakeholders because current care is poor, the treatment is inexpensive, and current training protocols ensure high penetration through providers. This treatment’s potential might be modeled in the following equation:

Implementation success = \( f \) of treatment effectiveness (moderate) + acceptability (high) + potential to improve care (high) + penetration (high).

Thus using implementation outcomes, the success of implementation may be modeled and tested, thereby making decisions about what to implement more explicit and transparent.

To increase the success of implementation, implementation strategies need to be employed strategically. For
example, implementation strategies could be employed to increase provider acceptance, improve penetration, reduce implementation costs, and achieve sustainability of the treatment being implemented. Understanding how to achieve implementation outcomes requires the kind of work now underway by Michie et al. (2009) advance a taxonomy of implementation strategies and reflect their demonstrated effects.

Summary and Implications

The science of implementation cannot be advanced without attention to implementation outcomes. All studies of implementation should explicate and measure implementation outcomes. Given the rudimentary state of the field, we chose a narrative approach to reviewing the literature and constructing a taxonomy. Our purpose is to advance the clarity of language, provoke debate, and stimulate more systematic work toward the aims of advancing the conceptual, linguistic, and methodological clarity in the field. A taxonomy of implementation outcomes can help organize the key variables and frame research questions required to advance implementation science. Their measurement and empirical test can help specify the mechanisms and causal relationships within implementation processes and advance an evidence base around successful implementation.

Acknowledgments

Preparation of this manuscript was supported by grants P30 MH068579 and UL1RR024992.

Open Access

This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References


Institute of Medicine Committee on Crossing the Quality Chasm. (2006). *Adaptation to mental health and addictive disorder: Improving the quality of health care for mental and substance-use conditions*. Washington, DC: Institute of Medicine, National Academies Press.


Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research

Nicole Gray Weiskopf, Chunhua Weng

ABSTRACT
Objective To review the methods and dimensions of data quality assessment in the context of electronic health record (EHR) data reuse for research.

Materials and methods A review of the clinical research literature discussing data quality assessment methodology for EHR data was performed. Using an iterative process, the aspects of data quality being measured were abstracted and categorized, as well as the methods of assessment used.

Results Five dimensions of data quality were identified, which are completeness, correctness, concordance, plausibility, and currency, and seven broad categories of data quality assessment methods: comparison with gold standards, data element agreement, data source agreement, distribution comparison, validity checks, log review, and element presence.

Discussion Examination of the methods by which clinical researchers have investigated the quality and suitability of EHR data for research shows that there are fundamental features of data quality, which may be difficult to measure, as well as proxy dimensions. Researchers interested in the reuse of EHR data for clinical research are recommended to consider the adoption of a consistent taxonomy of EHR data quality, to remain aware of the task-dependence of data quality, to integrate work on data quality assessment from other fields, and to adopt systematic, empirically driven, statistically based methods of data quality assessment.

Conclusion There is currently little consistency or potential generalizability in the methods used to assess EHR data quality. If the reuse of EHR data for clinical research is to become accepted, researchers should adopt validated, systematic methods of EHR data quality assessment.

As the adoption of electronic health records (EHRs) has made it easier to access and aggregate clinical data, there has been growing interest in conducting research with data collected during the course of clinical care. The National Institutes of Health has called for increasing the reuse of electronic health record (EHR) data reuse for research. Due to differences in measurement, recording, and still found a great deal of variability. The completeness of blood pressure recordings, for example, fell anywhere between 0.1% and 51%. Due to differences in measurement, recording, and still found a great deal of variability. The completeness of blood pressure recordings, for example, fell anywhere between 0.1% and 51%.
quality of an EHR-derived dataset for a given research task.

Our review primarily differs from those highlighted above in its focus. The previous reviews looked at data quality findings, while ours instead focuses on the methods that have been used to assess data quality. In fact, the earlier reviews were explicitly limited to studies that relied on the use of a reference standard, while we instead explore a range of data quality assessment methods. The contributions of this literature review are an empirically based conceptual model of the dimensions of EHR data quality studied by clinical researchers and a summary and critique of the methods that have been used to assess EHR data quality, specifically within the context of reusing clinical data for research. Our goal is to develop a systematic understanding of the approaches that may be used to determine the suitability of EHR data for a specific research goal.

METHODS
We identified articles in the literature by performing a search of the literature using standard electronic bibliographic tools. The literature search was performed by the first author on PubMed in February of 2012. As observed by Hogan and Wagner in their literature review, there is no medical subheadings (MeSH) term for data quality, so a brief exploratory review was performed to identify relevant keywords. The final list included ‘data quality’, ‘data accuracy’, ‘data reliability’, ‘data validity’, ‘data consistency’, ‘data completeness’, and ‘data error’. The MeSH heading for EHR was not introduced until 2010, so the older and more general MeSH heading ‘medical record systems, computerized’ was used instead. The phrases ‘EHR’, ‘electronic medical record’, and ‘computerized medical record’ were also included in order to capture articles that may not have been tagged correctly. We searched for articles including at least one of the quality terms and at least one of the EHR terms. Results were limited to English language articles. The full query is shown below:

’data quality’ OR ‘data accuracy’ OR ‘data reliability’ OR ‘data validity’ OR ‘data consistency’ OR ‘data completeness’ OR ‘data errors’ OR ‘data error’) AND (EHR OR electronic medical record OR computerized medical record OR medical records systems, computerized [mh]) AND English[lang]

This search produced 230 articles, all of which were manually reviewed by the first author to determine if they met the selection criteria. In particular, the articles retained for further review: (1) included original research using data quality assessment methods; (2) focused on data derived from an EHR or related system; and (3) were published in a peer-reviewed journal. Articles dealing with data from purely administrative systems (eg, claims databases) were not included. These inclusion criteria resulted in 44 relevant articles. Next, we performed an in-depth ancestor search, reviewing the references of all of the articles in the original pool of 44. This allowed us to identify an additional 51 articles, resulting in a final pool of 95 articles meeting our inclusion criteria that were then used to derive results in this study.

From each article we abstracted the features of data quality examined, the methods of assessment used, and basic descriptive information including about the article and the type of data being studied. Through iterative review of the abstracted data, we derived broad dimensions of data quality and general categories of assessment strategies commonly described in the literature. Finally, we reviewed the 95 articles again, categorizing every article based on the dimension or dimensions being assessed, as well as the assessment strategies used for each of those dimensions.

Before beginning this analysis, we searched for preexisting models of EHR data quality, but were unable to find any. We decided that the potential benefits of adapting a data quality model from another field were outweighed by the risks of approaching our analysis through the lens of a model that had not been validated in the area of EHR data quality. Furthermore, using an existing model to guide analysis is a deductive approach, which has the potential to obscure information contained in the data. By imposing an existing model from a different discipline, we would have run the risk of missing important findings. Therefore, we decided to use an inductive, data-driven coding approach. This approach provides advantages over the deductive approach by allowing us better coverage of the dimensions and methods of data quality assessment.

RESULTS
The majority of papers reviewed (73%) looked at structured data only, or at a combination of structured and unstructured data (22%). For our purposes, unstructured data types include free-entry text, while structured data types include coded data, values from pre-populated lists, or data entered into fields requiring specific alphanumeric formats.

Ignoring variations due to lexical categories and negation, the articles contained 27 unique terms describing dimensions of data quality. Features of data quality that were mentioned or described but not assessed were not included in our analysis. We grouped the terms together based on shared definitions. A few features of good data described in the literature, including sufficient granularity and the use of standards, were not included in our analyses. This decision was made due to the limited discussion of these features, the fact that they could be considered traits of good data practice instead of data quality, and because no assessment methods were described. Overall, we empirically derived five substantively different dimensions of data quality from the literature. The dimensions are defined below:

▸ Completeness: Is a truth about a patient present in the EHR?
▸ Correctness: Is an element that is present in the EHR true?
▸ Concordance: Is there agreement between elements in the EHR, or between the EHR and another data source?
▸ Plausibility: Does an element in the EHR makes sense in light of other knowledge about what that element is measuring?
▸ Currency: Is an element in the EHR a relevant representation of the patient state at a given point in time?

The list of data quality terms and their mappings to the five dimensions described above are shown in table 1. The terms chosen to denote each of the dimensions were the clearest and

Table 1 Terms used in the literature to describe the five common dimensions of data quality

<table>
<thead>
<tr>
<th>Completeness</th>
<th>Correctness</th>
<th>Concordance</th>
<th>Plausibility</th>
<th>Currency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility</td>
<td>Accuracy</td>
<td>Agreement</td>
<td>Accuracy</td>
<td>Recency</td>
</tr>
<tr>
<td>Accuracy</td>
<td>Corrections made</td>
<td>Consistency</td>
<td>Believability</td>
<td>Timeliness</td>
</tr>
<tr>
<td>Availability</td>
<td>Errors</td>
<td>Consistency</td>
<td>Reliability</td>
<td>Trustworthiness</td>
</tr>
<tr>
<td>Omission</td>
<td>Misleading</td>
<td>Predictive value</td>
<td>Validity</td>
<td></td>
</tr>
<tr>
<td>Omission</td>
<td>Positive</td>
<td>Predictive value</td>
<td>Validity</td>
<td></td>
</tr>
<tr>
<td>Presence</td>
<td>Quality</td>
<td>Validity</td>
<td>Validity</td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td>Validity</td>
<td>Rate of recording</td>
<td>Validity</td>
<td></td>
</tr>
<tr>
<td>Sensitivity</td>
<td>Validity</td>
<td>Validity</td>
<td>Validity</td>
<td></td>
</tr>
<tr>
<td>Validity</td>
<td>Validity</td>
<td>Rate of recording</td>
<td>Validity</td>
<td></td>
</tr>
</tbody>
</table>
least ambiguous from each of the groups. There was a great deal of variability and overlap in the terms used to describe each of these dimensions. ‘Accuracy’, for example, was sometimes used as a synonym for correctness, but in other articles meant both correctness and completeness. The dimensions themselves, however, were abstracted in such a way as to be exhaustive and mutually exclusive based on their definitions. Every article identified could be matched to one or more of the dimensions.

A similar process was used to identify the most common methods of data quality assessment. The strategies used to assess the dimensions of data quality fell into seven broad categories of methods, many of which were used to assess multiple dimensions. These general methods are listed and defined below:

- **Gold standard**: A dataset drawn from another source or multiple sources, with or without information from the EHR, is used as a gold standard.
- **Data element agreement**: Two or more elements within an EHR are compared to see if they report the same or compatible information.
- **Element presence**: A determination is made as to whether or not desired or expected data elements are present.
- **Data source agreement**: Data from the EHR are compared with data from another source to determine if they are in agreement.
- **Distribution comparison**: Distributions or summary statistics of aggregated data from the EHR are compared with the expected distributions for the clinical concepts of interest.
- **Validity check**: Data in the EHR are assessed using various techniques that determine if values ‘make sense’.
- **Log review**: Information on the actual data entry practices (eg, dates, times, edits) is examined.

A summary of which methods were used to assess which dimensions is shown in table 2. The graph in figure 1 shows the strength of the pairwise relationships between the dimensions and methods. Some of the methods were used to assess only certain dimensions of data quality, whereas other methods were applied more broadly. Element presence, for example, was used to assess completeness, but none of the other dimensions. Data element agreement and data source agreement, however, were applied more broadly. Most of the dimensions were assessed using an assortment of methods, but currency was only measured using a single approach.

### Completeness

Completeness was the most commonly assessed dimension of data quality and was an area of focus in 61 (64%) of the articles. Generally speaking, completeness referred to whether or not a truth about a patient was present in the EHR. Most of the articles used the term completeness to describe this dimension, but some also referred to data availability or missing data. In others, completeness was subsumed into more general concepts like accuracy or quality. Some articles cited the statistical definition of completeness suggested by Hogan and Wagner,2 in which completeness is equivalent to sensitivity.

Many articles assessed EHR data completeness by using another source of data as a gold standard. The gold standards used included concurrently kept paper records,12–18 information supplied by patients,19–21 review of data by patients,22–25 clinical encounters with patients,26–28 information presented by trained standard patients,29 30 information requested from the treating physician,31 and alternative data sources from which EHR elements were abstracted.32 33 A similar approach involved triangulating data from multiple sources within the EHR to create a gold standard.34 35

Other researchers simply looked at the presence or absence of elements in the EHR. In some cases, these were elements that were expected to be present, even if they were not needed for any specific task.36–38 57 58 68 69 74–81 In other situations, the elements examined were dependent upon the task at hand, meaning that the researchers determined whether or not the EHR data were

---

### Table 2  The dimensions of data quality and methods of data quality assessment

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Completeness</th>
<th>Correctness</th>
<th>Concordance</th>
<th>Plausibility</th>
<th>Currency</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Method</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gold standard</td>
<td>12–35</td>
<td>12–23 25 26 28–48</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data element agreement</td>
<td>49–55 56</td>
<td>49 50 52–54 56–67</td>
<td>26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Element presence</td>
<td>36–39 49 57–59 68 69 74–86</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data source agreement</td>
<td>79 87–89</td>
<td>90</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distribution comparison</td>
<td>97–100</td>
<td>31 57 97</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validity checks</td>
<td>32 62 104–106</td>
<td>73</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log review</td>
<td>73</td>
<td>16</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>61</td>
<td>57</td>
<td>16</td>
<td>7</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

In decreasing order of frequency, the dimensions are listed from left to right, and the methods from top to bottom. The numbers in the cells correspond to the article references featuring each dimension—method pair.
complete enough for a specific purpose. Other methods for assessing completeness included looking at agreement between elements from the same source, agreement between the EHR and paper records, agreement between the EHR and another electronic source of data, and comparing distributions of occurrences of certain elements between practices or with nationally recorded rates.

### Correctness

The second most commonly assessed dimension of data quality was correctness, which was included in 57 (60%) of the articles. EHR data were considered correct when the information they contained was true. Other terms that were commonly used to describe this concept included accuracy, error, and quality. Occasionally, correctness included completeness, due to the fact that some researchers consider missing data to be incorrect (i.e., errors of omission). The definition of correctness suggested by Hogan and Wagner states that data correctness is the proportion of data elements present that are correct, which is equivalent to positive predictive value.

Comparison of EHR data with a gold standard was by far the most frequently used method for assessing correctness. These gold standards included: paper records, information supplied by patients through interviews, or direct data entry; clinical encounters with patients; information presented by trained and credentialed staff members; automatically recorded data; contact with the treating physician; and alternative data sources from which information matching EHR elements were abstracted. Some researchers developed gold standards by extracting and triangulating data from within the EHR.

The second most common approach to assessing correctness was to look at agreement between elements within the EHR. This involved verifying a diagnosis by looking at associated procedures, medications, or laboratory values. Similarly, some articles reported on agreement between related elements and errors identified through the examination of the use of copy and paste practices. Other researchers looked specifically at agreement between structured elements and unstructured data within EHRs. One of the more formal approaches described for assessing correctness was the data quality probe, proposed by Brown and Warmington, which is a query that, when run against an EHR database, only returns cases with some disagreement between data elements.

A few articles described the use of validity checks to assess correctness. These included review of changes of sequential data over time, identifying end digit preferences in blood pressure values, and comparing elements with their expected value ranges. Two other approaches were to use corrections seen in log files as a proxy for correctness and comparing data on the same patients from a registry and an EHR.

### Concordance

Sixteen (17%) of the articles reviewed assessed concordance. Data were considered concordant when there was agreement or compatibility between data elements. This may mean that two elements recording the same information for a single patient have the same value, or that elements recording different information have values that make sense when considered together (e.g., biological sex is recorded as female and procedure is recorded as gynecological examination). Measurement of concordance is generally based on elements contained within the EHR, but some researchers also included information from other data sources. Common terms used in the literature to describe data concordance include agreement and consistency.

The most common approach to assessing concordance was to look at agreement between elements within the EHR, especially diagnoses and associated information such as medications or procedures. The second most common method used to assess concordance was to look at the agreement of EHR data with data from other sources. These other sources included billing information, paper records, patient-reported data, and physician-reported data. Another approach was to compare distributions of data within the EHR with distributions of the same information from similar medical practices or with national rates.

### Pseudonymity

Seven (7%) of the articles assessed the plausibility of EHR data. In this context, data were plausible if they were in agreement with general medical knowledge or information and were therefore feasible. In other words, assessments of plausibility were intended to determine whether or not data could be trusted or if they were of suspect quality. Other terms that were used to discuss and describe EHR data plausibility include data validity and integrity.

The most common approach to assessing the plausibility of EHR data was to perform some sort of validity check to determine if specific elements within the EHR were likely to be true or not. This included looking for elements with values that were outside biologically plausible ranges or that changed implausibly over time or zero-valued elements. Other researchers compared distributions of data values between practices and with national rates, or looked at agreement between related elements.

### Currency

The currency of EHR data was assessed in four (4%) of the 95 articles. Currency was often referred to in the literature as timeliness or recency. Data were considered current if they were recorded in the EHR within a reasonable period of time following measurement or, alternatively, if they were representative of the patient state at a desired time of interest. In all four articles, currency was assessed through the review of data entry logs. In three of the four, researchers reviewed whether desired data were entered into the EHR within a set time limit.

The fourth, researchers considered whether each type of data element was measured recently enough to be considered medically relevant.

### DISCUSSION

We identified five dimensions of data quality and seven categories of data quality assessment methods. Examination of the types of methods used, as well as overlap of the methods between dimensions, reveals significant patterns and gaps in knowledge. Below, we explore the major findings of the literature review, specifically highlighting areas that require further attention, and make suggestions for future research.

### Terminology and dimensions of data quality

One of the biggest difficulties in conducting this review resulted from the inconsistent terminology used to discuss data quality. We had not expected, for example, the overlap of terms between dimensions, or the fact that the language within a single article was sometimes inconsistent. The clinical research community has largely failed to develop or adopt a consistent taxonomy of data quality.
There is, however, overlap between the dimensions of data quality identified during this review and those described in preexisting taxonomies and models of data quality. Wang and Strong’s conceptual framework of data quality, for example, contains 15 dimensions, grouped into four categories: intrinsic, contextual, representational, and accessibility. Our review focused on intrinsic (inherent to the data) and contextual (task-dependent) data quality issues. The dimensions we identified overlapped with two of the intrinsic features: accuracy and believability, which are equivalent to correctness and plausibility; and two of the contextual features: timeliness, which is equivalent to currency, and completeness. The only dimension we identified that does not appear in Wang and Strong’s framework is concordance.

The Institute of Medicine identified four attributes of data quality relevant to patient records: completeness, accuracy, legibility, and meaning (related to comprehensibility). As the Institute of Medicine points out, electronic records by their nature negate many of the concerns regarding legibility, so we are left with three relevant attributes, two of which we identified through our review. Meaning is a more abstract concept and is likely to be difficult to measure objectively, which may be why we did not observe assessments of this dimension in the literature.

Although the five dimensions of data quality derived during our review were treated as mutually exclusive within the literature, we feel that only three can be considered fundamental: correctness, completeness, and currency. By this we mean that these dimensions are non-reducible, and describe core concepts of data quality as it relates to EHR data reuse. Concordance and plausibility, on the other hand, while discussed as separate features of data quality, appear to serve as proxies for the fundamental dimensions when it is not possible to assess them directly. This supposition is supported by the overlap observed in the methods used to assess concordance and plausibility with those used to assess correctness and completeness. A lack of concordance between two data sources, for example, indicates error in one or both of those sources: an error of omission, resulting in a lack of completeness, or an error of commission, resulting in a lack of correctness. Similarly, data that do not appear to be plausible may be incorrect, as in the case of a measurement that fails a range check, or incomplete, such as aggregated diagnosis rates within a practice that do not match the expected population rates. It may be that correctness, completeness, and currency are properties of data quality, while plausibility and concordance are methodological approaches to assessing data quality. In addition, researchers may refer to plausibility or concordance when they believe that there are problems with completeness or correctness, but have no way to be certain that errors exist or which data elements might be wrong.

**Data quality assessment methodology**

We observed a number of noteworthy patterns within the literature in terms of the types of data quality assessments used and the manner in which data quality assessment was discussed. For example, 37 of the 95 articles in our sample relied on a gold standard to assess data quality. There are a few problems with this approach. First, the data sources used could rarely be considered true gold standards. Paper records, for example, may sometimes be more trusted than electronic records, but they should not be considered entirely correct or complete. Perhaps more importantly, a gold standard for EHR data is simply not available in most cases. This will become more problematic as the use of de-identified datasets for research becomes more common. A ‘fitness for purpose’ approach, which suggests that the quality of each dataset compiled for a specific task must be assessed, necessitates the adoption of alternatives to gold standard-based methods.

In addition to the overreliance on gold standards, the majority of the studies we identified relied upon an ‘intuitive’ understanding of data quality and used ad hoc methods to assess data quality. This tendency has also been observed in other fields. Most of the studies included in this review presented assessment methodologies that were developed with a minimal empirical or theoretical basis. Only a few researchers made the effort to develop generalizable approaches that could be used as a step towards a standard methodology. Faulconer and de Luigan, for example, proposed a multistep, statistically driven approach to data quality assessment. Hogan and Wagner suggested specific statistical measures of the correctness and completeness of EHR data elements that have been adopted by other researchers. Certain methods, including comparing distributions of data from the EHR with expected distributions or looking for agreement between elements within the EHR, lend themselves more readily to generalization. Brown and Warmington’s data quality probes, for example, could be extended to various data elements, although they require detailed clinical knowledge to implement. Some researchers looking at the quality of research databases pulled from general practices in the UK have adopted relatively consistent approaches to comparing the distributions of data concerning specific clinical phenomena to information from registries and surveys.

In most cases, however, the specific assessment methods described in the literature would be difficult to apply to other datasets or research questions. If the reuse of EHR data for clinical research is to become common and feasible, development of standardized, systematic data quality assessment methods is vital.

In addition, if as a field we intend to adopt the concept of ‘fitness for purpose’, it is important to consider the intended research use of EHR data when determining if they are of sufficient quality. Some dimensions may prove to be more task dependent, or subjective, while others are essentially task independent, or objective. It will be important to develop a full understanding of the interrelationships of research tasks and data characteristics as they relate to data quality. For example, the completeness of a set of data elements required by one research protocol may differ from the completeness required for a different protocol. Many factors, including clinical focus, required resolution of clinical information, and desired effect size, can affect the suitability of a dataset for a specific research task.

**Future directions**

We believe that efforts to reuse EHR data for clinical research would benefit most from work in a few specific areas: adopting a consistent taxonomy of EHR data quality; increasing awareness of task dependence; integrating work on data quality assessment from other fields; and adopting systematic, statistically based methods of data quality assessment. A taxonomy of data quality would enable a structured discourse and contextu-

ize assessment methodologies. The findings in this review regarding the dimensions of data quality may serve as a stepping stone towards this goal. Task dependence is likely to become a growing issue as efforts to reuse EHR data for research increase, particularly as data quality assessment does not have a one-size-fits-all solution. One approach to addressing the problem of EHR data quality and suitability for reuse in research...
would be to look at what has been done outside of clinical research, because data quality has been an area of study in fields ranging from finance to industrial engineering. Finally, it is important that the clinical researchers begin to move away from ad hoc approaches to data quality assessment. Validated methods that can be adapted for different research questions are the ideal goal.

Limitations

There were a number of limitations to this review. First, the search was limited. Due to the lack of MeSH terms for data quality and the variation in terminology used to discuss data quality, it is possible that our original search may have missed some relevant articles. We believe that our decision to review the references of each article improved the saturation of our sample.

It is also important to note that our classification process was largely subjective and was performed by only one of the authors. It is possible that the original researchers might disagree with our interpretations. We chose to use an iterative process to label and categorize the dimensions of data quality and methods of assessment described in each article in an effort to develop a consistent coding scheme.

Finally, it is likely that the dearth of literature discussing data quality in the reuse of EHR data for clinical research is due partly to underreporting. A common first step in analyzing any dataset is to review distributions, summary statistics, and histograms, but this process is rarely described in publications. Such methods are therefore likely to be underrepresented in this review. Greater transparency regarding data cleaning or checking steps would be advisable, as it could help to establish acceptable reporting standards for the reuse of EHR data in research.

CONCLUSION

The secondary use of EHR data is a promising area of research. However, the problems with EHR data quality necessitate the use of quality assessment methodologies to determine the suitability of these data for given research tasks. In this review of the literature we have identified the major dimensions of data quality that are of interest to researchers, as well as the general assessment techniques that have been utilized. Data quality is not a simple problem, and if the reuse of EHR data is to become an accepted approach to medical research, the clinical research community needs to develop validated, systematic methods of EHR data quality assessment. We encourage researchers to be consistent in their discussion of the dimensions of data quality, systematic in their approaches to measuring data quality, and to develop and share best practices for the assessment of EHR data quality in the context of reuse for clinical research.

Acknowledgments

The authors would like to thank George Hippscuk, Adam Wilcox, and Suzanne Balken for their assistance in the preparation of this manuscript.

Funding

This research was supported by National Library of Medicine grants 5T15LM007079, R01LM009886, and R01LM010815.

Competing interests

None.

Provenance and peer review

Not commissioned; externally peer reviewed.

Open Access

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/3.0/

REFERENCES


Effects of Two Commercial Electronic Prescribing Systems on Prescribing Error Rates in Hospital In-Patients: A Before and After Study

Johanna I. Westbrook1,*, Margaret Reckmann1, Ling Li1, William B. Runciman2, Rosemary Burke3, Connie Lo1, Melissa T. Baysari4, Jeffrey Braithwaite5, Richard O. Day6

1 Centre for Health Systems and Safety Research, Australian Institute of Health Innovation, Faculty of Medicine, University of New South Wales, Sydney, Australia, 2 School of Psychology, Social Work & Social Policy, University of South Australia, Adelaide, Australia, 3 Pharmacy Department, Concord Repatriation General Hospital, Sydney, Australia, 4 Australian Institute of Health Innovation, Faculty of Medicine, University of New South Wales, Sydney, Australia, 5 Centre for Clinical Governance Research, Australian Institute of Health Innovation, Faculty of Medicine, University of New South Wales, Sydney, Australia, 6 Department of Clinical Pharmacology and Toxicology, St Vincent’s Hospital, Sydney, and Faculty of Medicine, University of New South Wales, Sydney, Australia

Abstract

Background: Considerable investments are being made in commercial electronic prescribing systems (e-prescribing) in many countries. Few studies have measured or evaluated their effectiveness at reducing prescribing error rates, and interactions between system design and errors are not well understood, despite increasing concerns regarding new errors associated with system use. This study evaluated the effectiveness of two commercial e-prescribing systems in reducing prescribing error rates and their propensities for introducing new types of error.

Methods and Results: We conducted a before and after study involving medication chart audit of 3,291 admissions (1,923 at baseline and 1,368 post e-prescribing system) at two Australian teaching hospitals. In Hospital A, the Cerner Millennium e-prescribing system was implemented on one ward, and three wards, which did not receive the e-prescribing system, acted as controls. In Hospital B, the iSoft MedChart system was implemented on two wards and we compared before and after error rates. Procedural (e.g., unclear and incomplete prescribing errors) and clinical (e.g., wrong dose, wrong drug) errors were identified. Prescribing error rates per admission and per 100 patient days; rates of serious errors (5-point severity scale, those ≥3 were categorised as serious) by hospital and study period; and rates and categories of postintervention “system-related” errors (where system functionality or design contributed to the error) were calculated. Use of an e-prescribing system was associated with a statistically significant reduction in error rates in all three intervention wards (respectively reductions of 66.1% [95% CI 53.9%–78.3%]; 57.5% [33.8%–81.2%]; and 60.5% [48.5%–72.4%]). The use of the system resulted in a decline in errors at Hospital A from 6.25 per admission (95% CI 5.23–7.28) to 2.12 (95% CI 1.71–2.54; p<0.0001) and at Hospital B from 3.62 (95% CI 3.30–3.93) to 1.46 (95% CI 1.20–1.73; p<0.0001). This decrease was driven by a large reduction in unclear, illegal, and incomplete orders. The Hospital A control wards experienced no significant change (respectively –12.8% [95% CI –41.1% to 15.5%]; –11.3% [–40.1% to 17.5%]; –20.1% [–52.2% to 12.4%]). There was limited change in clinical error rates, but serious errors decreased by 44% (0.25 per admission to 0.14; p=0.0002) across the intervention wards compared to the control wards (17% reduction; 0.30–0.25; p=0.40). Both hospitals experienced system-related errors (0.73 and 0.51 per admission), which accounted for 35% of postsystem errors in the intervention wards; each system was associated with different types of system-related errors.

Conclusions: Implementation of these commercial e-prescribing systems resulted in statistically significant reductions in prescribing error rates. Reductions in clinical errors were limited in the absence of substantial decision support, but a statistically significant decline in serious errors was observed. System-related errors require close attention as they are frequent, but are potentially remediable by system redesign and user training. Limitations included a lack of control wards at Hospital B and an inability to randomize wards to the intervention.

Please see later in the article for the Editors’ Summary.


Academic Editor: Aziz Sheikh, Edinburgh University, United Kingdom

Received June 28, 2011; Accepted December 16, 2011; Published January 31, 2012

Copyright: © 2012 Westbrook et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The research was funded by grants from the National Health and Medical Research Council (NHMRC) Program grant 568612 and Project grant 400929. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: WBR has in the past had shares, and acted as a consultant to give general advice on incident reporting to Patient Safety International (PSI), which was then bought by ISOFT, which in turn has been bought by Computing Services Corporation. He was a board member of PSI, which at that stage was, but is no longer, a subsidiary of the not-for-profit research organisation, the Australian Patient Safety Foundation (APSF), of which he has been and is President. RB spoke on “Electronic medication management and the Hospital pharmacist” at the Cerner Regional User Group meeting in 2010. This was at the request of her employer. She personally received no honorarium or fees from this speaking engagement. All other authors declared that no competing interests exist.

E-mail: JWESTBROOK@UNSW.EDU.AU

Current address: Information Management and Technology Division, Sydney South West Area Health Service, Sydney, Australia
Introduction

It is well over a decade since electronic prescribing systems were first shown to reduce medication errors [1,2], demonstrating their potential to address this long-standing, costly problem [3–5]. However, recent reviews [6–9] reveal that many questions remain unanswered regarding the extent to which systems deliver improvements in medication safety in different settings, important contextual and work practice factors associated with effectiveness, and the cost benefit of systems. To date, evidence of effectiveness rests largely on the experiences of a few hospitals using home-grown systems.

A central question is whether commercial e-prescribing systems can deliver the same benefits as home-grown systems. There is little work comparing commercial systems or the interactions between system design and error rates and types, despite increasing concerns regarding new errors associated with their use [8,10,11]. Implementation of these organisation-wide clinical information systems is complex [12,13] with a multitude of work process and cultural factors [14–16], which affect system adoption and use, driving both intended and unintended outcomes [10,11,17,18].

In 2011, the US Agency for Healthcare Research and Quality [8] released a review of the effects of health information technology on medication management and drew attention to the need for research that evaluates systems in everyday settings and allows comparisons between systems and study sites. Our aim was to evaluate two commercial e-prescribing systems with respect to their effectiveness in reducing prescribing errors and their propensities for introducing new types of error.

Methods

Sample and Data Collection

A before and after study design was implemented at two major teaching hospitals in Sydney, Australia. Hospital A had 400 beds and Hospital B 326 beds. At Hospital A data were collected from four wards pre and post e-prescribing system implementation (two geriatric, a renal/vascular, and a respiratory ward). One ward (geriatric) was assigned the intervention and the remaining three wards acted as controls. At Hospital B the intervention was implemented on two wards (psychiatry and cardiology), and error rates were evaluated in the pre and post e-prescribing implementation periods. Figure 1 outlines the study design.

A daily review of all inpatient medication charts (n = 3,291) was conducted by three pharmacists independent from the hospitals for at least two months pre- and postintervention, with the exception of the psychiatric ward (1 mo pre and post). Data collection at Hospital A was conducted between May–August 2006 (pre) and May–August 2008 (28 wk post e-prescribing system), and at Hospital B between November 2007–March 2008 (pre), and March 2008–February 2010 (16 and 10 wk post system introduction). Data collection was dictated by the hospitals’ e-prescribing system implementations, which experienced several delays. Human research ethics approval was received from both hospitals and the University of Sydney.

Error Classification

Errors were classified into procedural (three categories) or clinical errors (14 categories) (Table S1 lists error definitions). Prescribing errors identified in the intervention wards in the postperiod were additionally reviewed to assess whether or not they were “system-related” (see definitions Table S1). System-related errors were defined as errors where system functionality or design contributed to the error, and there was little possibility that another cause, such as a lack of knowledge, produced the error. For example, an order for an inappropriate drug located on a drop-down menu next to a likely drug selection was flagged as a system-related error. Thus all system-related errors underwent dual classification in terms of (1) their manifestation according to one of the 17 procedural or clinical error categories and (2) the system-related mechanism that was deemed to be associated with those errors. In this paper, the system-related errors are reported according to their clinical manifestation and are listed in a separate table, as strategies for their prevention are likely to relate to system redesign or improved functionality.

Figure 1. Outline of study design.
doi:10.1371/journal.pmed.1001164.g001
Inter-rater reliability tests were conducted at regular intervals and compared pharmacist reviewers’ agreement with respect to number and type of errors. These tests involved double audit of 10% of all admissions and produced kappa scores of 0.82–0.84. In the last stage of the research, 1,097 admissions (33% of the total sample) were re-reviewed in order to ensure consistency of data collection between the early and later data collection periods. Two pharmacists independently rated the actual or potential severity of errors (Box 1); disagreement was settled by consensus with input from a clinical pharmacist (ROD) when required. Severity review committees involving an emergency physician, hospital pharmacists, and nurses from both hospitals were also given subsets of errors to classify during the study.

Hospital Prescribing and the Interventions

In the preintervention period all wards used paper medication charts in which the prescribing doctors wrote orders. These charts were then used by nursing staff as the medication administration charts. There was no intermediate transcription step between a prescriber’s order and the final medication chart entry, as is the case in some countries.

Ward pharmacy services were provided during the weekdays but not on weekends. The research pharmacists’ daily review of the medication charts may have occurred either before or after the ward pharmacists had done their rounds. All interventions (corrections) made by the ward pharmacists in patients’ medication charts were identifiable and noted (i.e., errors detected by the ward pharmacists were included in the study).

Interventions consisted of the implementation of two e-prescribing systems (Cerner Millennium PowerOrders and iSoft MedChart) integrated with each hospitals’ computerised order entry system. Prescribers were required to use the systems to prescribe medications in the post period.

Hospital A implemented the Cerner system, where prescribing is mainly by menu selection of pre-prepared order sentences that are triggered upon drug selection and that can be modified by the prescriber. “Care sets” allow for a group of related orders to be selected and ordered simultaneously with a single click. Unlisted medications and prescribing order comments need to be generated by the prescriber. In the Cerner e-prescribing system, active decision support at the time of study consisted of allergy alerts and drug–drug interaction alerts set at the most severe level (using the Multum database). Medication orders could not be completed if the patient’s allergy status was not recorded. If a prescriber wished to over-ride an alert they needed to select an override reason from a drop-down list or free text boxes. Passive decision support included a drug information database, the highly drug–drug interaction alerts were not operational during the study. All alerts allowed the prescriber to continue with the order. Alerts were all “pop-ups” on the screen. Approximately half of the alerts were for information only; prescribers were not required to take action and just had to close the alert box. Others required the prescriber to respond by ticking an “override” box.

For approximately 10% of the alerts prescribers were required to enter a free-text reason for overriding the alert in order to proceed. Drug information references were available online as passive decision support.

During the intervention periods both sites used paper orders for a small subset of medications. At Hospital A, heparin infusions and patient-controlled analgesia remained on paper charts.

At Hospital B, orders for intravenous (IV) fluids, IV infusions (e.g., heparin infusion), variable dose regimes (such as titrated or reducing doses), insulins, oral anticoagulants (warfarin), chemotherapy, parenteral nutrition, and epidural or patient-controlled analgesia remained on paper charts. The prescriber was required to order an electronic prompt to signal the administration times for these drugs, but the actual drug orders were located on a paper chart. Errors related to these electronic prompts were included in the postperiod data collection.

Statistical Analysis

The error data were linked with the patient admission data, which matched the study periods. Rates of prescribing errors per admission and per 100 patient days were calculated for each error type and category, by period (pre/post), group (intervention/control), hospital, and ward. Serious errors [grade(d≥3)] (Box 1) were examined by group, error type, and period. System-related error rates per admission were examined for both systems. The 95% CIs for the average error rates per admission and per 100 patient days were calculated using the large sample approximation of mean ±1.96 × standard error. For the pre- and postanalysis, two-sample t-tests were used to compare baseline data with post e-prescribing system data with the level of significance set at 5%. The 95% CIs for percentage changes were calculated as per the Fieller CI [19]. All statistical analyses were carried out with SAS 9.2 [20].

Box 1. Severity Assessment Code [47]

**Minor errors**
1. Insignificant: Incident is likely to have little or no effect on the patient.
2. Minor: Incident is likely to lead to an increase in level of care e.g. review, investigations, or referral to another clinician.

**Serious errors**
3. Moderate: Incident is likely to lead to permanent reduction in bodily functioning, increased length of stay, surgical intervention.
4. Major: Incident is likely to lead to a major permanent loss of function.
5. Serious: Incident is likely to lead to death.
Results

Incidence, Type, and Severity of Prescribing Errors at Baseline

The 1,923 admissions across the six wards reviewed at baseline revealed 11,168 prescribing errors, an average of 5.8 per admission. The majority (n = 8,225; 73.6%; 4.28 per admission) were procedural (e.g., unclear, incomplete, or illegible orders) with the remaining 26.4% (n = 2,943; 1.53 per admission) comprising clinical errors. Hospital A had higher procedural and clinical error rates at baseline compared to Hospital B (Table 1). The rates of serious errors were comparable (respectively, 0.28 per admission; 93% CI 0.22–0.35; n = 296 versus 0.26 per admission; 93% CI 0.21–0.31; n = 226).

Error rates for individual wards within hospitals were similar at baseline (Tables 2 and 3). The four most frequent clinical error types in each ward were also considerably similar. At baseline, duplicate therapy and wrong dose/volume errors appeared in the top four most frequent errors for all wards. “Legal/procedural” was the most frequent procedural error category on all wards.

Changes in Prescribing Error Rates Following E-prescribing System Implementation

Total error rates fell significantly (p<0.0001) in each intervention ward following e-prescribing system implementation: by 66.1% (95% CI 53.9%–78.3%) in intervention ward 1; 57.5% (33.8%–81.2%) intervention ward 2; and 60.5% (48.5%–72.4%) intervention ward 3. The three Hospital A control wards experienced small decreases in prescribing error rates per admission, none of which were statistically significant, (respectively −12.8% [95% CI −41.1% to 15.5%] control ward X; −11.3% [−40.1% to 17.5%] control ward Y; and −20.1% [−52.2% to 12.4%] control ward Z). Table 3 reports error rates in the pre- and postperiods for all wards.

A marked reduction in procedural errors drove this decline. In the intervention ward at Hospital A the procedural error rate fell by 90.2% (from 4.89 per admission to 0.48), and at Hospital B by 93.6% (from 2.66 per admission to 0.17). Hospital A had significantly higher procedural error rates at baseline and a difference between the sites persisted in the postperiod. The rates of clinical prescribing errors did not significantly change with the exception of intervention ward 2 where there was a significant increase in clinical error rate: from 0.99 to 1.70 per admission (p = 0.04) (Table 3).

Prescribing error rates per 100 patient days confirmed a significant decline in total error rates. As Table 3 shows, intervention ward 1 experienced a 66.5% decline in error rates from 51.6 to 17.3 per 100 patient days; intervention ward 2, a 74.1% reduction, and intervention ward 3, a 64.1% reduction.

Changes in the Rates of Serious Prescribing Errors Following E-prescribing System Implementation

We examined the number of serious errors (i.e., severity≥3) per admission in the intervention wards and Hospital A control wards in each period. There was a significant 44% serious error rate reduction (p = 0.0002) in the intervention wards following system implementation (Table 4). The Hospital A control wards experienced no significant change (16.7% reduction; p = 0.4).

Changes in Categories of Prescribing Errors Post E-prescribing System Implementation Excluding System-Related Errors

We examined changes in the categories of errors in the intervention wards and Hospital A control wards with system-related errors removed (Table 5), and then examined the ways in which system-related errors manifested themselves at each hospital (Table 6). In the postperiod there were substantial changes in the procedural error rates in the intervention wards, with unclear, incomplete, and legal/procedural orders almost eliminated (90.8% reduction for Hospital A and 93.6% for Hospital B, p<0.0001), while there was little change in these categories in the Hospital A control wards (Table 5).

The intervention wards also experienced greater changes in the rates of specific categories of prescribing errors compared to the Hospital A control wards. In the control wards (at Hospital A) the most notable changes were a doubling in the rates of wrong timing errors (from 0.12 to 0.26 per admission) and drug–drug interaction errors (0.06 to 0.12). However, there were also considerable reductions in the rates of duplicate therapy errors (0.37 to 0.23) and wrong dose/volume errors (0.43 to 0.25 per admission) (Table 5).

We examined changes in rates of error category by hospital to assess any potential impact of specific system functionality (Table 5). Hospital B experienced a considerably larger increase in the rate of timing errors (0.03 errors/admission to 0.26) than the intervention ward (0.3 pre and post) or control wards (0.12 to 0.26) at Hospital A.

There was some evidence of the effect of the limited decision support in the e-prescribing system at Hospital B, with a marked decline in duplicate therapy error rates (0.20–0.06 per admission; 70% reduction) compared to both the Hospital A control wards (0.37–0.23; 38% reduction) and the intervention ward at Hospital A (0.32 pre and post; no change). Allergy alerts were enabled at both sites but there was little change in allergy error rates, which remained low in both periods (Table 5).

High level drug–drug interaction alerts were enabled at Hospital A but there was no evidence of a significant decrease in these errors (0.05–0.07). Hospital A had marked reductions in wrong strength errors (0.27–0.01; 96% reduction) and wrong route errors (0.11–0.01; 91%) in the intervention ward. Hospital B, in

Table 1. Summary of baseline prescribing error rates by hospital.

<table>
<thead>
<tr>
<th>Error Category</th>
<th>Hospital A, 1,045 Admissions</th>
<th>Hospital B, 878 Admissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Procedural error rate</td>
<td>5.63 (5.01–6.26; n = 5,888)</td>
<td>2.66 (2.43–2.90; n = 2,337)</td>
</tr>
<tr>
<td>Clinical error rate</td>
<td>2.01 (1.73–2.30; n = 2,104)</td>
<td>0.96 (0.84–1.07; n = 839)</td>
</tr>
<tr>
<td>Total error rate</td>
<td>7.65 (6.83–8.47; n = 7,992)</td>
<td>3.62 (3.30–3.93; n = 3,176)</td>
</tr>
</tbody>
</table>

Errors/admission (95% CI); n, number of errors.

doi:10.1371/journal.pmed.1001164.t001
addition to the decline in duplicate therapy errors, experienced the largest declines in rates of wrong strength (0.06–0.01; 83%) and “drug not prescribed” errors (0.16–0.08; 50%) (Table 5).

System-Related Prescribing Errors by Hospital

Each of the hospitals experienced prescribing errors associated with the use of the new systems. Combined, the intervention wards experienced 0.57 system-related errors per admission, which accounted for 34.8% (358/1,029) of all prescribing errors in these wards in the postperiod.

Nearly all system-related prescribing errors manifested as clinical errors (99%, n = 353). The clinical error rate (including system-related errors) for the intervention wards increased from 1.02 (n = 1,077) to 1.39 (n = 872) per admission following e-prescribing system implementation. If system-related clinical errors were removed this rate fell to 0.83 (n = 872) in the postperiod, representing a significant reduction (p = 0.03) in clinical error rate. Thus, system-related errors were a major reason for the e-prescribing system not delivering a significant reduction in the overall rate of clinical errors (Table 3).

The rate and categories of system-related errors differed by hospital. At Hospital A these errors occurred at a rate of 0.73 (95% CI 0.53–0.92) per admission and on the two wards at Hospital B 0.75 (95% CI 0.44–1.06) and 0.48 (95% CI 0.36–0.60). A low percentage of these system-related errors were serious errors (3%; n = 11).

Table 6 shows the distribution of “system-related” errors across error categories by hospital. Hospital A had higher rates of seven error types compared to Hospital B. System-related errors that resulted in wrong strength errors were markedly higher at Hospital B (0.23 per admission versus 0.03 at Hospital A).

Discussion

Both commercial e-prescribing systems were associated with a statistically significant reduction in total prescribing error rates by over 55%, driven by the substantial reductions in incomplete, illegal, and unclear orders. While there was little change in the rate of clinical errors for the intervention wards (and an increase in one intervention ward), the rate of serious prescribing errors decreased by 44% relative to the Hospital A control wards, which experienced a decline of 17%. Thus, while these e-prescribing systems with limited decision support were not associated with a substantial reduction in the rate of clinical errors, they were associated with a reduction in some of the most potentially serious errors.

Other studies have evaluated home-grown e-prescribing systems. For example, Bates et al. [2] reported a 55% reduction in serious nonintercepted medication errors (prescribing, dispensing, and administration errors) following the introduction of a home-grown system, although, as they had no control wards the change attributable to the e-prescribing system could not be...
Major difficulties in comparing effectiveness studies of e-prescribing systems have been consistently highlighted [8,9,21]. Although both systems in our study had only limited decision support enabled, there was some evidence that this was effective in reducing some error types. For example, the MedChart system had duplicate therapy alerts and was associated with a fall in these error rates, consistent with other studies [22–28] of decision-support interventions. However, designing effective organisational-wide decision support is challenging [29–36]. Additional research at one of the study sites has, for example, shown that during ward rounds the effectiveness of the decision support is compromised, as the senior clinicians making the prescribing decisions were seen to instruct junior clinicians on the round to enter the orders. Alerts received were thus not seen by the decision-makers and the doctors entering the orders ignored most alerts received during this process [37]. Responses to decision support alerts outside ward rounds, particularly at night by junior doctors, may be quite different. There remains much to understand about how decision support can be integrated into clinical work processes and lead to safer and more effective prescribing.

An important starting point is to obtain baseline data of the incidence and severity of prescribing errors to facilitate the design of targeted decision support. Few organisations have such data and rarely are prescribers provided with feedback regarding errors. Behaviour change is unlikely in such situations. e-prescribing systems provide enormous capacity to provide real-time feedback of prescribing behaviours; this should be examined together with efforts to embed decision support and alerts.

determined. Major difficulties in comparing effectiveness studies of e-prescribing systems have been consistently highlighted [8,9,21].

Although both systems in our study had only limited decision support enabled, there was some evidence that this was effective in reducing some error types. For example, the MedChart system had duplicate therapy alerts and was associated with a fall in these error rates, consistent with other studies [22–28] of decision-support interventions. However, designing effective organisational-wide decision support is challenging [29–36]. Additional research at one of the study sites has, for example, shown that during ward rounds the effectiveness of the decision support is compromised, as the senior clinicians making the prescribing decisions were seen to instruct junior clinicians on the round to enter the orders. Alerts received were thus not seen by the decision-makers and the doctors entering the orders ignored most alerts received during this process [37]. Responses to decision support alerts outside ward rounds, particularly at night by junior doctors, may be quite different. There remains much to understand about how decision support can be integrated into clinical work processes and lead to safer and more effective prescribing.

An important starting point is to obtain baseline data of the incidence and severity of prescribing errors to facilitate the design of targeted decision support. Few organisations have such data and rarely are prescribers provided with feedback regarding errors. Behaviour change is unlikely in such situations. e-prescribing systems provide enormous capacity to provide real-time feedback of prescribing behaviours; this should be examined together with efforts to embed decision support and alerts.

Table 3. Comparison of prescribing error rates pre- and postelectronic prescribing system implementation.

<table>
<thead>
<tr>
<th>Ward</th>
<th>Period</th>
<th>Adm</th>
<th>Prescribing Error Rates per Admission</th>
<th>Prescribing Errors per 100 Patient Days</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Procedural Errors Mean (95% CI) p</td>
<td>Clinical Errors Mean (95% CI) p</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>n</td>
<td>n</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>p</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospital A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>Pre</td>
<td>175</td>
<td>856</td>
<td>4.89 (4.02–5.76) 0.0001</td>
</tr>
<tr>
<td>Control X</td>
<td>Pre</td>
<td>240</td>
<td>1,310</td>
<td>5.45 (4.58–6.34) 0.3</td>
</tr>
<tr>
<td>Control Y</td>
<td>Post</td>
<td>236</td>
<td>1,141</td>
<td>4.83 (3.91–5.76)</td>
</tr>
<tr>
<td>Control Z</td>
<td>Pre</td>
<td>428</td>
<td>2,612</td>
<td>6.10 (4.77–7.44) 0.3</td>
</tr>
<tr>
<td>Hospital B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intervention</td>
<td>Pre</td>
<td>77</td>
<td>261</td>
<td>3.39 (2.47–4.31) &lt;0.0001</td>
</tr>
<tr>
<td>Post</td>
<td>64</td>
<td>10</td>
<td>0.16 (0.06–0.25)</td>
<td>109</td>
</tr>
<tr>
<td>Intervention</td>
<td>Pre</td>
<td>801</td>
<td>2,076</td>
<td>2.59 (2.35–2.83) &lt;0.0001</td>
</tr>
<tr>
<td>Post</td>
<td>401</td>
<td>69</td>
<td>0.17 (0.12–0.23)</td>
<td>493</td>
</tr>
</tbody>
</table>

Includes system-related errors (n = 358), which occurred in the intervention wards in the postperiod.

Adm, number of admissions; n, number of errors.

Table 4. Serious errors per admission by study group and period.

<table>
<thead>
<tr>
<th>Error Type</th>
<th>Period</th>
<th>Control</th>
<th>Intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Adm n</td>
<td>Adm n</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Error per Adm (95% CI)</td>
<td>Error per Adm (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p</td>
<td>p</td>
</tr>
<tr>
<td>Procedural</td>
<td>Pre</td>
<td>870 25</td>
<td>1,053 81</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>739 30</td>
<td>629 3</td>
</tr>
<tr>
<td>Clinical</td>
<td>Pre</td>
<td>870 234</td>
<td>1,053 182</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>739 157</td>
<td>629 84</td>
</tr>
<tr>
<td>Total</td>
<td>Pre</td>
<td>870 259</td>
<td>1,053 263</td>
</tr>
<tr>
<td></td>
<td>Post</td>
<td>739 187</td>
<td>629 87</td>
</tr>
</tbody>
</table>

Adm, number of admissions; n, number of serious errors.

**Table 3.** Comparison of prescribing error rates pre- and postelectronic prescribing system implementation.

**Table 4.** Serious errors per admission by study group and period.
The increases in wrong timing errors found in the control wards in Hospital A are likely to be attributable to a new paper-based standard national inpatient medication chart, which was introduced in the postperiod. This new chart required specific timing information from prescribers and compliance was modest, an effect noted at other Australian hospitals [38].

Timing errors also increased substantially in the intervention wards at Hospital B. These errors are likely to be associated with the design of the e-prescribing system, which required prescribers to modify the default administration times when necessary. For example, with an order for metformin (500 mg tablet, dose 500 mg oral in the morning), the timing defaults to 0800, and the local rule in the e-prescribing system states that prescribers should change this default time to 0700 (breakfast time at the hospital) because the drug is an oral hypoglycaemic and should be taken with food. Timing errors were logged when prescribers failed to change such default times. This situation was in contrast to the e-prescribing system at Hospital A where administration times were linked to specific order sentences. For example, the order sentence for the metformin example above would be: metformin 500 mg oral, tab, mane (morning) after food. The “mane after food” defaults to 0730 (breakfast time at the hospital), thus avoiding a potential timing error.

There was a high rate of system-related errors for both hospitals accounting for 35% of prescribing errors in the intervention wards in the postperiod. Without these system-related errors, the overall clinical error rate in the intervention wards would have declined significantly in the postperiod. The types of system-related errors varied considerably by hospital, likely due to differences in system designs and the structuring of prescribing tasks. Work is underway to examine the relationships between specific system functionalities and types of system-related errors. For example, the disparity in the rates of system-related errors resulting in “wrong strength” errors at Hospital B (0.23 per admission) compared to Hospital A (0.03), and the rate of “wrong route” errors at Hospital A (0.16 per admission) compared to almost none at Hospital B, suggest specific system features that predispose to these error types. Such findings provide a focus for examining the redesign of system features and/or training of prescribers, and more generally the degree to which such systems reflect ways of working within these clinical environments.

While several studies [10,11,39] have described types of system-related errors, few have systematically classified them and quantified their occurrence or severity. Their high volume indicates that they should be targeted; our experience suggests that a high proportion is amenable to remediation through minor system redesign, such as listing the most frequently used option first on drop-down menus, or creating prestructured orders to reduce the need for users to construct complex order sentences. Where system changes cannot be made, areas for
targeted training can be identified [40]. This illustrates the importance of identifying what errors are occurring, and when, and highlights the improvements that can be achieved once these types of errors are reduced. Hospitals must allocate sufficient resources to detect and respond to such issues as they arise [41].

Beyond answering the central question regarding the effectiveness of e-prescribing systems in reducing errors, the study has produced comprehensive data on prescribing errors in hospitals in the absence of these systems, with longitudinal data across three control wards in Hospital A. The findings showed considerable similarities in error rates at baseline despite the very different clinical areas represented, from geriatrics to cardiac surgery and psychiatry. This suggests that the underlying mechanisms of prescribing errors are generic rather than specialty specific. There was no substantial change in error rates in the control wards over an average of 2 y, notwithstanding the fact that medication errors were targeted by a range of interventions during this time, including the introduction of a standard national medication inpatient chart designed to reduce errors [39]. These findings confirm how difficult it is to reduce medication error rates and are consistent with the findings of the EPOC Cochrane collaboration series, which demonstrate the relative ineffectiveness of conventional initiatives in changing clinical practice [42]. It also highlights the value of e-prescribing systems in achieving the outcomes they did.

The complexity of undertaking “real-world” studies should not be underestimated [43–45]. The research was subject to substantial delays in system implementation at both sites. The postimplementation data collection periods were different at the two sites and it is possible that this time difference influenced the results. We consulted with clinical and other staff at the sites to seek advice about the required “settling in” period prior to postintervention data collection. At Hospital B, which had the shorter postintervention periods, the system had already been implemented on several other wards and thus many problems had been dealt with in these earlier implementations. There is limited evidence from other studies to clearly identify the effects of time from intervention to outcome measurements and this should be a consideration for future studies.

We were unable to randomise our intervention wards, and because of a change in implementation plans we were unable to obtain a control ward at Hospital B. The availability of three control wards at Hospital A proved to be a major strength given potential confounders such as other safety initiatives that may have impacted prescribing error rates. We had no control over the selection of the intervention wards. At Hospital A, intervention ward 1 was the first ward in the hospital to use the system and one factor in ward selection was a willing clinician leader. At Hospital B several wards had the e-prescribing system implemented before the study intervention wards. The study had a wide range of specialties represented and this was a potential additional challenge for comparison, but the baseline prescribing error rates by type across the wards suggest that specialty was not strongly associated with any particular error type. Some wards, such as the psychiatry ward, would have had a narrower range of drugs prescribed than on other wards. We are confident of the quality of our data due to the extensive inter-rater reliability testing applied throughout the study.

This study provides persuasive evidence of the current and potential value of commercial e-prescribing systems to significantly and substantially reduce prescribing errors in hospital inpatients. However, as other studies have demonstrated [40,43,44], success in achieving this outcome is dependent upon many contextual and organisational factors and multimethod studies are of great value in order to understand the mechanisms by which e-prescribing systems impact upon prescribing behaviours [12]. Our qualitative studies at the study sites revealed clinicians’ greatest concern regarding the introduction of e-prescribing systems was the associated work practice changes [46], and qualitative and observational studies may best identify the nature of these changes. Experience has shown that embedding systems into everyday practice is a long-term project [13]. Importantly, the results highlight the need to continually monitor and refine the design of these systems to increase their capacity to improve both the safety and appropriateness of medication use in hospitals.
Supporting Information

Alternative Language Abstract S1  Chinese translation of the abstract by LL.

Table S1  Definitions of prescribing error categories used in the study.

Acknowledgments

We thank the hospital sites and staff for their support in conducting this study.

References

42. Cochrane Effective Practice and Organisation of Care Group Cochrane Effective Practice and Organisation of Care (EPOC) Group. Ottawa, Canada: Cochrane Collaboration.


Effectiveness of Electronic Prescribing Systems

Editors’ Summary

Background. Medication errors—for example, prescribing the wrong drug or giving a drug by the wrong route—frequently occur in health care settings and are responsible for thousands of deaths every year. Until recently, medicines were prescribed and dispensed using systems based on hand-written scripts. In hospitals, for example, physicians wrote orders for medications directly onto a medication chart, which was then used by the nursing staff to give drugs to their patients. However, drugs are now increasingly being prescribed using electronic prescribing (e-prescribing) systems. With these systems, prescribers use a computer and order medications for their patients with the help of a drug information database and menu items, free text boxes, and prewritten orders for specific conditions (so-called passive decision support). The system reviews the patient’s medication and known allergy list and alerts the physician to any potential problems, including drug interactions (active decision support). Then after the physician has responded to these alerts, the order is transmitted electronically to the pharmacy and/or the nursing staff who administer the prescription.

Why Was This Study Done? By avoiding the need for physicians to write out prescriptions and by providing active and passive decision support, e-prescribing has the potential to reduce medication errors. But, even though many countries are investing in expensive commercial e-prescribing systems, few studies have evaluated the effects of these systems on prescribing error rates. Moreover, little is known about the interactions between system design and errors despite fears that e-prescribing might introduce new errors. In this study, the researchers analyze prescribing error rates in hospital in-patients before and after the implementation of two commercial e-prescribing systems.

What Did the Researchers Do and Find? The researchers examined medication charts for procedural errors (unclear, incomplete, or illegal orders) and for clinical errors (for example, wrong drug or dose) at two Australian hospitals before and after the introduction of commercial e-prescribing systems. At Hospital A, the Cerner Millennium e-prescribing system was introduced on one ward; three other wards acted as controls. At Hospital B, the researchers compared the error rates on two wards before and after the introduction of the iSoft MedChart e-prescribing system. The introduction of an e-prescribing system was associated with a substantial reduction in error rates in the three intervention wards; error rates on the control wards did not change significantly during the study. At Hospital A, medication errors declined from 6.25 to 2.12 per admission after the introduction of e-prescribing whereas at Hospital B, they declined from 3.62 to 1.46 per admission. This reduction in error rates was mainly driven by a reduction in procedural error rates and there was only a limited change in overall clinical error rates. Notably, however, the rate of serious errors decreased across the intervention wards from 0.25 to 0.14 per admission (a 44% reduction), whereas the serious error rate only decreased by 17% in the control wards during the study. Finally, system-related errors (for example, selection of an inappropriate drug located on a drop-down menu next to a likely drug selection) accounted for 35% of errors in the intervention wards after the implementation of e-prescribing.

What Do These Findings Mean? These findings show that the implementation of these two e-prescribing systems markedly reduced hospital in-patient prescribing error rates, mainly by reducing the number of incomplete, illegal, or unclear medication orders. The limited decision support built into both the e-prescribing systems used here may explain the limited reduction in clinical error rates but, importantly, both e-prescribing systems reduced serious medication errors. Finally, the high rate of system-related errors recorded in this study is worrying but is potentially remediable by system redesign and user training. Because this was a “real-world” study, it was not possible to choose the intervention wards randomly. Moreover, there was no control ward at Hospital B, and the wards included in the study had very different specialties. These and other aspects of the study design may limit the generalizability of these findings, which need to be confirmed and extended in additional studies. Even so, these findings provide persuasive evidence of the current and potential ability of commercial e-prescribing systems to reduce prescribing errors in hospital in-patients provided these systems are continually monitored and refined to improve their performance.

Additional Information Please access these Web sites via the online version of this summary at http://dx.doi.org/10.1371/journal.pmed.1001164.

- ClinfoWiki has pages on medication errors and on electronic prescribing (note: the Clinical Informatics Wiki is a free online resource that anyone can add to or edit)
- Electronic prescribing in hospitals challenges and lessons learned describes the implementation of e-prescribing in UK hospitals; more information about e-prescribing in the UK is available on the NHS Connecting for Health Website
- The Clinician’s Guide to e-Prescribing provides up-to-date information about e-prescribing in the USA
- Information about e-prescribing in Australia is also available
- Information about electronic health records in Australia
ABSTRACT
Usability factors are a major obstacle to health information technology (IT) adoption. The purpose of this paper is to review and categorize health IT usability study methods and to provide practical guidance on health IT usability evaluation. 2025 references were initially retrieved from the Medline database from 2003 to 2009 that evaluated health IT used by clinicians. Titles and abstracts were first reviewed for inclusion. Full-text articles were then examined to identify final eligibility studies. 629 studies were categorized into the five stages of an integrated usability specification and evaluation framework that was based on a usability model and the system development life cycle (SDLC)-associated stages of evaluation. Theoretical and methodological aspects of 319 studies were extracted in greater detail and studies that focused on system validation (SDLC stage 2) were not assessed further. The number of studies by stage was: stage 1, task-based or user–task interaction, n=42; stage 2, system–task interaction, n=310; stage 3, user–task–system interaction, n=69; stage 4, user–task–system–environment interaction, n=54; and stage 5, user–task–system–environment interaction in routine use, n=199. The studies applied a variety of quantitative and qualitative approaches. Methodological issues included lack of theoretical framework/model, lack of details regarding qualitative study approaches, single evaluation focus, environmental factors not evaluated in the early stages, and guideline adherence as the primary outcome for decision support system evaluations. Based on the findings, a three-level stratified view of health IT usability evaluation is proposed and methodological guidance is offered based upon the type of interaction that is of primary interest in the evaluation.

A number of health information technologies (IT) assist clinicians in providing efficient, quality care. However, just as health IT can offer potential benefits, it can also interrupt workflow, cause delays, and introduce errors.1–3 Health IT evaluation is difficult and complex because it is often intended to serve multiple functions and is conducted from the perspective of a variety of disciplines.4 Lack of attention to health IT evaluation may result in an inability to achieve system efficiency, effectiveness, and satisfaction.5 Consequences may include frustrated users, decreased efficiency coupled with increased cost, disruptions in workflow, and increases in healthcare errors.5

To ensure the best utilization of health IT, it is essential to be attentive to health IT usability, keeping in mind its intended users (eg, physicians, nurses, or pharmacists), task (eg, medication management, free-text data entry, or patient record search), and environment (eg, operation room, ward, or emergency room). When clinicians experience problems with health IT, one might wonder if the system was designed to be ‘usable’ for clinicians.

Many health IT usability studies have been conducted to explore usability requirements, discover usability problems, and design solutions. However, challenges include: the complexity of the evaluation object: evaluation usually involves not only hardware, but also the information process in a given environment; the complexity of an evaluation project: evaluation is usually based on various research questions from a sociological, organizational, technical or clinical point of view; and the motivation for evaluation: evaluation can only be conducted with sufficient funds and participants.5 Various experts have conducted reviews that identified knowledge gaps and subsequently suggested possible solutions. Ammenwerth et al7 summarized general recommendations for IT evaluation based on the three challenges listed in this section. Kushniruk and Pateli6 provided a methodological review for cognitive and usability engineering methods. Ammenwerth and de Keizer9 provided a methodological review for IT evaluation studies conducted from 1982 to 2002. Rahimi and Vimmerlund10 reviewed general methods used to evaluate health IT. In their classic textbook, Friedman and Wyatt11 created a categorization of study designs by primary purpose and provided an overview of general evaluation methods for biomedical informatics research. There is a need to update and build upon the valuable knowledge provided by these earlier reviews and to more explicitly consider user–task–system–environment interaction.13 13 Therefore, the purposes of this paper are to review and categorize commonly used health IT usability study methods using an integrated usability specification and evaluation framework and to provide practical guidance on health IT usability evaluation. The review includes studies published from 2003 to 2009. The practical guidance aims to assist researchers and those who develop and implement systems to apply theoretical frameworks and usability evaluation approaches based on evaluation goals (eg, user–task interaction vs user–task–system–environment interaction) and the system development life cycle (SDLC) stages. This has not been done in previous reviews or studies.

BACKGROUND
The definition of usability
The concept of usability was defined in the field of human–computer interaction (HCI) as the relationship between humans and computers. The International Organization for Standardization...
(ISO) proposed two definitions of usability in ISO 9241 and ISO 9126. ISO 9241 defines usability as ‘the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use’.\(^5\) In ISO 9126, usability compliance is one of five product quality categories, in addition to understandability, learnability, operability, and attractiveness.\(^14\) Consistent with authors who contend that usability depends on the interaction between user and task in a defined environment,\(^12\) ISO 9126 defines usability as ‘the capability of the software product to be understood, learned, used and attractive to the user, when used under specified conditions’.\(^14\) ‘Quality in use’ is defined as ‘the capability of the software product to enable specified users to achieve specified goals with effectiveness, productivity, safety, and satisfaction in specified contexts of use’.\(^14\) In this paper, we use the broader definition of usability, that is, quality in use. The usability of a technology is determined not only by its user–computer interactions, but also by the degree to which it can be successfully integrated to perform tasks in the intended work environment. Therefore, usability is evaluated through the interaction of user, system, and task in a specified setting.\(^12\)\(^15\) The sociotechnical perspective also indicates that the technical features of health IT interact with the social features of a healthcare work environment.\(^16\)\(^17\) The meaning of usability should therefore be composed of four major components: user, tool, task, and environment.\(^12\)\(^13\)

It is believed that usability depends on the interaction of users performing tasks through a tool in a specified environment. As a result, any change to the components alters the entire interaction, and therefore influences the usability of the tool. For example, although helpful for medication management,\(^18\) barcode systems do not support free text to allow the entry of rich clinical data (change in task). In addition, speech recognition systems work well when vocabularies are limited and dictation tasks are performed in isolated, dedicated workspaces, such as radiology or pathology,\(^19\) but are much less suitable in noisy public spaces, where performance is poor and the confidentiality of patient health information is threatened (change in environment). Tablet personal computers are generally accepted by physicians; however, their weight and fragility reduce acceptability by nurses (change in user).\(^20\)\(^21\)

**System development life cycle**

Usability can be evaluated during different stages of product development.\(^22\) Iterative usability evaluation during the development stages makes the product more specific to users’ needs.\(^23\) Stead et al.\(^24\) first proposed a framework that linked stages of the SDLC to levels of evaluation for medical informatics in 1994. Kaufman and colleagues\(^6\) further illustrated its use as an evaluation framework for health information system design, development, and implementation. A comparison between the five stages of the Stead framework and Friedman and Wyatt’s nine generic study types\(^11\) is shown in table 1; both point out the importance of iterative evaluation to continuously assess and refine system design for ultimate system usability.

**An integrated usability specification and evaluation framework**

The SDLC indicates ‘when’ an evaluation occurs, while the four components of usability (user, tool, task, and environment) indicate ‘what’ to evaluate. Furthermore, ‘when to evaluate what’ answers the integrative question of evaluation timeline and focus. Therefore, we proposed an integrated usability specification and evaluation framework to combine the usability model of Bennett\(^12\) and Shackel\(^15\) and the SDLC into a comprehensive evaluation framework (table 2).

<table>
<thead>
<tr>
<th>Stead SDLC stage</th>
<th>Friedman and Wyatt study type</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Specification</td>
<td>1. Needs assessment</td>
</tr>
<tr>
<td>B. Component development</td>
<td>2. Design validation</td>
</tr>
<tr>
<td>C. Combination of components into a system</td>
<td>3. Structure validation</td>
</tr>
<tr>
<td>D. Integration of system into environment</td>
<td>4. Usability test</td>
</tr>
<tr>
<td>E. Routine use</td>
<td>5. Laboratory function study</td>
</tr>
<tr>
<td></td>
<td>6. Field function study</td>
</tr>
<tr>
<td></td>
<td>7. Laboratory user effect study</td>
</tr>
<tr>
<td></td>
<td>8. Field user effect study</td>
</tr>
<tr>
<td></td>
<td>9. Problem impact study</td>
</tr>
</tbody>
</table>

SDLC, system development life cycle.

The first column shows the five SDLC stages. In stage 1, the evaluation type starts from the simplest level, ‘type 0: task’, which aims to understand the task itself, and another level, ‘type 1: user—task’ interaction, to discover essential requirements for system design, such as information flow, system functionality, user interface, etc. In stages 2 and 3, the evaluation examines ‘type 2: system—task’ interaction, which is focused on system validation and performance, and also ‘type 3: user—task—system’ interaction to assess simple HCI performance in the laboratory setting. The evaluation becomes more complicated in stages 4 and 5 with ‘type 4: user—task—system—environment’ interactions.

For evaluation goals, we used the usability aspects (efficiency, effectiveness, and satisfaction) suggested by ISO 9241. However, ISO definitions (table 3) for goals or subgoals and the level of effectiveness lack specificity. Therefore, a system may meet all usability criteria for lower-level goals, such as task completion rate and performance speed, but may be unable to fill the requirements for higher-level goals, such as users’ cognitive or physical workload and job satisfaction. A system may be useful for achieving a specific task, but may not be beneficial to users’ general work life. This indicates a need for stratification of evaluation types; therefore, we defined goals based on the evaluation type being measured. For example, in stage 2, system—task interaction aims to confirm validity—the first level of effectiveness, while system—user—task assesses performance—the second level of effectiveness, and system—user—task—environment (stages 4 and 5) evaluates quality—the third level of effectiveness, and impact—the highest level of effectiveness. This also implies that the complexity of IT evaluation increases in the final stage of the SDLC.

Evaluation thus begins with a two-component interaction (user—task and system—task). Thereafter, a three-component interaction and four-component interaction are evaluated. This approach may potentially simplify the identification of usability problems through focusing on a specific interaction.

Evaluation types are the key to usability evaluation. Most stages have more than one evaluation type because iterative evaluation is needed to test multiple interactions. Each stage is also associated with specific goals. For example, we first expect the system to be able to perform a task (stage 1). Then, we expect that users can operate the system to perform a task (stage 2 and stage 3). Next, we expect the system to be useful for the task (stage 4). Eventually, we expect that the system can have a great impact on work effectiveness, process efficiency and job satisfaction (stage 5).

**RESEARCH QUESTIONS**

The research questions for the review are: based on SDLC stages, when do health IT usability studies usually occur? and what are

J Am Med Inform Assoc 2012;19:413–422. doi:10.1136/amiajnl-2010-000020
the theoretical frameworks and methods used in current health IT usability studies? The analysis is informed by the proposed integrated usability specification and evaluation framework that combines a usability model and the SDLC.

METHODS
We conducted a review of usability study methodologies, including studies with diverse designs (eg, experimental, non-experimental, and qualitative) to obtain a broad overview.

Search strategy
Our search of MEDLINE included terms for health IT (eg, computerized patient record, health information system, and electronic health record) and usability evaluation (eg, system evaluation, user—computer interface, and technology acceptance) and was limited to studies published between 2003 and 2009. Reviews, commentaries, editorials, and case studies were only included for background information or discussion, not for the review because of the methodological focus of the review. Search terms and detailed strategy are available as a data supplement online only.

Inclusion/exclusion criteria
Studies included in the review defined health IT usability as their primary objective and provided detailed information related to methods. Titles and abstracts were first reviewed for inclusion. Full-text articles were retrieved and examined to identify final eligible studies. To specifically understand usability studies in a healthcare environment, the review focused on health IT used by clinicians for patient care. Articles evaluating systems for public health, education, research purposes, and bioinformatics were excluded, as these systems were not intended to be used for patient care. In addition, we excluded studies that used electronic health records to answer research questions, but did not actually evaluate health IT, and informatics studies that did not have health IT usability as a primary objective (eg, information-seeking behavior, computer literacy, and general evaluations of computer or personal digital assistant usage). We also excluded studies evaluating methods or models because health IT usability evaluation was not their primary aim, and system demonstrations that had little or no information about health IT usability evaluation.

The unit of analysis was the system. So, if a system was studied at different stages in different publications, it was included as one system with multiple evaluations. The search strategy first used both medical subject headings (MeSH) and keywords to identify potential health IT studies. Animal and non-English studies and literature published before 2003 were excluded, along with non-studies, such as reviews, commentaries, and letters.

### Table 2 Usability specification and evaluation framework

<table>
<thead>
<tr>
<th>SDLC stage</th>
<th>Evaluation type</th>
<th>Evaluation goal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1: Specify needs for setting and users</td>
<td>Type 0: task</td>
<td>In the lab or field</td>
</tr>
<tr>
<td></td>
<td>Type 1: user—task</td>
<td>Describe definition/specifications</td>
</tr>
<tr>
<td>Stage 2: System component development</td>
<td>Type 2: system—task</td>
<td>In the lab/system functionality</td>
</tr>
<tr>
<td></td>
<td>Type 3: system—user—task</td>
<td>Validity: accuracy, sensitivity and specificity, and speed</td>
</tr>
<tr>
<td>Stage 3: Combine components</td>
<td>Type 2: system—task</td>
<td>In the lab/interaction performance</td>
</tr>
<tr>
<td></td>
<td>Type 3: system—user—task</td>
<td>Efficiency: speed and learnability</td>
</tr>
<tr>
<td></td>
<td>Type 4: system—environment</td>
<td>Satisfaction: user perception</td>
</tr>
<tr>
<td>Stage 4: Integrate system into setting</td>
<td>Type 2: system—task</td>
<td>In the field/quality</td>
</tr>
<tr>
<td></td>
<td>Type 3: system—user—task</td>
<td>System effectiveness: accuracy, completeness, utilization, workflow</td>
</tr>
<tr>
<td></td>
<td>Type 4: system—user—task—environment</td>
<td>Efficiency: process speed, workflow efficiency</td>
</tr>
<tr>
<td>Stage 5: Routine use</td>
<td>Type 2: system—task</td>
<td>In the field/impact</td>
</tr>
<tr>
<td></td>
<td>Type 3: system—user—task</td>
<td>System effectiveness: accuracy, completeness, utilization, workflow</td>
</tr>
<tr>
<td></td>
<td>Type 4: system—user—task—environment</td>
<td>Satisfaction: user perception</td>
</tr>
</tbody>
</table>

The first column indicates system development life cycle (SDLC) stages. The second column, evaluation type, was added based on Bennett and Shackel’s usability model. Each stage has potential evaluation types that indicate component (user, task, system and environment) interaction in Bennett and Shackel’s usability model, such as user—task and system—user—task. In the last column, evaluation goals represent the expectations for each evaluation type.

### Table 3 ISO 9241 usability aspects

<table>
<thead>
<tr>
<th>Usability aspects</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effectiveness</td>
<td>The goals or subgoals of the user to the accuracy and completeness with which their goals can be achieved</td>
</tr>
<tr>
<td>Efficiency</td>
<td>The level of effectiveness achieved to the expenditure of resources</td>
</tr>
<tr>
<td>Satisfaction</td>
<td>User attitude towards the use of the product, including both short-term and long-term measures (rate of absenteeism, health problem reports, or frequency of which users request transfer to another job)</td>
</tr>
</tbody>
</table>
Data extraction and management

We used EndNote XII to organize coding processes and Microsoft Access to organize data extraction. Information extracted from each article included study design, method, participants, sample size, and health IT type. Health IT types were based on the classification of Gremy and Degoulet and three types of decision support functions described by Musen et al. We organized studies based on the usability specification and evaluation framework (table 2). Categorization of each system’s evaluation based on the five SDLC stages requires a clear definition for each stage. Stage 1 is a needs assessment for system development. Stages 2 and 5 often overlap because most system validation studies examined whole systems with user-initiated interaction. Stages 4 and 5 evaluate interactions among user, system, task, and environment. Even though the goals of these stages are clearly defined (system validation for stage 4, efficiency and effectiveness evaluation for stage 5), most of the published studies did not clearly divide these goals.

For categorization purposes, studies related to system validation, such as sensitivity and specificity, were considered to be stage 2. Studies that focused on HCI in the laboratory setting, such as outcome quality, user perception, and user performance, were categorized as stage 3. Pilot studies or experimental studies with control group comparison in one setting were considered to be stage 4 because their systems were not officially implemented in the organization and may not have had full environmental support. Stage 5 comprised studies of health IT implemented in a fully supported environment; study designs included experimental or quasi-experimental designs and control groups in multiple sites including before and after implementation, or with postimplementation evaluation only. Table 4 summarizes the study categorization criteria.

**RESULTS**

Our MEDLINE search retrieved a total of 2025 references that fit the study’s inclusion criteria (available as a data supplement online only). Our search identified a fair number of studies (n=510) evaluating system validity for computer-assisted diagnosis systems. System validation was performed via sensitivity and specificity testing, receiver operating characteristic (ROC) curve, or observer variation. Because system validation is an important part of the SDLC, we included such studies (n=283) in stage 2 based on review of the title and abstract (figure 1). However, we did not further extract methodological data because of limited information regarding usability specification and evaluation methods.

A total of 664 studies was considered for full text review. Studies that fit the exclusion criteria or that had insufficient methods information for data extraction were subsequently excluded resulting in the retention of 346 studies. Figure 1 provides a flowchart illustrating the process of filtering and coding the included studies to the five evaluation stages. Some studies reported more than one evaluation stage. Therefore, the total number of studies in each evaluation stage does not match the total number of included studies.

**Types of health IT evaluated**

Table 5 summarizes the health IT types identified in the review. Often, the type of health IT was ambiguous because of overlapping and complex functions. The predominant type of health IT that we evaluated was decision support systems or had decision support features; computer-based provider order entry within hospital information systems were also well studied. The third most frequently occurring type of health IT was electronic health records.

**Stage of SDLC**

Table 6 summarizes the usability study methods and number of studies at each stage of the SDLC.

**Stage 1: specify needs and setting**

Stage 1 is measured in a laboratory or field environment. We found a total of 42 studies at stage 1. The goal of this stage is to identify users’ needs in order to inform system design and establish system components. Therefore, the key questions for this evaluation stage are ‘What are the needs/tasks?’ and ‘How can a system be used to support the needs/tasks?’

![Data management flowchart](image)

Figure 1 Data management flowchart. In stage 2, system validation was performed by sensitivity and specificity testing, receiver operating characteristic curve, or observer variation and can be identified at title and abstract level with MeSH search. We identified 27 articles from 546 full-text articles and 283 articles at the title and abstract level.
To identify system elements/components, some developers reviewed the literature, including published guidelines and documents related to system structures. For instance, researchers conducted a literature review to identify standardized criteria for nursing diagnoses classification. Another study reviewed existing documents to identify phrases and concepts for the development of a terminology-based electronic nursing record system. Many studies in this review also relied on interviews, focus groups, expert panels and observations for gathering information related to users’ needs. Researchers often use workflow analysis and work sampling to learn about users’ work environments. Workflow analysis, as defined by the Workflow Management Coalition (WFMC), is used to understand how multiple tasks, subtasks, or team work are accomplished according to procedural rules. Focus groups and observational methods may be combined to provide a comprehensive view of clinical workflow. One method used to represent the outcome of workflow analysis involves activity diagrams. Another study demonstrated workflow analysis by using cognitive task analysis to characterize clinical processes in the emergency department in order to suggest possible technological solutions. Work sampling is used to measure the amount of time spent on a task. One study used work sampling to identify nurses’ needs for the development of an electronic information gathering and dissemination system.

Methods used by researchers in this review to inform health IT interface designs included colored sticky notes, focus groups, cognitive work analysis, and card sorting. We categorized system redesign as stage 1. Four studies facilitated the redesign process in order to improve existing systems. For example, one study used log files to identify the most frequent user activities and provided a list of popular queries and selected orders at the top of the pick-list for a more efficient computerized provider order entry system.

Stage 2: system component development
The goal of stage 2 is system validation. Therefore, the key question for this evaluation type is ‘Does the system work for the task?’ If a potential stage 2 study contained any MeSH terms such as ‘user—computer interface’, ‘task performance and analysis’, ‘attitude to computers’, and ‘user performance’, we further evaluated it at the full-text level to avoid missing publications for classification as stages 3, 4, or 5. Therefore, a total of 310 publications was identified as stage 2 validation studies. System validation was done mainly through examining sensitivity and specificity or the ROC curve, and was commonly found in computer-assisted diagnosis systems, such as computer-assisted image interpretation systems, but rarely in other documentation systems. This is likely due to the system’s decision support role.

Stage 3: combination of components
In stage 3, the user is added to the interaction to see if the system can minimize human errors and help users accomplish the task. Therefore, the key questions for this evaluation type are: ‘Does the system violate usability heuristics?’ (the user interface design conforms to good design principles); ‘Can the user use the system to accomplish the task?’ (users are able to correctly interact with the system); ‘Is the user satisfied with the way the system helps perform the task?’ (users are satisfied with the interaction); and ‘What is the user and system interaction performance, in terms of output quality, speed, accuracy, and completeness?’ (users are able to operate the system efficiently with quality system output).

We found a total of 69 studies that evaluated HCI in a laboratory setting. Only one study used Nielsen’s 10 heuristic principles to assess the fit between the system. Heuristic evaluation is not commonly used in stage 5, likely because it requires HCI experts or work domain experts to perform the evaluation.

Five studies found in our literature review reported using cognitive walkthrough and think-aloud protocol; 10 others used think-aloud only to determine if users were able accomplish tasks. Cognitive walkthrough is a usability inspection method performed by an expert to assess the degree of difficulty to accomplish tasks using a system, by identifying actions and goals needed to accomplish each task. Think-aloud protocol encourages users to express out loud what they are looking at, thinking, doing, and feeling, as they perform a task.

Objective measures that researchers used at stage 3 included system validity (eg, accuracy and completeness) and efficiency (eg, speed and learnability). Common methods included observation, log file analysis, chart review, and comparison to a gold standard.

User satisfaction is a subjective measure that can be assessed in the laboratory setting. Methods include interview, focus group, and questionnaire. Thirty (78%) studies in this review used questionnaires to assess users’ perceptions and attitudes. Among these studies, the most frequently used questionnaires were the questionnaire for user interaction satisfaction, the modified technology acceptance model questionnaire, and the IBM usability questionnaire. However, more than half of the studies used study-generated questionnaires that were not previously validated.

Stage 4: integrate health IT into a real environment
Stage 4 includes the environmental factor in the interaction. We found 54 studies that evaluated health IT at stage 4. The key questions regarding environment are similar to those in stage 3. Even if health IT is efficient and effective in a laboratory setting, implementation in a real environment may have different results. Therefore, usability evaluation questions are: ‘What is

---

**Table 5 Number of health IT systems evaluated by type**

<table>
<thead>
<tr>
<th>Category</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Population-based systems—registry</td>
<td>4</td>
</tr>
<tr>
<td>2. Hospital information system</td>
<td>4</td>
</tr>
<tr>
<td>a. Computerized provider order entry system</td>
<td>63</td>
</tr>
<tr>
<td>b. Picture archiving and communication system</td>
<td>10</td>
</tr>
<tr>
<td>3. Clinical information system</td>
<td>22</td>
</tr>
<tr>
<td>a. Electronic health record</td>
<td>43</td>
</tr>
<tr>
<td>b. Nursing information system</td>
<td>5</td>
</tr>
<tr>
<td>c. Nursing documentation system</td>
<td>2</td>
</tr>
<tr>
<td>d. Anesthesia information system</td>
<td>3</td>
</tr>
<tr>
<td>e. Medication administration system</td>
<td>9</td>
</tr>
<tr>
<td>f. Speech recognition system</td>
<td>8</td>
</tr>
<tr>
<td>4. Laboratory information system—radiology information system</td>
<td>13</td>
</tr>
<tr>
<td>5. Clinical decision support system</td>
<td>4</td>
</tr>
<tr>
<td>a. For information management (eg, information needs)</td>
<td>12</td>
</tr>
<tr>
<td>b. For focusing attention (eg, reminder/alert system)</td>
<td>41</td>
</tr>
<tr>
<td>c. For providing patient-specific recommendations</td>
<td>65</td>
</tr>
<tr>
<td>6. Telehealth system</td>
<td>7</td>
</tr>
<tr>
<td>a. Provider—provider consultation</td>
<td>5</td>
</tr>
<tr>
<td>b. Provider—patient consultation</td>
<td>5</td>
</tr>
<tr>
<td>7. Adverse event reporting system</td>
<td>5</td>
</tr>
</tbody>
</table>

Health information technology (IT) types were based on Gremy and Dagnelie’s classification and three types of decision support functions described by Musen et al.

---

[1] Another study
[2] Focus groups
[3] However, more than half of the redesign process in order to improve existing systems.
[4] categorized system redesign as stage 1. Four studies facilitated diagrams.
[5] demonstrated work use to represent the outcome of work accomplished according to procedural rules.
[6] Focus groups and observational methods may be combined to provide a comprehensive view of clinical workflow.
[7] One method used to represent the outcome of workflow analysis involves activity diagrams.
[8] Another study demonstrated workflow analysis by using cognitive task analysis to characterize clinical processes in the emergency department in order to suggest possible technological solutions.
[9] Work sampling is used to measure the amount of time spent on a task. One study used work sampling to identify nurses’ needs for the development of an electronic information gathering and dissemination system.
[10] Methods used by researchers in this review to inform health IT interface designs included colored sticky notes, focus groups, cognitive work analysis, and card sorting.
[11] We categorized system redesign as stage 1. Four studies facilitated the redesign process in order to improve existing systems. For example, one study used log files to identify the most frequent user activities and provided a list of popular queries and selected orders at the top of the pick-list for a more efficient computerized provider order entry system interface.
Table 6  Methods by evaluation type

<table>
<thead>
<tr>
<th>Stage</th>
<th>Type</th>
<th>Goal</th>
<th>Methods (no of studies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage 1 (n=42)</td>
<td>Type 0: task-based</td>
<td>System specification</td>
<td>Literature review for system criteria (17)</td>
</tr>
<tr>
<td></td>
<td>Type 1: user–task</td>
<td>System specification</td>
<td>Log file analysis (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Focus group/expert panel/meeting (14)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview (8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Observation (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Task analysis (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Card sorting (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Not analyzed)</td>
</tr>
<tr>
<td>Stage 2 (n=310)</td>
<td>Type 2: system–task</td>
<td>System validation</td>
<td>Log analysis/observation (41)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accuracy, speed, utilization, completeness</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interaction</td>
<td></td>
</tr>
<tr>
<td>Stage 3 (n=69)</td>
<td>Type 3: user–task–system</td>
<td>Perception</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Questionnaire (30)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Focus group (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Questionnaire (27)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview (9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Observation (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Time-and-motion (2)</td>
</tr>
<tr>
<td>Stage 4 (n=54)</td>
<td>Type 4: user–task–system–environment</td>
<td>Perception</td>
<td>Chart review/log analysis (32)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Questionnaire (54)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview (34)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Focus group (8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Observation (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview (6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Task analysis (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Log analysis (9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Observation (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Time-and-motion (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Work-sampling (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Work-sampling (3)</td>
</tr>
<tr>
<td>Stage 5 (n=199)</td>
<td>Type 4: user–task–system–environment</td>
<td>Perception</td>
<td>Chart review/log analysis (115)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Questionnaire (54)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview (34)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Focus group (8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Observation (4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview (6)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Task analysis (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Log analysis (9)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Observation (7)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Time-and-motion (3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Work-sampling (1)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Work-sampling (3)</td>
</tr>
</tbody>
</table>

is to understand the impact of health IT beyond the short-term measures of system-user-task-environment interaction. Therefore, the key question is ‘How does the system impact healthcare?’

One hundred and fifteen stage 5 studies evaluated the efficacy of health IT using log files or chart reviews. Other methods measured outcomes including guideline adherence, utilization, accuracy, document quality, medication error, patient outcomes, and cost-effectiveness.

Study design and data analysis in stages 4 and 5

Because stages 4 and 5 involve field testing, which examines higher levels of health IT effectiveness and efficiency, such as utilization, process efficiency and prescribing behaviors, we further analyzed studies in stages 4 and 5. More quantitative studies were conducted than qualitative studies. Qualitative methods included observation, interview, and focus group. Fourteen studies (20%) in stage 4 and 50 studies (26%) in stage 5 were qualitative. We also categorized the quantitative studies based on their objectivity or subjectivity (appendix IV, available as a data supplement online only). The most common study design was a quasi-experimental study design using one group with pretest and posttest comparison, or with only posttest assessment. The MEDLINE search retrieved an equal number of subjective evaluations (eg, self-report questionnaire) and objective evaluations (eg, log file analysis, cognitive walkthrough, and chart review).

the user, task, system, and environment interaction performance in terms of output quality, speed, accuracy, and completeness?’; ‘Is the user satisfied with the way the system helps perform a task in the real setting?’; and ‘Does the system change workflow effectiveness or efficiency?’

Twelve studies used observation, log files, and/or chart review to assess interaction performance, such as accuracy, time, completeness, and general workflow. Methods to evaluate user satisfaction included questionnaires (n=57), interviews (n=9), and focus groups (n=4). The third question aims to understand users’ work quality, such as workflow and process efficiency. One study used workflow analysis to evaluate the work process before and after health IT was implemented.

The motivation for adopting nursing information systems is often increased time for direct patient care. Researchers use time-and-motion studies to collect work activity information for time efficiency before and after using health IT. Two stage 4 studies in this literature review conducted randomized controlled trials to measure the efficiency of health IT. Although system impact is not the primary focus of stage 4, 32 studies assessed system utilization, patient outcomes, guideline adherence, and medication errors.

Stage 5: routine use

Most studies (n=196) included in this review evaluated health IT at stage 5 of the SDLC. The main purpose of evaluation at this stage...
With regard to analysis methods, only a small number of studies reported using multivariate analysis (eg, linear regression, logistical regression, general equation modeling); most used descriptive and comparative methods. For example, in studies in which questionnaires were used to assess clinicians’ perceptions of health IT, descriptive analysis was usually used with comparisons between different types of users, such as physicians versus nurses.

**Theories and methods used in usability studies**

Theories being applied in health IT usability studies can be grouped into four categories: general system development/design framework, HCI, technology acceptance and technology adoption. Researchers use these theories/models to support their study rationales. We present these theories/models, their references and example studies in appendix V (available as a data supplement online only). Methods used in the usability studies address research questions to understand system specification, interface design, task/workflow identification, user–task–system interaction, field observation and multitask performance. We also categorized methods used in the usability studies, their references and example studies in appendix VI (available as a data supplement online only). Two additional methods that were not within our original search scope, but which were subsequently discovered are also included because of their potential value to health IT usability evaluation. Online appendices V and VI can be used to guide researchers in the selection of theories and methods for usability studies in the future.

**Guidance for health IT usability evaluation**

The usability specification and evaluation framework (table 2), which we adapted for application in our review, identifies the time points and interaction types for evaluation and facilitates the selection of theoretical models, outcome measures, and evaluation methods. A guide based upon the framework is summarized in table 7. This guide has the potential to assist researchers and those who develop and implement systems to design usability studies that are matched to specific system and evaluation goals.

**DISCUSSION**

Most studies evaluated health IT at stage 4 and stage 5. Some noted that the health IT had been evaluated before it went into field testing. Health IT that is evaluated only in stages 4 or 5 may be commercial products, and therefore, organizations may have lacked the opportunity for earlier evaluation. Studies identified adoption barriers due to system validation, usefulness, ease of use, system flow, workflow interruption, insufficient system training, or technology support. Some of these barriers may have been minimized by applying evaluation methods during stage 1 to stage 3. In this section we summarize methodological issues in existing studies and describe a stratified view of health IT usability evaluation.

**Methodological issues in existing studies**

We discovered several problems in existing studies. Even though there is an increased awareness regarding the importance of usability evaluation, most studies in this review provided limited information about associated evaluations relevant to each SDLC stage.

**Lack of theoretical framework/model**

Theoretical frameworks/models are essential to research studies, suggesting rationale for hypothesized relationships and providing the basis for verification. However, the majority of publications in this review lacked an explicit theoretical framework/model. This is consistent with a previous literature review of evaluation designs for health information systems. Most theoretical frameworks used were adapted from HCI, information system management, or cognitive behavioral science. In addition, there are no clear guidelines or recommendations for the utilization of a theoretical framework in health IT usability evaluation.

**Table 7** Guide for selection of theories, outcomes, and methods based on type of interaction

<table>
<thead>
<tr>
<th>Type</th>
<th>Theory applied (online supplement)</th>
<th>Examples of outcomes measured</th>
<th>Methods (online supplement)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 0: task-based</td>
<td>NA</td>
<td>System specification</td>
<td>System element identification</td>
</tr>
<tr>
<td>Type 1: user–task</td>
<td>General system development/design framework</td>
<td>System specification</td>
<td>Interface design</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Task/workflow identification</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Questionnaire</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Interview</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Focus group</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Sensitivity and specificity</td>
</tr>
<tr>
<td>Type 2: system–task</td>
<td>NA</td>
<td>System validation</td>
<td>Chart review/log analysis</td>
</tr>
<tr>
<td>Type 3: user–task–system</td>
<td>Human–computer interaction Technology acceptance</td>
<td>Accuracy</td>
<td>User–system–task interaction</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speed</td>
<td>Questionnaire</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Completeness</td>
<td>Interview</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interaction</td>
<td>Focus group</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Perception</td>
<td>Task/workflow analysis</td>
</tr>
<tr>
<td>Type 4: user–task–system–environment</td>
<td>Human–computer interaction Technology acceptance Technology adoption</td>
<td>Utilization</td>
<td>Chart review/log analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Patient outcome</td>
<td>Field observation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Guideline adherence</td>
<td>Questionnaire</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Medication errors</td>
<td>Interview</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Accuracy</td>
<td>Focus group</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Document quality</td>
<td>Chart review/log analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cost-effectiveness</td>
<td></td>
</tr>
</tbody>
</table>
studies. The guidance provided in table 7 can potentially assist with this issue.

Lack of details regarding qualitative study approaches
Qualitative studies generally use interviews or observations to explore in-depth knowledge of users’ experiences, work patterns and human behaviors that methods such as surveys or system logs cannot capture. However, most qualitative studies did not provide detailed information regarding the study approaches applied to answer the research questions. For example, phenomenology, ethnography and ground theory are commonly used to understand users’ experience, culture and decision-making process, respectively. Research in HCI and computer-supported cooperative work (CSCW) has also established user-centered approaches such as participatory design and contextual enquiry to understand the relationship between technology and human activities in health care.67—71 In our review, 29 qualitative studies integrated clinicians from the beginning stages of the system development process to define clinician needs, system specification or workflow as is done with participatory design approaches. However, the studies reviewed did not directly refer to a specific study approach such as participatory design. Explication of the qualitative approach is important to determine study rigor with regard to criteria such as credibility, trustworthiness, and auditability.

Single evaluation focus
Although usability evaluation ideally examines the relationship of users, tools, and tasks in a specific working environment, most of the studies in this review focused on single measures, such as time efficiency or user acceptance, which cannot convey the whole picture of usability. Furthermore, some potentially useful methods (eg, task analysis and workflow analysis) were rarely used in the studies in our review.

Environmental factors not evaluated in the early stages
Among studies in the development stage (stage 1), only four reported conducting task or workflow analysis. Even though identifying system features, functions, and interfaces from the user perspective is essential for the design of usable system, the perspective is frequently lacking. As a possible consequence, studies have identified workarounds and barriers to adoption due to lack of fit between system and workflow or environment.72 75

Guideline adherence as the primary outcome for clinical decision support systems
Clinical decision support systems were typically evaluated for their effectiveness by clinician adherence to guideline recommendations. However, guideline adherence is influenced by providers’ estimates of the time required to resolve reminders resulting in low adherence rates when providers estimate a long resolution time.74 Other usability barriers related to computerized reminders in this review included workflow problems, such as receipt of reminders while not with a patient, thus impairing data acquisition and/or the implementation of recommended actions; and poor system interface. Facilitators included: limiting the number of reminders; strategic location of the computer workstations; integration of reminders into workflow; and the ability to document system problems and receive prompt administrator feedback.65

Physicians are more adherent to guideline recommendations when they are less familiar with the patient.75 However, greater adherence does not suggest better treatment, because additional knowledge beyond the electronic health record was found to be the major reason for non-adherence.76 Therefore, measuring guideline adherence rates may not capture pure usability; adjustments should be made for other confounders, such as familiarity with patients and physicians’ experience. Moreover, as with other application areas, relevant techniques should be applied at earlier stages of SDLC to tease out user interface versus workflow issues (see appendix IV, available online only). For example, application of a Wizard of Oz approach77 78 in stage 3 would allow the examination of HCI issues without the presence of a fully developed rules engine.

A stratified view of health IT usability evaluation
The SDLC provides a commonly accepted guide for system development and evaluation. However, it does not focus on the type of interactions that are evaluated. Consequently, it is difficult to ascertain if a system issue is due to system—task interaction, user—system—task interaction or user—system—task—environment interaction.

Inspired by Bennett and Shackel’s usability model, the health IT usability specification and evaluation framework (table 2), which we adapted and applied in our review, provides a categorization of study approaches by evaluation types. Existing models or frameworks successfully identify potential factors that influence health IT usability. However, because of the varied manner in which the factors influence interactions, it is difficult to determine if problems stem from health IT usability, user variance, or organizational factors. Therefore, a stratified view of health IT usability evaluation (figure 2) may potentially provide a better understanding for health IT evaluation.

In the stratified view, level 1 targets system specification to understand user—task interaction for system development, which is usually conducted in SDLC stage 1. Level 2 examines the task performance to assess system validation and HCI, which is generally evaluated in SDLC stages 2 and 3. Level 3 aims to incorporate environmental factors to identify work processes and system impact in a real setting, which is commonly assessed in SDLC stages 4 and 5. However, in many situations health IT is evaluated only at SDLC stages 4 and 5;
evaluation also should occur at earlier stages in order to determine which level of interaction is problematical.

Task/expectation complexity, user variances, and organizational support are factors that influence the use of the system, but are not problems of the system itself, and need to be differentiated from system-related issues. For example, at level 1, through application of user-centered design, we can control some user variance by recruiting the targeted users as participants. In addition, task/expectation complexity can be measured to identify system specifications. At level 3, we can minimize user variance by user training and providing sufficient organizational support. An evaluation of perceived usability based on the level of task/expectations would then be able to reveal the system usability at each level of task/expectations.

Friedman and Wyatt’s table of nine generic study types describes similar ideas of incorporating users and tasks into the usability evaluation at certain system development time points. The stratified view of health IT usability evaluation extends the concept of evaluating with users and tasks to considering levels of user—task—system—environment interaction, as well as identifying confounding factors, task/expectation complexity, user variances, and organizational support, that directly or indirectly influence the results of usability evaluation. The stratified view potentially provides a clearer explanation of interactions and factors influencing interactions, and can assist those conducting usability evaluations to focus on the interactions without overlooking the ultimate goal, health IT adoption, which is also influenced by non-interaction factors.

Limitations

There are a number of limitations to the review. First, the usability review used only one database (MEDLINE). Therefore, we may have missed methods that were used only in HCI or CSCW studies published in other scientific databases. To estimate the extent of this limitation, we searched Scopus to identify any additional theories and methods in HCI and CSCW-related journals or conference proceedings that were not found in the studies that we retrieved from MEDLINE. Although we found only one theoretical framework and one method not covered in our review, approaches such as participatory design and ethnography provide overall frameworks that link sets of methods together in a manner that is greater than the sum of their parts. Therefore, a review that included searches of additional databases such as PubMed, Web of Science and Scopus would provide additional practical guidance beyond that derived from our analysis in this review. Second, because we used MeSH and keywords to retrieve results, studies with inappropriate indexing or that lacked sufficient information in their abstracts may not have been retrieved. Third, the number of reviewers was small. Studies were analyzed by one author (PYY) with review by the second (SB). Both authors agreed upon the study extraction and categorization. Fourth, our review focused on identifying methods and scales used, but did not evaluate best methods or scales for usability evaluation because the selection of methods and scales also depends on the type of health IT, evaluation goals, and other variables. However, this was the first review of health IT studies using an integrated usability specification and evaluation framework, and it provides an inventory of evaluation frameworks and methods and practical guidance for their application.

CONCLUSION

Although consideration of user—task—system—environment interaction for usability evaluation is important, existing reviews do not provide guidance based on the user—task—system—environment interaction. Therefore, we reviewed and categorized current health IT usability studies and provided a practical guide for health IT usability evaluation that integrates theories and methods based on user—task—system—environment interaction. To better identify usability problems at different levels of interaction, we also provided a stratified view of health IT usability evaluation to assist those conducting usability evaluations to focus on the interactions without overlooking some non-interaction factors. There is no doubt that the usability of health IT is critical to achieving its promise in improving health care and the health of the public. Toward such goals, these materials have the potential to assist conducting usability studies at different SDL stages and in measuring different evaluation outcomes for specific evaluation goals based on users’ needs and levels of expectation.

Funding The research was supported by the Center for Evidence-based Practice in the Underserved (P30NR010677).

Competing interests None.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES


J Am Med Inform Assoc 2012;19:413–422. doi:10.1136/amiajnl-2010-000020 421


