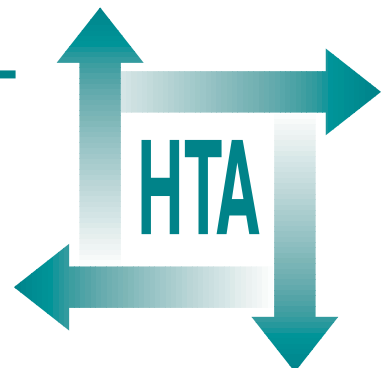


# Methods for the analysis of quality-of-life and survival data in health technology assessment

LJ Billingham  
KR Abrams  
DR Jones



Health Technology Assessment  
NHS R&D HTA Programme



## Standing Group on Health Technology

### Current members

---

**Chair:**

**Professor Sir Miles Irving,**  
Professor of Surgery, University  
of Manchester, Hope Hospital,  
Salford

Professor Martin Buxton,  
Professor of Economics,  
Brunel University

Professor Francis Creed,  
School of Psychiatry  
& Behavioural Sciences,  
University of Manchester

Professor Charles Florey,  
Department of Epidemiology  
& Public Health, Ninewells  
Hospital & Medical School,  
University of Dundee

Professor John Gabbay,  
Director, Wessex Institute for  
Health Research & Development

Professor Sir John  
Grimley Evans,  
Department of  
Geriatric Medicine,  
Radcliffe Infirmary, Oxford

Dr Tony Hope,  
The Medical School,  
University of Oxford

Professor Richard Lilford,  
Regional Director, R&D,  
West Midlands

Dr Jeremy Metters,  
Deputy Chief Medical Officer,  
Department of Health

Professor Maggie Pearson,  
Regional Director of R&D,  
NHS Executive North West

Mr Hugh Ross,  
Chief Executive, The United  
Bristol Healthcare NHS Trust

Professor Trevor Sheldon,  
Director, NHS Centre for  
Reviews & Dissemination,  
University of York

Professor Mike Smith,  
Director, The Research  
School of Medicine,  
University of Leeds

Dr John Tripp,  
Department of Child Health,  
Royal Devon & Exeter  
Healthcare NHS Trust

Professor Tom Walley,  
Department of  
Pharmacological Therapeutics,  
University of Liverpool

Dr Julie Woodin,  
Chief Executive,  
Nottingham Health Authority

Professor Kent Woods  
(**Chair Designate**),  
Regional Director of R&D,  
NHS Executive, Trent

### Past members

---

Dr Sheila Adam,  
Department of Health

Professor Angela Coulter,  
Director, King's Fund, London

Professor Anthony Culyer,  
Deputy Vice-Chancellor,  
University of York

Dr Peter Doyle,  
Executive Director, Zeneca Ltd,  
ACOST Committee on Medical  
Research & Health

Professor John Farndon,  
Professor of Surgery,  
University of Bristol

Professor Howard  
Glennester,  
Professor of Social Science  
& Administration, London  
School of Economics &  
Political Science

Mr John H James,  
Chief Executive,  
Kensington, Chelsea &  
Westminster Health Authority

Professor Michael Maisey,  
Professor of Radiological  
Sciences, Guy's, King's & St  
Thomas's School of Medicine  
& Dentistry, London

Mrs Gloria Oates,  
Chief Executive,  
Oldham NHS Trust

Dr George Poste,  
Chief Science & Technology  
Officer, SmithKline Beecham

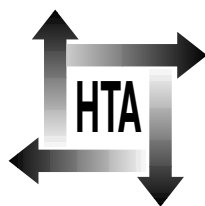
Professor Michael Rawlins,  
Wolfson Unit of  
Clinical Pharmacology,  
University of Newcastle-  
upon-Tyne

Professor Martin Roland,  
Professor of General Practice,  
University of Manchester

Professor Ian Russell,  
Department of Health Sciences  
& Clinical Evaluation,  
University of York

Dr Charles Swan,  
Consultant Gastroenterologist,  
North Staffordshire  
Royal Infirmary

Details of the membership of the HTA panels, the NCCHTA Advisory Group and the HTA Commissioning Board are given at the end of this report.



### **How to obtain copies of this and other HTA Programme reports.**

An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (<http://www.nchta.org>). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents, York Publishing Services.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents, York Publishing Services by:

- fax (with **credit card** or **official purchase order**)
- post (with **credit card** or **official purchase order** or **cheque**)
- phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

### **Contact details are as follows:**

York Publishing Services  
PO Box 642  
YORK YO31 7WX  
UK

Email: [nchta@yps-publishing.co.uk](mailto:nchta@yps-publishing.co.uk)  
Tel: 0870 1616662  
Fax: 0870 1616663  
Fax from outside the UK: +44 1904 430868

NHS and public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please contact York Publishing Services at the address above. Subscriptions can only be purchased for the current or forthcoming volume.

### **Payment methods**

#### *Paying by cheque*

If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *York Publishing Distribution* and drawn on a bank with a UK address.

#### *Paying by credit card*

The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

#### *Paying by official purchase order*

You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

### **How do I get a copy of HTA on CD?**

Please use the form on the HTA website ([www.nchta.org/htacd.htm](http://www.nchta.org/htacd.htm)). Or contact York Publishing Services (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

---

The website also provides information about the HTA Programme and lists the membership of the various committees.



# Methods for the analysis of quality-of-life and survival data in health technology assessment

LJ Billingham<sup>1,2</sup>

KR Abrams<sup>1</sup>

DR Jones<sup>1</sup>

<sup>1</sup> Department of Epidemiology and Public Health,  
University of Leicester, UK

<sup>2</sup> Cancer Research Campaign Trials Unit, University of  
Birmingham, UK

Published July 1999

---

This report should be referenced as follows:

Billingham LJ, Abrams KR, Jones DR. Methods for the analysis of quality-of-life and survival data in health technology assessment. *Health Technol Assess* 1999;**3**(10).

Health Technology Assessment is indexed in Index Medicus/MEDLINE and Excerpta Medica/EMBASE. Copies of the Executive Summaries are available from the NCCHTA web site (see overleaf).

# NHS R&D HTA Programme

The overall aim of the NHS R&D Health Technology Assessment (HTA) programme is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and work in the NHS. Research is undertaken in those areas where the evidence will lead to the greatest benefits to patients, either through improved patient outcomes or the most efficient use of NHS resources.

The Standing Group on Health Technology advises on national priorities for health technology assessment. Six advisory panels assist the Standing Group in identifying and prioritising projects. These priorities are then considered by the HTA Commissioning Board supported by the National Coordinating Centre for HTA (NCCHTA).

This report is one of a series covering acute care, diagnostics and imaging, methodology, pharmaceuticals, population screening, and primary and community care. It was identified as a priority by the Methodology Panel and funded as project number 93/50/04.

The views expressed in this publication are those of the authors and not necessarily those of the Standing Group, the Commissioning Board, the Panel members or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for the recommendations for policy contained herein. In particular, policy options in the area of screening will be considered by the National Screening Committee. This Committee, chaired by the Chief Medical Officer, will take into account the views expressed here, further available evidence and other relevant considerations.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

Series Editors: Andrew Stevens, Ruairidh Milne and Ken Stein  
Editorial Assistant: Melanie Corris

The editors have tried to ensure the accuracy of this report but cannot accept responsibility for any errors or omissions. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

© Crown copyright 1999

Enquiries relating to copyright should be addressed to the NCCHTA (see address given below).

Published by Core Research, Alton, on behalf of the NCCHTA.

Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

---

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,  
Mailpoint 728, Boldrewood,  
University of Southampton,  
Southampton, SO16 7PX, UK.  
Fax: +44 (0) 1703 595 639 Email: [hta@soton.ac.uk](mailto:hta@soton.ac.uk)  
<http://www.hta.nhsweb.nhs.uk>



# Contents

<b>List of abbreviations</b> .....	i	<b>8 Longitudinal modelling methods that deal with informative drop-out</b> .....	35
<b>Executive summary</b> .....	iii	Introduction .....	35
<b>1 Introduction</b> .....	1	Methods .....	35
Aims .....	1	Discussion .....	36
Rationale for the study .....	1	References .....	36
Introduction to this report .....	1	<b>9 Quality-of-life analysis in the presence of informative drop-out: summary and discussion</b> .....	37
References .....	2	References .....	37
<b>2 Literature search</b> .....	3	<b>10 Analysis of survival data adjusting for quality of life: introduction</b> .....	39
Introduction .....	3	<b>11 Comparison of treatments in terms of survival</b> .....	41
Search methodology .....	3	Introduction to survival analysis .....	41
Results of the search .....	5	Survivor and hazard functions .....	41
Further potential searching .....	5	Comparison of treatments using Kaplan–Meier estimates .....	42
Conclusions .....	6	Non-parametric tests for comparing survival .....	43
<b>3 Background to quality-of-life and survival data</b> .....	7	References .....	44
The meaning of quality of life .....	7	<b>12 Comparison of treatments in terms of survival adjusting for covariates</b> .....	45
Measuring quality of life .....	7	Introduction .....	45
Quality-of-life data .....	8	Stratified survival analysis .....	45
Survival data .....	11	Cox proportional hazards model with fixed covariates .....	46
References .....	11	Cox model with time-dependent covariates .....	47
<b>4 Background to the illustrative example – the MIC study</b> .....	13	References .....	49
Introduction .....	13	<b>13 Analysis of survival data adjusting for quality of life: summary and discussion</b> ....	51
Background .....	13	<b>14 Simultaneous analysis of quality-of-life and survival data: introduction</b> .....	53
Measuring quality of life .....	13	<b>15 Quality-adjusted life years</b> .....	55
Missing data in the quality-of-life study .....	15	Introduction .....	55
References .....	16	Utilities .....	55
<b>5 Quality-of-life analysis in the presence of informative drop-out: introduction</b> .....	19	QALY methodology .....	57
<b>6 Descriptive methods for the analysis of longitudinal quality-of-life data</b> .....	21	Example: calculating QALYs in the MIC study .....	58
Introduction .....	21	Application of QALYs .....	60
Patient profiles .....	21	Critical appraisal of QALYs .....	61
Group profiles .....	23	Alternative and related methods to QALYs .....	62
References .....	27	References .....	63
<b>7 Standard methods for the analysis of longitudinal data</b> .....	29		
Introduction .....	29		
Analysis of each time point separately .....	29		
Summary measures .....	29		
Repeated measures analysis of variance .....	31		
More complex modelling techniques for longitudinal data .....	31		
References .....	32		

<b>16 Quality-adjusted survival analysis</b> .....	71	<b>19 Modelling quality of life and survival as two simultaneous processes</b> .....	99
Introduction .....	71	References .....	99
Quality-adjusted survival analysis using a QALY model .....	71	<b>20 Simultaneous analysis of quality-of-life and survival data: summary and discussion</b> .....	101
TWiST .....	73	References .....	102
Q-TWiST .....	75	<b>21 Conclusions and recommendations</b> .....	103
Critical appraisal of quality-adjusted survival analysis .....	81	Overall summary and discussion .....	103
References .....	82	Recommendations .....	104
<b>17 Multistate survival analysis</b> .....	85	References .....	106
Introduction .....	85	<b>Acknowledgements</b> .....	107
Defining the model .....	85	<b>Bibliography</b> .....	109
Modelling the transition rates .....	87	<b>Appendix I</b> Details of search terms used in electronic database searching .....	145
Modelling transition rates in the MIC data ...	88	<b>Health Technology Assessment reports published to date</b> .....	147
Alternative modelling approaches when exact transition times are not known .....	91	<b>Health Technology Assessment panel membership</b> .....	149
Critical appraisal of multistate survival analysis .....	93		
References .....	94		
<b>18 Incorporating results from a multistate model into a Q-TWiST analysis</b> .....	97		
Introduction .....	97		
Methodology .....	97		
Conclusions .....	97		
Reference .....	98		





## List of abbreviations

BIDS	Bath Information and Data Service
CI	confidence interval
CT	chemotherapy arm (MIC study)
EM	an algorithm consisting of an expectation step and a maximisation step
EORTC	European Organisation for Research and Treatment of Cancer
HYE	healthy-years equivalent
MAL	malaise variable (MIC study)
MIC	mitomycin, ifosfamide and cisplatin
MQS	mean quality-of-life score
PAL	palliative treatment arm (MIC study)
QAL	quality-adjusted life
QALY	quality-adjusted life year
QoL	quality of life*
Q-TWiST	quality-adjusted TWiST
REL	time spent with symptoms of disease following disease relapse
SAVE	saved young life equivalent
TOX	time spent with toxicity from treatment
TR	time to disease relapse
TTO	time trade-off
TWiST	time spent without symptoms of disease and toxicity of treatment
TWiST(L)	accumulated TWiST

\* Used only in tables and figures





## Executive summary

### Objectives

Quality of life has become an important issue in health care, especially in studies of chronic diseases. Substantial amounts of quality-of-life data are now being gathered in clinical trials, using a variety of instruments. In longitudinal studies of quality of life in which survival is also an endpoint, patients are generally severely ill and it is common for participants to drop out of the study because of illness or death. In such situations, the drop-out process may depend on the quality of life being experienced, rather than being random; hence the incomplete follow-up of patients is called informative drop-out. This must be appropriately accounted for in any analysis of the data to avoid the introduction of bias.

This study identifies and reviews critically the methods proposed for the analysis of quality-of-life and survival data in health technology assessment, particularly those that assess both these endpoints simultaneously. In this way methodology that requires wider dissemination can be identified together with areas requiring further research. It was not within the remit of this study to address issues related to the meaning, definition and measurement of quality of life.

### Methods

The scientific and medical literature was searched for relevant methodological articles. Electronic searches were carried out systematically using Science Citation Index, Social Science Citation Index and the EMBASE database provided by BIDS (Bath Information and Data Service). The searches were supplemented by exploded references, personal collections and handsearching of the journal *Quality of Life Research*.

### Results

Methods for analysing quality-of-life and survival data were found to fall into three broad categories, as described below, according to the research question underlying the study; this in turn depends on the disease and treatments under investigation.

### Quality-of-life analysis in the presence of informative drop-out

The use of standard methods for the analysis of longitudinal data is discussed in terms of their application to longitudinal quality-of-life data. All methods, from simple descriptive analysis to complex modelling techniques, will give biased results when informative drop-out is present in the data. Standard methods should therefore be used with caution when analysing longitudinal quality-of-life data. Modelling techniques that deal with informative drop-out have been developed and their application to quality-of-life data is discussed.

### Analysis of survival data adjusting for quality of life

In comparing treatments in terms of survival, it is often necessary to adjust for other patient-related factors, known as covariates, that could potentially affect the survival time of a patient. In some situations the survival analysis may need to adjust for baseline measures of quality of life (fixed covariates), while in others, allowance for changing quality of life over time may be required (time-dependent covariates). If assessments of quality of life are infrequent or data are missing for reasons other than death, then it may be difficult to adjust for changes in quality of life with any degree of accuracy. Modelling quality of life and survival as two simultaneous processes may improve the analysis in this situation.

### Simultaneous analysis of quality-of-life and survival data

In studies in which quality of life and survival are both important endpoints, it may be advantageous to assess health technologies in terms of these endpoints simultaneously. Three different approaches can be used to achieve this:

- combining quality and quantity of life into a single endpoint and using quality-adjusted survival analysis methods to compare treatments
- using multistate models to model the movement of patients between various health states, defined by levels of quality of life and by death, and exploring how treatments differ in terms of these movements

- considering quality of life and survival as two simultaneous processes and describing the data in terms of two interlinked models.

Quality and quantity of life can be combined into a single endpoint by weighting periods of survival time according to the quality of life experienced. The resulting outcome measures are generally referred to as QALYs (quality-adjusted life years) with special forms known as TWiST (time spent without symptoms of disease and toxicity of treatment) and Q-TWiST (quality-adjusted TWiST). The use of standard survival analysis techniques on the QALY endpoint will generally give biased results because individuals with a worse quality of life will be censored earlier than those with a good quality of life, resulting in informative censoring. Methods of overcoming this problem, including partitioned survival analysis, are discussed. Quality-adjusted survival analysis overcomes problems of informative drop-out due to death and has the potential to be extended to deal with other disease- or treatment-related reasons for drop-out.

Multistate models are defined by categorising the period of follow-up of patients in a trial into a number of different health states defined in terms of levels of quality of life and death. The movement between health states is described by transition rates, which are modelled using the transition times for patients. Various modelling approaches are discussed. The inclusion of death as a health state in the model enables the analysis to deal with informative drop-out due to death and the inclusion of a 'drop-out' state could cover other reasons.

The most recently developed, and potentially most powerful, approach to analysing quality-of-life and survival data is to model the longitudinal quality-of-life data and the drop-out process, which includes drop-out due to death, as two simultaneous processes. Such an approach has the advantage of allowing quality-of-life data to be assessed longitudinally while adjusting for informative drop-out. In addition, the inter-relationship between the two can be explored.

## Conclusions and recommendations

### Obtaining appropriate data

- The method of analysis needs to be decided at the design stage of a study so that appropriate

quality-of-life data can be collected. Issues to consider are:

- the quality-of-life instrument to be used
- the frequency and timing of quality-of-life assessments
- the need to minimise non-compliance
- the collection of additional information, such as reason for drop-out
- the sample size required.

### Choosing the appropriate method

- The choice of method should be based on the research question that the study aims to answer. The advantages and disadvantages of each method should be considered carefully together with the relevance and interpretability of the results to clinicians and patients.
- Methods used to analyse longitudinal quality-of-life data must allow for informative drop-out.

### Reporting the analysis

- Methods used should be reported clearly, with details of definitions and assumptions used in the analysis.
- A sensitivity analysis should be carried out to assess the robustness of conclusions to any critical assumptions made in the analysis.

### Recommendations for further research

- Further experience in the application of quality-adjusted survival analysis techniques to quality-of-life data is needed to enable a proper evaluation of such methods.
- Further research is needed in order to develop hierarchical models, multistate models and simultaneous modelling methods in their practical application to quality-of-life and survival data using both classical and Bayesian approaches. Consideration should be given to how methods could deal with the multivariate nature of the quality-of-life endpoint.
- A full review of available software for methods that simultaneously analyse quality-of-life and survival data is needed to highlight areas requiring further development.
- Progress in the most rapidly developing areas of multistate survival analysis and simultaneous modelling should be monitored, together with parallel areas of methodological development such as in the field of AIDS research.

# Chapter I

## Introduction

### Aims

The methods that have been proposed for the analysis of quality of life and survival data in health technology assessment are reviewed, with particular reference to those that assess these two endpoints simultaneously. The key objectives were to:

- identify proposed methodology for the simultaneous analysis of quality and length of life
- review critically the proposed methodology
- illustrate such methodology where possible by application to data from a previously conducted study
- identify on the basis of the review:
  - methodology that requires dissemination within the health research community and the NHS
  - areas of non-existent or deficient methodology
  - areas that require further research.

### Rationale for the study

Quality-of-life assessment has become an important issue in healthcare research and the study of new technologies, especially in many chronic diseases. Although survival is usually the standard endpoint for assessing treatments in clinical trials, informed clinical decisions generally require quantification and interpretation of quality-of-life measures, especially with respect to variation over time and the interrelationship with length of life. The role of quality of life will become even more prominent in the future as, for many diseases, improvements in survival due to treatment are either unlikely to be dramatic or likely to be made at the expense of quality of life.

There has been much research into the development of instruments with which to measure quality of life. This has resulted in a plethora of instruments and substantial amounts of quality-of-life data being gathered in trials. Hence there is a need for methods that enable the effective and efficient analysis and interpretation of such data. It is therefore timely that the range of statistical methods which have been proposed for dealing with such data are systematically reviewed and

evaluated. There are a number of previous reviews of the analysis of quality of life in clinical trials.<sup>1-7</sup>

In longitudinal studies of quality of life in which survival is also an endpoint, the patient population will not be stable over time. Patients are generally severely ill and individuals may have incomplete follow-up of quality of life for reasons related to disease or treatment, including death. This drop-out process may be informative and can make statistical analysis particularly problematic. There is, therefore, a need to identify appropriate methods which will yield unbiased and clinically relevant assessment of health technologies in such situations.

This study will therefore yield two benefits.

- The identification of existing methodology that has been shown to benefit the assessment of health technologies with respect to both quality and length of life, and thus enable more informed decision-making within the NHS.
- The identification of those areas in which further work is required so that existing methods can either be more appropriately applied or be realistically developed.

### Introduction to this report

#### Coverage

The first four chapters of this report provide an introduction and background. They include details of the methodology used to carry out the review and they provide a background to both types of data encountered in the report – quality-of-life data and survival data. The early chapters also provide a background to the clinical trial from which quality-of-life and survival data were taken to illustrate some of the proposed methodology.

Health technology assessment is the evaluation of any intervention intended to improve health. The investigation of quality of life and survival most often takes place in the context of a randomised clinical trial that compares treatments and the report focuses on this scenario. The main body of the report, in which the proposed methods for

analysing quality-of-life and survival data are discussed, is in three parts according to the definition of the research question of interest. This research question depends on the disease and treatments under investigation.

The first approach, discussed in chapters 5–9, covers studies in which the primary aim is to compare health technologies in terms of quality of life. Survival is also considered but only in terms of the problematical impact it has on the quality-of-life assessment data. Deaths of patients during the course of a study may result in informative drop-out, causing quality-of-life data to be missing. The problems for analysis caused by informative drop-out are discussed.

The second approach, discussed in chapters 10–13, covers studies in which the primary aim is to compare health technologies in terms of survival. Quality of life may also be measured and any survival analysis may need to be adjusted for this.

The third approach, discussed in chapters 14–20, includes studies in which quality of life and survival are both important endpoints for assessing health technologies. Instead of analysing each endpoint separately, it may be more appropriate to analyse both endpoints simultaneously. The focus of the report is on this type of methodology.

In the final chapter of the report the methods discussed are summarised and the implications for the design and conduct of health technology assessment research are considered. Methodology that requires wider dissemination is identified and recommendations for further research are made.

### Format of the report and guidance for readers

The report can be approached in a variety of ways depending on the needs of the individual reader. Although the report is cross-referenced, the section addressing each approach can be read as a single entity, thus enabling the reader to focus on just that part of the report that relates to the particular problem that they wish to address.

Technical detail has been kept to a minimum in order to make the report accessible to a wide range

of readers. For clarification, most key methods are illustrated with worked examples, using data from a real study. However, the analysis presented here is purely to illustrate the methodology and should not be interpreted as a report of the results of this study; these will be presented elsewhere.

It was not within the remit of this study to address issues relating to the meaning, definition and measurement of quality of life. Use of the term, quality of life, in the literature has thus been accepted uncritically. The term, quality of life, is used here to mean anything that purports to measure health-related quality of life or some aspect of it. This does not affect any discussion of the methodological problems of analysing such data when it is assessed longitudinally in conjunction with survival data.

### References

1. Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Spiegelhalter DJ, Jones DR. Quality-of-life assessment: can we keep it simple? *J R Stat Soc (A)* 1992;**155**:353–93.
2. Hopwood P, Stephens RJ, Machin D. Approaches to the analysis of quality of life data: experiences gained from a Medical Research Council Lung Cancer Working Party palliative chemotherapy trial. *Qual Life Res* 1994;**3**:339–52.
3. Fletcher A, Gore SM, Jones DR, Fitzpatrick R, Spiegelhalter DJ, Cox DR. Quality of life measures in health care. II: design, analysis and interpretation. *BMJ* 1992;**305**:1145–8.
4. Olschewski M, Schumacher M. Statistical analysis of quality of life data in cancer clinical trials. *Stat Med* 1990;**9**:749–63.
5. Schumacher M, Olschewski M, Schulgen G. Assessment of quality of life in clinical trials. *Stat Med* 1991;**10**:1915–30.
6. Pocock SJ. A perspective on the role of quality-of-life assessment in clinical trials. *Control Clin Trials* 1991;**12**:S257–65.
7. Scott CB, Stetz J, Bruner DW, Wasserman TH. Radiation Therapy Oncology Group quality of life assessment: design, analysis, and data management issues. *Qual Life Res* 1994;**3**:199–206.

# Chapter 2

## Literature search

### Introduction

The literature search forms the basis of the review and in this chapter details are presented of the methodology used in the search. The aim was to identify all statistical methodology that has been proposed for the analysis of data for health technology assessment with respect to both quality and length of life. The focus was on methods that analyse quality-of-life data over time in situations in which length of survival is an issue and, in particular, methods that simultaneously analyse quality of life and survival.

### Search methodology

#### Introduction

A systematic and repeatable strategy was devised to search the literature for relevant articles. A complete list of all papers citing relevant statistical methods was not necessary, since enumeration of usage was not an objective. However, the search did seek to identify the full range of methods and so was kept as broad and as thorough as possible within the time constraints of the study.

Statistical methodology is represented, not only in theoretical but also in applied literature. The literature of greatest interest therefore was essentially of a statistical or medical nature. The search, however, was not limited to these areas since methods found in other fields, such as industry, agriculture and education, could be adapted to be applicable in a health context. Journal articles comprised the main body of literature but all types, including books, reports, conference papers and theses, were considered for inclusion.

The search strategy consisted of a variety of different approaches. The main part was undertaken using electronic databases but a considerable number of references were also obtained by other methods, such as handsearching journals, exploring references, and personal recommendations. When identified, references were stored and managed on a database using Reference Manager bibliographic software v. 7 (Research Information Systems, USA, 1995).

#### Electronic database searching

A variety of electronic databases provided by Bath Information and Data Service (BIDS) were used to search the scientific literature. BIDS Science Citation Index provides access to over 4400 journals in natural, physical and biomedical science and technology, BIDS EMBASE to over 3500 pharmacological and biomedical journals, and BIDS Social Science Citation Index to over 1400 journals in behavioural and social sciences.

The search was kept as broad as possible with articles in any language being included and searches going back in time as far as databases would allow (1980 for EMBASE; 1981 for Science Citation Index and Social Science Citation Index). A broad strategy was maintained by searching articles for relevant words and phrases (search terms), not only in keywords but also in titles and abstracts.

The search terms fell into three main categories; quality of life, survival and methodology (see appendix 1 for details). The terms were chosen to achieve an adequate balance between sensitivity and precision (or specificity). Sensitivity is the proportion of relevant references captured by the search, while precision is the proportion of articles captured by the search which are relevant. Searches need to be sensitive, in order to achieve a comprehensive result overall, and precise, so that the number of references to check for relevance is not impractical.

The search strategy, applicable to each database, was split into three stages as follows.

**Stage 1** Journals in the fields of statistics, epidemiology, clinical trials and biometry (see appendix 1 for details on how these were selected) were searched for articles containing quality-of-life search terms.

**Stage 2** All journals on the database were searched for articles containing methodology search terms, that is, terms relating to known statistical methodologies used in the simultaneous analysis of quality-of-life and survival data.

**Stage 3** All journals on the database were searched for articles containing both quality-of-life and survival search terms.

The first and second stages were devised to retrieve methodological articles. At Stage 1 this was achieved by restricting the search to the more methodological journals, while at the same time allowing any article that had a reference to quality of life. Conversely, at Stage 2 the search allowed any type of journal available on the database but restricted to those articles that specifically used a known relevant methodology. Both stages were devised to produce a relatively small number of references with a high degree of precision. Applying the first two stages of the search strategy to Science Citation Index, Social Science Citation Index and EMBASE produced a manageable number of references, which contained the bulk of the references for the review, including many key papers.

The third stage was devised to make the overall search more comprehensive. It proved, however, to be problematic by producing an unmanageable number of references (approximately 1400 new references from Science Citation Index alone). Attempts were made to make the number of references more manageable by splitting them according to whether or not they contained a longitudinal search term (see appendix 1), with the expectation that papers containing such a term would have a greater chance of being relevant to the review. In practice, however, this did not prove to be helpful since relevant references were found in both categories. Given the time constraints of the study, it was decided to abandon this part of the search strategy.

### Other searching methods

The electronic database searching was supplemented by the following methods:

- the journal *Quality of Life Research* was handsearched from its first issue (1992) for relevant articles
- references known by the researchers or colleagues to be relevant to the study were added to the database
- reference lists of articles already identified were checked and added to the database where relevant ('exploded' references).

### Selection of relevant literature

The criteria thought to make an article 'relevant' to the study are defined formally below, together with a description of the process of checking an article for relevance.

#### Criteria

The criteria for judging the relevance of an article were as follows.

- A paper had to include some sort of quality-of-life assessment over time. Papers in which quality of life was studied at a single time point or as a change from baseline, and in which standard non-longitudinal statistical methods of analysis were used, were rejected.
- Papers in which quality of life was assessed longitudinally in circumstances where survival was not an issue were rejected.
- A paper had to use either a known methodology of interest or a new and clearly detailed methodology. Papers that purely discussed quality of life were rejected.
- Papers that described an application for which the methodology was not clearly detailed were rejected.
- Papers that were essentially discussions of issues relating to quality-of-life instruments, including validity and reliability, were rejected.

As a general rule, if there was any doubt as to whether a paper should be included or not, then it was included.

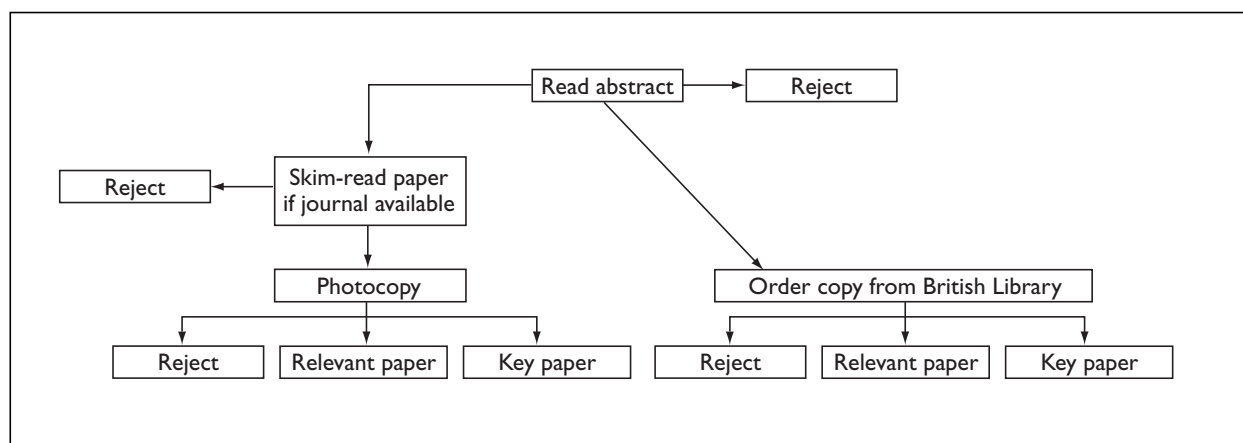
A paper was defined as 'key' if it contained all the elements relevant to the study, that is, it described a methodology used to simultaneously assess health technologies in terms of longitudinal quality of life and survival.

#### Methodology

A strategy was devised (see *Figure 1*) to assess papers for their relevance to the project using the above criteria. For most articles, abstracts were available and these were used initially to assess the article's relevance. If the relevance was not clear from the abstract and the journal containing it was readily accessible, then it was assessed by skim-reading the whole paper. In some cases it was not until a paper was obtained and read in detail that it became apparent that it was not relevant; such papers were discarded. Many papers were in journals that were not available from the university libraries and departmental collections at Leicester and Birmingham; copies of these papers were obtained from the British Library.

One author (LJB) made the first trawl, discarding the obviously non-relevant references and including the obviously relevant ones. All three authors then checked the abstracts of the more questionable references and if at least one author thought a paper might be relevant then it was included.





**FIGURE 1** Strategy to obtain relevant papers from the literature search

## Results of the search

The search produced 1127 references in total, of which 361 were included in the report as relevant (see *Table 1*). Key papers were obtained from all parts of the search strategy.

At Stage 1 of the database search, statistics and biometry journals produced a small number of papers, of which most were key references, whereas the clinical trial and epidemiology journals produced a large number of references, of which the majority were not relevant.

At Stage 2 of the search a large number of references were brought up, of which a large proportion were associated with quality-adjusted life years (QALYs) and, in particular, their use in decision analysis and health economics (see page 61). These papers were of no direct interest and only a selected group of these references (in general, those that were in readily accessible journals) were included.

## Further potential searching

Time limitations resulted in only three electronic databases being searched. These provided good coverage of the literature but, for complete comprehensiveness, other databases should be searched. A variety of medical databases were considered (but not formally used) for the review: MEDLINE, CINAHL, CANCERLIT, AIDSLINE and HEALTHPLAN. In addition, other specialised databases were considered, such as the psychological database PsycLIT, the social sciences database ASSIA, and the statistical databases MATHSCI and CIS.

The overall search strategy identified some grey literature, such as reports, conference papers and theses, which were mainly acquired following recommendation by colleagues. A fully comprehensive search would require a more thorough searching of the grey literature. This would include contacting individuals and departments who are known to be working in the field, together with more formal methods, such as using the special electronic database for grey literature, SIGLE.

**TABLE 1** Numbers of references obtained from each part of the search strategy

Part of search strategy	References retrieved	Relevant references
<b>BIDS search stage 1</b>		
QoL search terms in statistical journals	40	17
QoL search terms in biometry journals	4	4
QoL search terms in clinical trial journals	69	5
QoL search terms in epidemiology journals	198	10
<b>BIDS search stage 2</b>		
Methodology search terms in all journals	600	165
Handsearching <i>Quality of Life Research</i>	29	12
Recommended references	92	65
Exploded references	95	83
<b>Total</b>	<b>1127</b>	<b>361</b>

The electronic searches only went back until 1980. This may well be adequate since quality-of-life research is a relatively new phenomenon and exploded references should identify the most important earlier articles. More formal searching of databases or journals, however, could be carried out to check for all relevant articles published before 1980.

Other approaches to searching electronic databases were considered but were not carried out because of time limitations. An alternative to using search terms would be to search by author, using the names of people active in the area of quality-of-life research. Also, citation searches of key papers found from the existing search could be tried. In addition, an Internet search could be attempted.

The timing of the electronic searches was such that references were restricted to all those

published before 1997. Searching is an ongoing process and the search strategy could be re-run to create a more up-to-date reference database. Relevant articles published after the search that have been brought to our attention have been included in our reference database but have not necessarily been referenced in the report. The bibliography (see page 109) highlights key papers that were published too late for inclusion in this report.

## Conclusions

The search aimed to identify all methods proposed for the analysis of quality-of-life and survival data. It is not possible to determine if this has been achieved but the broadness of the search strategy used, together with the method of exploding references, should ensure a reasonably complete coverage of relevant material.

## Chapter 3

# Background to quality-of-life and survival data

### The meaning of quality of life

In its most general context, quality of life is a concept incorporating all the factors that might impact on an individual's life. In health service research it is more usual to consider health-related quality of life, which includes only those factors that affect an individual's health. There is no general agreement regarding the identification of such factors.<sup>1,2</sup>

WHO defines health as 'a state of complete physical, mental and social well-being, and not merely the absence of disease and infirmity'.<sup>3</sup> Quality of life is often referred to in these terms. For example, Schumacher and colleagues<sup>4</sup> specify the dimensions comprising quality of life as follows:

- symptoms of disease and side-effects of treatment (e.g. nausea, pain, anorexia)
- physical and functional status (e.g. mobility, self-care, fatigue)
- emotional status (e.g. anxiety, depression, satisfaction with care)
- social functioning (e.g. family interaction, work/recreation, time with friends).

In this study, a pragmatic view is taken and quality of life is accepted as any measure that purports to reflect health-related quality of life or some aspect of it.

### Measuring quality of life

There are many issues to consider when attempting to measure quality of life, such as, what questions should be asked, how should responses be recorded, when should questions be asked, of whom should questions be asked and who should do the asking. These, together with other aspects of measuring quality-of-life data in clinical trials, have been extensively reviewed and discussed elsewhere.<sup>5-7</sup>

The quality of life of a patient is usually measured using an instrument in the form of a questionnaire designed for patient completion. The questionnaire generally comprises sets of questions or

items relating to the various dimensions of quality of life, such as physical, psychological or social. The format of responses may be 'yes/no', a series of ordered categories, or a linear analogue scale.

Many instruments are used in the assessment of quality of life,<sup>8-10</sup> for example, Rotterdam Symptom Checklist, Nottingham Health Profile, Sickness Impact Profile, Hospital Anxiety and Depression Scale, Short Form (SF) 36. Generic quality-of-life instruments measure general aspects of quality of life and are applicable in a wide range of research settings, while non-generic instruments are relevant for a specific disease or treatment. Some questionnaires are dimension-specific, that is, they only ask questions relating to a particular aspect of quality of life (e.g. Hospital Anxiety and Depression Scale). A review of the use made of quality-of-life instruments is currently being undertaken.<sup>11</sup>

Much work has been done to test the validity and reliability of some such instruments. An instrument is valid if it is actually measuring what it is designed to measure and it is reliable if, all things being equal, it measures consistently from one occasion to the next. There has been much discussion regarding the requirements of quality-of-life measures and methods for assessing such requirements, for example, Cronbach alpha coefficients and factor analysis.<sup>2,12-15</sup> The criteria for assessing quality-of-life instruments are reviewed in another NHS R&D Health Technology Assessment report.<sup>16</sup>

Another approach to measuring quality of life is based on the valuations of or preferences for different quality-of-life states. The advantage of this approach is that it yields a single value, often referred to as a **utility**, as a measure of quality of life. The utility value is a single score ranging from 0 – representing a quality of life equivalent to death, to 1 – representing perfect health. The relationship between descriptive and valuation approaches has been explored.<sup>17,18</sup> This method of measurement is usually used in the context of the QALY (as this is covered later in chapter 15, details of the methodology are included there).

## Quality-of-life data

### Nature of quality-of-life data

Quality-of-life data is generally longitudinal in nature. Some studies assess quality of life at one time point only or take a baseline measure and a follow-up measure but, generally, quality of life is recorded at more than two time points during the course of a study. There can be any number of time points and these will not necessarily be evenly spaced or consistent across individuals.

Quality-of-life data is generally multivariate in nature. At one extreme, quality of life can be measured by a single global measurement such as the Karnofsky Index,<sup>19</sup> while at the other, assessment is made by a multitude of items measuring a variety of conceptual dimensions, for example, the Sickness Impact Profile measures, in 136 items, 12 dimensions. Items within each dimension are sometimes combined as a weighted or unweighted sum to create dimension-specific global measures, or sometimes all items from the questionnaire are combined to create an overall global quality-of-life score. Thus, at each time point, quality-of-life data may comprise a single measure or a large set of measures.

Quality-of-life data may take various forms (binary, ordinal or continuous) for which a variety of distributional assumptions are appropriate. Responses to each item on a questionnaire may yield binary data from a yes/no response, ordinal data from a categorical scale or continuous data from a linear analogue scale. The aggregation of items, to give a global measurement of a dimension or of overall quality of life, results in data that is usually treated as continuous, despite the fact that the global measure may take only a restricted set of values from a restricted range. The distribution of quality-of-life variables as measured from a linear analogue scale or as an aggregated global score may or may not be normally distributed. Sometimes data may be transformed to create normally distributed data for analysis but on other occasions (e.g. when there is a heavy preponderance of a particular response value, perhaps corresponding to 'no problem') this may be difficult.

### Problem of missing data

One of the main problems in analysing longitudinal quality-of-life data is caused by missing data. The validity of the analysis of data with missing values is dependent on the mechanism associated with the missing data. Responses missing at a particular time point,  $t$ , may be categorised in three ways, as defined by Little and Rubin:<sup>20</sup>

- (i) **missing completely at random**, when the probability of response at time  $t$  is independent of both the previously observed values and unobserved values at time  $t$
- (ii) **missing at random**, when the probability of response at time  $t$  depends on the previously observed values but not the unobserved values at time  $t$
- (iii) **non-ignorable non-response**, when the probability of response at time  $t$  depends on the unobserved values at time  $t$  and possibly on the previously observed values as well.

There are three different forms of missing data that can arise in a longitudinal quality-of-life study designed to take at least three serial measures of quality of life from each participant:

- (i) single missing items from an otherwise complete questionnaire
- (ii) intermittent missing whole questionnaires
- (iii) missing data resulting from a patient dropping out of the study.

The problems associated with each form of missing data and the ways of handling them are discussed in general below and, in relation to the illustrative example, in chapter 4 (see pages 15–16).

### Single missing items

Missing single items mainly cause problems in terms of calculating global scores since a dimension-specific score is based on the values of **all** items within the dimension. If values for items are missing in the dataset, then they either need to be imputed or the calculation of global scores needs to accommodate them (see page 10).

Problems will also occur if the item with missing values is to be analysed independently as a measure of quality of life. Single missing items may not prove to be a major problem in this case, since it is usually not unreasonable to assume that this sort of missing data is missing completely at random. This assumption would be untenable in situations where consistent non-response to an item suggests that the question is inappropriate or difficult to answer.

### Intermittent missing whole questionnaires

An intermittent missing questionnaire occurs if a patient does not complete a questionnaire at the required time point but has completed questionnaires at time points before and after the missing form. It may be possible to assume that this sort of data is missing completely at random, since the patient has not dropped out of the study altogether.

The probability of response may, however, depend on covariates, such as treatment, or may depend on the quality of life experienced at that time, in which case it would be invalid to assume that the missing data mechanism was ignorable.

### **Missing data resulting from patient drop-out**

A drop-out is defined as a patient who withdraws from a study before they have completed all planned assessments. Patients may withdraw from a study for a number of reasons, illness, death, cessation of treatment, lack of treatment effect, lost to follow-up, or they may reach the censor date of the study. In studies where survival, as well as quality of life, is an endpoint, patients are generally severely ill; thus drop-out caused by death or illness will be a common occurrence. In such situations, the drop-out process may depend on the unobserved measurements (i.e. those measurements that would have been observed had the patient not dropped out) and the incomplete follow-up of subjects is called **informative drop-out** (other terms used include **informative censoring** and **non-ignorable drop-out**).

Once a patient has dropped out of a study, no more information on quality of life is available from that point onwards. Quality-of-life information could be considered as censored at the date of drop-out. In terms of analysis, drop-outs cause problems because they create missing data which are not just missing at random. The missing data mechanism is likely to depend upon the health status of the patient and is therefore non-ignorable.

One way of dealing with missing data from drop-outs is to impute values to replace the missing data from data that already exist. There are a variety of methods for doing this<sup>20</sup> but it may be difficult if there is a large amount of missing data. Otherwise the method of analysis must allow for informative drop-out. Methods of analysis which simultaneously assess quality-of-life and survival data (see chapters 14–20) overcome the problems associated with drop-out due to death but informative drop-out for other reasons also needs to be considered in any analysis.

### **Handling the multivariate nature of the quality-of-life endpoint**

Quality-of-life instruments consist of many items, often grouped into a number of dimensions. Because quality of life can be considered in terms of individual items or in terms of separate dimensions, it is, as an endpoint, potentially multivariate in nature. In some situations it may be desirable to consider each item or dimension as a separate

quality-of-life endpoint. In a descriptive analysis, this will only cause problems of presentation and interpretation if the different items or dimensions give conflicting conclusions. If hypothesis testing is involved, analysis of multiple endpoints will lead to the problem of multiple testing, where the probability of a finding false-positives increases as the number of tests performed increases. When quality-of-life measures consist of only a few dimensions, the problem of multiple testing should not be a major one.

In some studies it may be possible to limit the amount of hypothesis testing by specifying in advance a few key quality-of-life measures on which hypotheses will be tested, leaving the remaining variables to be analysed purely descriptively.<sup>21–24</sup> If this approach is not practical or desirable then the analysis will need to account for the multivariate nature of the quality-of-life endpoint.

There are a variety of ways of handling multiple endpoints in clinical trials.<sup>25</sup> The application of such approaches to quality-of-life data has been discussed<sup>4,22</sup> and some methods have been applied and compared in the analysis of quality-of-life data.<sup>26</sup> One approach is to combine multiple endpoints to create global scores before analysis. Another approach is the post-analysis combination of results from the univariate analysis of each separate endpoint. Alternatively, a hierarchical approach to the analysis can be used. These are now considered in turn.

### **Combining multiple endpoints to create global scores**

For each individual, the values of the items that make up a quality-of-life endpoint can be combined in some way to form a global score. In some cases the items within each quality-of-life dimension may be combined to create dimension-specific global scores, while in others an overall quality-of-life global score may be created either by combining all items on a questionnaire or by combining dimension-specific global scores. Use of a single global quality-of-life score simplifies statistical analysis and should be aimed at when sensible and justifiable.<sup>27</sup> If treatments are likely to affect dimensions differently, then combining dimensions into a global score may not be sensible.<sup>24</sup> Global scores can be analysed using standard univariate techniques.

Global scores can be calculated using either an unweighted or a weighted sum. It is suggested that unweighted sums should only be used to combine items which are highly positively

correlated.<sup>27</sup> However, a weighted sum may make interpretation difficult and the weights used may be controversial.<sup>24</sup> Weights can be determined either from the data, using scores from a factor analysis, or from decision theory, using utility analysis techniques, or arbitrarily.<sup>21</sup> Olschewski and Schumacher<sup>27</sup> recommend aggregation using data-oriented procedures.

A method for calculating global scores proposed by O'Brien<sup>28</sup> has been applied to quality-of-life data.<sup>26,29</sup> It is a non-parametric approach and creates global scores from ranks rather than actual data values. The data for all treatment groups are pooled and, for each variable in the multivariate quality-of-life endpoint, the values across all individuals are ranked. A global score is created for each individual by summing the ranks for all variables.

Problems in calculating global scores can occur if data are missing on some of the items within the score. One way to tackle the problem is to impute the missing values. This is only feasible if they are limited in number. If a subject has a missing value for an item then the value could reasonably be imputed from:

- values of the other items within the dimension for that patient
- values of the other items in the patient's questionnaire
- values of the item on the patient's questionnaires at time points on either side of the missing value.

If missing values are not imputed then the formula for calculating global scores needs to allow for the number of items involved in the calculation. Using a mean rather than a sum allows accommodation of missing values into the global score since the mean can be calculated for a reduced number of items. Alternatively, expressing the sum as the percentage of the maximum achievable score<sup>24</sup> allows for the possibility of a reduced number of items. Otherwise the global score should be recorded as missing if any item within it is missing.

### **Combining results from univariate tests on multiple endpoints**

The simplest method for testing a global null hypothesis of no treatment effect, using the results from multiple univariate tests, is to use a Bonferroni-type adjustment. The  $p$ -values from the multiple univariate tests are adjusted by multiplying each  $p$ -value by the number of tests carried out. Each endpoint can be assessed using these adjusted  $p$ -values or a global null hypothesis can be assessed using the minimum  $p$ -value.

This method has been recommended for use with quality-of-life data<sup>21</sup> and has been applied and compared to other methods in a quality-of-life setting.<sup>26</sup> The main drawback with the global null hypothesis approach is that it confines attention to the smallest  $p$ -value and may be too conservative.

A parametric method for combining results from multiple univariate  $t$  tests, originally proposed by O'Brien<sup>28</sup> but developed by Pocock and colleagues,<sup>30</sup> has been applied to quality-of-life data.<sup>26,29</sup> The following test statistic can be used to assess a global null hypothesis of no treatment effect:

$$\mathbf{J}^T \mathbf{S}^{-1} \mathbf{t} / (\mathbf{J}^T \mathbf{S}^{-1} \mathbf{J})^{\frac{1}{2}} \quad (1)$$

where  $\mathbf{J}$  is a vector of ones,  $\mathbf{S}$  is an estimated correlation matrix and  $\mathbf{t}$  is a vector of  $t$  statistics from the separate univariate  $t$  tests. The test statistic has an asymptotic standard normal distribution.

The main drawback of these methods is that they do not give an estimate of the treatment effect, they just provide a test statistic.

### **Hierarchical approach**

Multilevel models have been advocated for the analysis of data that have a hierarchical structure.<sup>31</sup> Longitudinal data can be thought of as hierarchical data, with level one of the hierarchy being observations over time within a patient and level two being the patient. Multilevel models have been used to analyse longitudinal quality-of-life data<sup>32,33</sup> and their application is discussed in more detail in chapter 7.

The hierarchical approach also provides a means of handling the multivariate nature of the quality-of-life endpoint.<sup>32,33</sup> The multiple dimensions that constitute quality of life can simultaneously be analysed in a multilevel model by adding an extra level to the standard longitudinal data model, with level one becoming the various quality-of-life dimensions, level two the observations over time and level three the patients. Models are fitted to the data in the usual way using a variety of assumptions.

Multilevel models have an advantage over methods discussed previously in that they provide estimates of the treatment effect as well as test statistics. Treatment effects are estimated for each dimension separately and, if appropriate, an overall summary estimate may be obtained. The model also allows the correlation between dimensions to be estimated. Multilevel models are flexible in

that they can cope with situations where some of the dimension scores may be missing for some patients. However, the application of multilevel models to quality-of-life data in general is problematic since the method assumes the missing data mechanism is ignorable, which is not generally true of quality-of-life data.

## Survival data

### Nature of survival data

Measurements of the time between two events, an initial occasion and an endpoint of interest, are known as survival data. In a clinical trial, the initial occasion will usually be defined as the same event for all individuals, such as date of randomisation, date of starting treatment or date of diagnosis. The endpoint of interest will depend on the nature of the disease and treatments under investigation but is often death or relapse from a period of disease remission.

Survival data are different from other types of continuous data because over the period of a study the endpoint of interest is not necessarily observed in all subjects. This may occur because:

- (a) some patients are lost to follow-up, that is, they are not followed to the end of the study and, when last seen, have not experienced the event of interest, or
- (b) the event has not occurred in some patients by the time the study closes for analysis.

Such data are referred to as censored survival times and are different from missing data in that they provide a lower bound for the actual non-observed survival times. Any analysis carried out on survival data should use statistical methods that do not disregard censored data and, indeed, make the fullest possible use of it to avoid loss of information.

This study is specifically interested in the analysis of quality-of-life and survival data where survival generally means time to death. Thus in the context of this report, unless specified otherwise, survival will refer to time to death from a fixed origin, typically randomisation in a clinical trial. Most findings, however, will also be applicable to survival defined in other ways such as, for example, time from randomisation to relapse or death (whichever occurs first).

### Informative censoring

Most analytical methods used for survival data with censored observations are only valid if censoring is

non-informative. This means that the censoring is not related to any factors associated with the actual survival time, that is, the actual survival time,  $t$ , of an individual is independent of any mechanism which causes that individual's survival time to be censored at time  $c$ , where  $c < t$ .<sup>34</sup> When the censoring mechanism is not independent of survival time, informative censoring occurs and standard methods used for survival analysis are invalidated.

Informative censoring is a particular problem when analysing quality-of-life and survival data simultaneously and will be discussed in more detail in the context of each proposed statistical methodology.

## References

1. Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987;**40**:593–600.
2. Naji SA, Sheldon TA. The measurement of patient outcomes. *Proc R Soc Ed B Biol Sci* 1993;**101**:55–75.
3. World Health Organization. The constitution of the World Health Organization. *WHO Chron* 1947;**1**:29.
4. Schumacher M, Olschewski M, Schulgen G. Assessment of quality of life in clinical trials. *Stat Med* 1991;**10**:1915–30.
5. Fayers PM, Jones DR. Measuring and analysing quality of life in cancer clinical trials: a review. *Stat Med* 1983;**2**:429–46.
6. Aaronson NK. Quality of life assessment in clinical trials: methodologic issues. *Control Clin Trials* 1989;**10**:S195–208.
7. Barnett DB. Assessment of quality of life. *Am J Cardiol* 1991;**67**:C41–4.
8. Bowling A. Measuring health: a review of quality of life measurement scales. Buckingham: Open University Press; 1991.
9. Bowling A. Measuring disease: a review of disease-specific quality of life measurement scales. Buckingham: Open University Press; 1995.
10. Fallowfield L. The quality of life: the missing measurement in health care. London: Souvenir Press (E&A); 1990.
11. Campbell M, Gibbard N. Frequency of use of quality of life instruments. To be published.
12. Hays RD, Anderson R, Revicki D. Psychometric considerations in evaluating health-related quality of life measures. *Qual Life Res* 1993;**2**:441–9.
13. Fitzpatrick R, Fletcher A, Gore SM, Jones DR, Spiegelhalter DJ, Cox DR. Quality of life measures in health care. I: applications and issues in assessment. *BMJ* 1992;**305**:1074–7.

14. Chow SC, Ki FYC. Statistical issues in quality-of-life assessment. *J Biopharm Stat* 1996;**6**:37–48.
15. Chow SC, Ki FYC. On statistical characteristics of quality of life assessment. *J Biopharm Stat* 1994;**4**:1–17.
16. Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess* 1998;**2**(14).
17. Bosch JL, Hunink MGM. The relationship between descriptive and valualational quality-of-life measures in patients with intermittent claudication. *Med Decis Making* 1996;**16**:217–25.
18. Revicki DA, Kaplan RM. Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. *Qual Life Res* 1993;**2**:477–87.
19. Karnofsky D, Burchenal JH. Clinical evaluation of chemotherapeutic agents in cancer. In: Macleod CM, editor. Evaluation of chemotherapeutic agents. New York: Columbia University Press; 1949.
20. Little RJA, Rubin DB. Statistical analysis with missing data. New York: John Wiley; 1987.
21. Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Spiegelhalter DJ, Jones DR. Quality-of-life assessment: can we keep it simple? *J R Stat Soc A* 1992;**155**:353–93.
22. Pocock SJ. A perspective on the role of quality-of-life assessment in clinical trials. *Control Clin Trials* 1991;**12**:S257–65.
23. Nayfield SG, Ganz PA, Moinpour CM, Cella DF, Hailey BJ. Report from a National Cancer Institute (USA) workshop on quality of life assessment in cancer clinical trials. *Qual Life Res* 1992;**1**:203–10.
24. Fletcher A, Gore SM, Jones DR, Fitzpatrick R, Spiegelhalter DJ, Cox DR. Quality of life measures in health care. II: design, analysis and interpretation. *BMJ* 1992;**305**:1145–8.
25. Zhang J, Quan H, Ng J, Stepanavage ME. Some statistical methods for multiple endpoints in clinical trials. *Control Clin Trials* 1997;**18**:204–21.
26. Tandon PK. Applications of global statistics in analysing quality of life data. *Stat Med* 1990;**9**:819–27.
27. Olschewski M, Schumacher M. Statistical analysis of quality of life data in cancer clinical trials. *Stat Med* 1990;**9**:749–63.
28. O'Brien PC. Procedures for comparing samples with multiple endpoints. *Biometrics* 1984;**40**:1079–87.
29. Tandon PK, Stander H, Schwarz RP. Analysis of quality of life data from a randomized, placebo-controlled heart failure trial. *J Clin Epidemiol* 1989;**42**:955–62.
30. Pocock SJ, Geller NL, Tsiatis AA. The analysis of multiple endpoints in clinical trials. *Biometrics* 1987;**43**:487–98.
31. Goldstein H. Multilevel statistical models. 2nd ed. London: Edward Arnold; 1995.
32. Beacon HJ, Thompson S. Multi-level models for repeated measurement data: application to quality of life data in clinical trials. *Stat Med* 1996;**15**:2717–32.
33. Beacon HJ. The statistical analysis of self assessed quality of life data in cancer clinical trials [PhD thesis]. London School of Hygiene and Tropical Medicine, University of London; 1996.
34. Collett D. Modelling survival data in medical research. London: Chapman and Hall; 1994.



## Chapter 4

# Background to the illustrative example – the MIC study

### Introduction

The methods identified by the review and discussed in this report are illustrated by application to data from a previously conducted study, referred to here as the MIC study. This illustrative example is, in fact, the second of two concurrent Phase III trials of MIC (mitomycin, ifosfamide and cisplatin) chemotherapy which were conducted in patients with non-small cell lung cancer. The results from both trials are reported elsewhere<sup>1</sup> and this report focuses on the quality-of-life and survival data from the second trial (MIC2) only.

The general background to the trial is presented below, including the characteristics of patients in the quality-of-life study. The measurement of quality of life in the MIC study is described together with details of missing data.

### Background

In 1988, a randomised Phase III trial was initiated at the Cancer Research Campaign Clinical Trials Unit, Birmingham, to evaluate the role of chemotherapy in the treatment of non-small cell lung cancer in patients with extensive stage disease. Patients were randomly allocated to receive either standard palliative treatment (PAL arm), usually radiotherapy, or MIC chemotherapy, to a maximum of four courses, followed by palliative care (CT arm). The aim of the study was to compare treatments in terms of both survival and quality of life. Quality of life was an important endpoint in the study because both treatments were considered largely palliative and MIC chemotherapy was considered by some clinicians to be highly toxic.

The trial closed in March 1996, by which time 359 patients had been randomised into the study. For practical reasons associated with the availability of the research nurse, the quality-of-life component of the study was carried out only on a subset of trial patients, essentially consisting of patients treated at three main oncology centres. Quality-of-life data were collected for 109 patients from the trial, 67 on the CT arm and 42 on the PAL arm.

**TABLE 2** Characteristics of patients in the MIC quality-of-life study

Characteristic	CT arm (n = 67)	PAL arm (n = 42)
<b>Sex</b>		
Male	47 (70%)	32 (76%)
Female	20 (30%)	10 (24%)
<b>Age</b>		
Median	62	66
Range	42–75	49–75
Inter-quartile range	56–68	61–71
<b>Histology</b>		
Squamous	35 (52%)	26 (62%)
Adenocarcinoma	28 (42%)	10 (24%)
Large cell undifferentiated	2 (3%)	1 (2%)
Not known	2 (3%)	5 (12%)
<b>WHO performance status</b>		
0	20 (30%)	7 (17%)
1	29 (43%)	17 (40%)
2	17 (25%)	18 (43%)
Not known	1 (2%)	0 (0%)

The sex, age, histology and performance status of the patients on entry to the quality-of-life study are given in *Table 2*.

The palliative group, in comparison to the chemotherapy group, had a slightly greater proportion of males, was on average an older group and had a greater proportion of squamous tumours. In addition, the group appeared to have poorer performance status at baseline compared with the chemotherapy group. These differences in patient characteristics may need to be considered in any comparison of treatments.

### Measuring quality of life

#### Instrument

Quality of life was assessed using questionnaires completed by the patients with help of a dedicated quality-of-life research nurse. The questionnaire was designed specifically for the trial but was based on the EORTC QLQ-LC13, the lung cancer module

of the quality-of-life questionnaires designed by the European Organisation for Research and Treatment of Cancer (EORTC).<sup>2</sup>

The questionnaire consisted essentially of 11 questions evaluating specific physical and psychological aspects of a patient's quality of life, plus a more general question on malaise (see *Table 3*). All questions related to how the patient had been feeling over the previous 3 weeks. Responses were on a four-level ordered categorical scale ('None', 'A little', 'Quite a bit' and 'Very much') and for analytical purposes have been coded from 0 to 3, respectively.

The questionnaire also asked about the patient's performance status using WHO categories and asked how they felt in comparison to the time of the previous questionnaire; however, these questions will not be considered for the purposes of this report.

Two measures of quality of life from the MIC study, one ordinal and one continuous, are used here to demonstrate the various methodologies: these are, respectively, the malaise question (MAL) and the

mean quality-of-life score (MQS). The MQS was calculated from the answers to the 12 questions presented in *Table 3*. If a patient did not respond to all 12 items, then the mean was calculated as the mean of the reduced number of non-missing values. In total, 399 MQSs were calculated and 11% of these were calculated from incomplete data (see page 15).

### Timing of assessments

In the MIC study, the quality of life of patients was assessed only during their treatment period. There were several reasons for this decision, which was made when the study was designed:

- (i) the main aim of the quality-of-life part of the study was to assess both treatments in terms of their immediate impact on quality of life
- (ii) the survival time of patients in the study in general is short (median survival time for all patients in trial was 5.6 months) and thus long-term quality of life was not thought to be an issue when comparing treatments
- (iii) the collection of data during this period was most practicable in that the patients were attending clinic for treatment and were therefore readily accessible.

**TABLE 3** Questions from the MIC quality-of-life questionnaire

Item	Question
Cough	Do you have a cough?
Severe dyspnoea	Do you get breathless on mild activity like dressing?
Moderate dyspnoea	Do you get breathless when walking on the flat?
Mild dyspnoea	Do you get breathless on stairs or walking uphill?
Haemoptysis	Have you coughed blood?
Pain	How much pain are you getting?
Appetite	Have you noticed any loss of appetite?
Anxiety	Have you been worrying?
Depression	Have you been depressed?
Dysphagia	Have you any difficulty swallowing?
Nausea	Did you feel sick during or since your last treatment? (CT arm) Have you been feeling sick? (PAL arm)
Malaise	Have you been feeling generally ill?

The study was designed so that patients completed questionnaires on entry to the trial and then every 3 weeks thereafter, completing five forms in total on the CT arm and four forms in total on the PAL arm. In theory, all questionnaires should thus have been completed within 12 weeks of entering the trial; however, in reality the timing of the questionnaires varied considerably from this planned administration (see *Table 4*). On average, the baseline questionnaire was completed 1 week after entry to the trial and then subsequent questionnaires were completed every 3 weeks after this.

The analysis of the data needed to focus on the period during which quality-of-life data was collected and applicable. A cut-off point defining

**TABLE 4** Timing of questionnaires (in weeks) from date of entry to trial

	Q1	Q2	Q3	Q4	Q5
Planned	0	3	6	9	12
Median	0.9	4.0	7.2	10.3	13.0
Minimum	0	1.9	5.4	8.7	10.9
Maximum	6.7	9.0	26.0	27.0	23.3

this period needed to be established. Of the 399 completed questionnaires, 41 questionnaires were not completed within the 12 weeks from date of entry to the trial. If a 12-week cut-off was chosen for the analysis then the information from these questionnaires would be lost. Investigation of the distribution of times of the questionnaires indicated that 18 weeks from entry to trial would be the optimal cut-off value. This value enabled the majority of the questionnaires to be incorporated in the analysis, with information from only four (three participants) being lost, while at the same time not creating a situation in which individuals had large time spans with no quality-of-life information. It should be noted that 37 participants (34%) died during this 18-week period.

## Missing data in the quality-of-life study

The number of returned questionnaires in the MIC study diminishes with time. A reduction in numbers from one time point to the next results either from intermittent missing whole questionnaires or from individuals dropping out of the study (see *Table 5*). For example, with the second questionnaire on the CT arm, only 63 questionnaires from an expected 67 were returned; three were intermittently missing and one was missing because the patient had dropped out of the study, leaving 66 patients left in for the third questionnaire. These types of missing quality-of-life data were discussed in general in chapter 3 (see page 8). A third type of missing data, single missing items from otherwise complete questionnaires, was also

**TABLE 5** Number of returned questionnaires over time in the MIC study

	Q1	Q2	Q3	Q4	Q5	Total
<b>CT arm</b>						
Patients in study	67	67	66	53	46	
Returned questionnaires	67	63	52	45	38	265
<b>Reasons for missing questionnaires</b>						
Intermittent missing	0	3	1	1	0	5
Patient drop-out	0	1	13	7	8	29
<b>PAL arm</b>						
Patients in study	42	42	36	32	–	
Returned questionnaires	42	35	32	25	–	134
<b>Reasons for missing questionnaires</b>						
Intermittent missing	0	1	0	0	–	1
Patient drop-out	0	6	4	7	–	17

covered in this chapter and all three are discussed here in the context of analysis and interpretation of quality-of-life data from the MIC study.

## Single missing items

For some returned questionnaires in the study, responses were not given to all questions. The extent of this problem was not great (see *Table 6*), with 89% of questionnaires being completed fully and only 3% having more than one item missing. The maximum number of missing items on any one questionnaire, five out of 12, occurred just once. This type of missing data caused problems in calculating MQSs (see page 14) and when data were missing on a questionnaire, the calculation of the mean was based on the reduced number of non-missing items. This assumes that the missing data is ignorable, that is, the reason for it being missing is not related to the quality of life at that time.

**TABLE 6** Number of missing items from returned questionnaires

Number of missing items	Frequency (%)
0	356 (89)
1	33 (8)
2	8 (2)
3–5	2 (1)
<b>Total</b>	<b>399</b>

Because this report analyses the malaise item independently, the missing data for this specific item causes problems. Missing values for the MAL variable could be considered either as intermittent missing data or as drop-out data depending on the questionnaire from which the item was missing. The malaise item was missing on eight questionnaires. In four of these, the missing item could be classed as intermittent, in that MAL values were available on questionnaires both before and after that with the missing value. These missing values were tackled in the same way as if the whole questionnaire was missing (see below). The remaining four missing values occurred in the final questionnaires of three patients (two values were on consecutive questionnaires belonging to one participant) and, for the purposes of this report, were treated in the same way as missing data resulting from drop-out (see below). In general, the nature of this type of missing data is different from drop-out data since, with this type of data, the patient is still participating in the study; treating them as drop-outs is, thus, questionable.

### Intermittent missing whole questionnaires

In the MIC study, five patients had intermittent missing whole questionnaires, with one having two consecutive questionnaires missing. In the analysis of the data, the values from the questionnaire prior to the intermittent missing one were assumed to carry over until the questionnaire after the missing one. In this way intermittent missing questionnaires were effectively ignored. This assumes that the reason for them being missing was not related to the patient's quality of life at that time.

### Missing data resulting from drop-out

In the MIC study, a drop-out is defined as a participant who did not complete a final questionnaire (fifth on CT arm and fourth on PAL arm). Participants dropped out at various times during the study (see *Table 5*).

There are four categories of drop-out:

- drop-out directly caused by death (i.e. patient died within 3 weeks of last completed questionnaire before the next planned assessment)
- drop-out not due to death (i.e. patient was alive at next planned assessment) and patient died before the analysis cut-off time (i.e. within 18 weeks of entry to trial)
- drop-out not due to death (i.e. patient was alive at next planned assessment) and patient died after the analysis cut-off time (i.e. after 18 weeks from entry to trial)
- drop-out patient last known to be still alive (after 18 weeks).

In the MIC study, 46 patients dropped out of the study at some point. They were balanced across the two treatment arms: 43% (29/67) of patients on the CT arm and 40% (17/42) on the PAL arm. The type of drop-out, in terms of time of death relative to last returned questionnaire, appeared to be slightly different on the two arms (see *Table 7* and *Figures 2* and *3*), with almost half of the drop-

**TABLE 7** Categories of drop-out in the MIC study

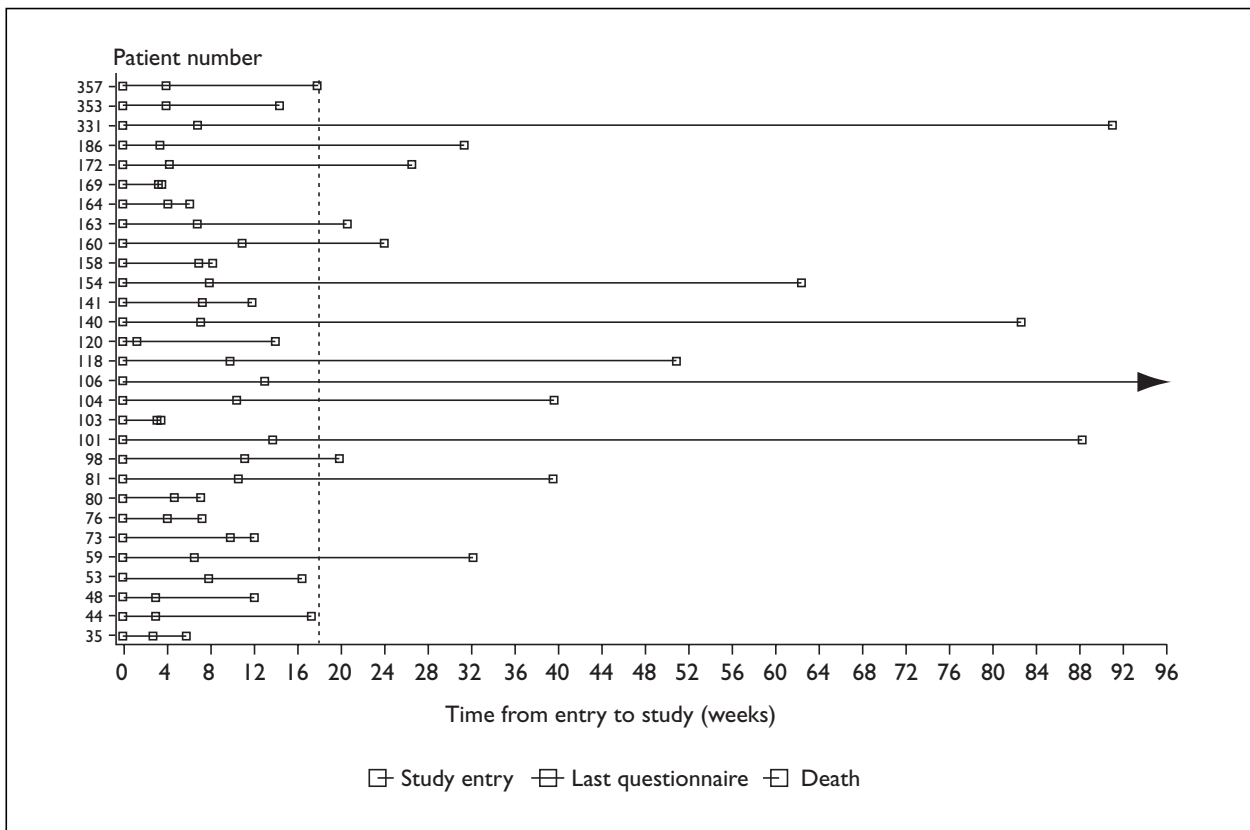
	CT arm	PAL arm	Total
Drop-out due to death	6	4	10
Drop-out not due to death but died within 18 weeks	9	10	19
Drop-out not due to death and died after 18 weeks	13	3	16
Drop-out last known to be alive	1	0	1
<b>Total</b>	<b>29</b>	<b>17</b>	<b>46</b>

outs on the CT arm (14/29) surviving the 18-week period from entry to trial but only three surviving the 18-week period of analysis on the PAL arm.

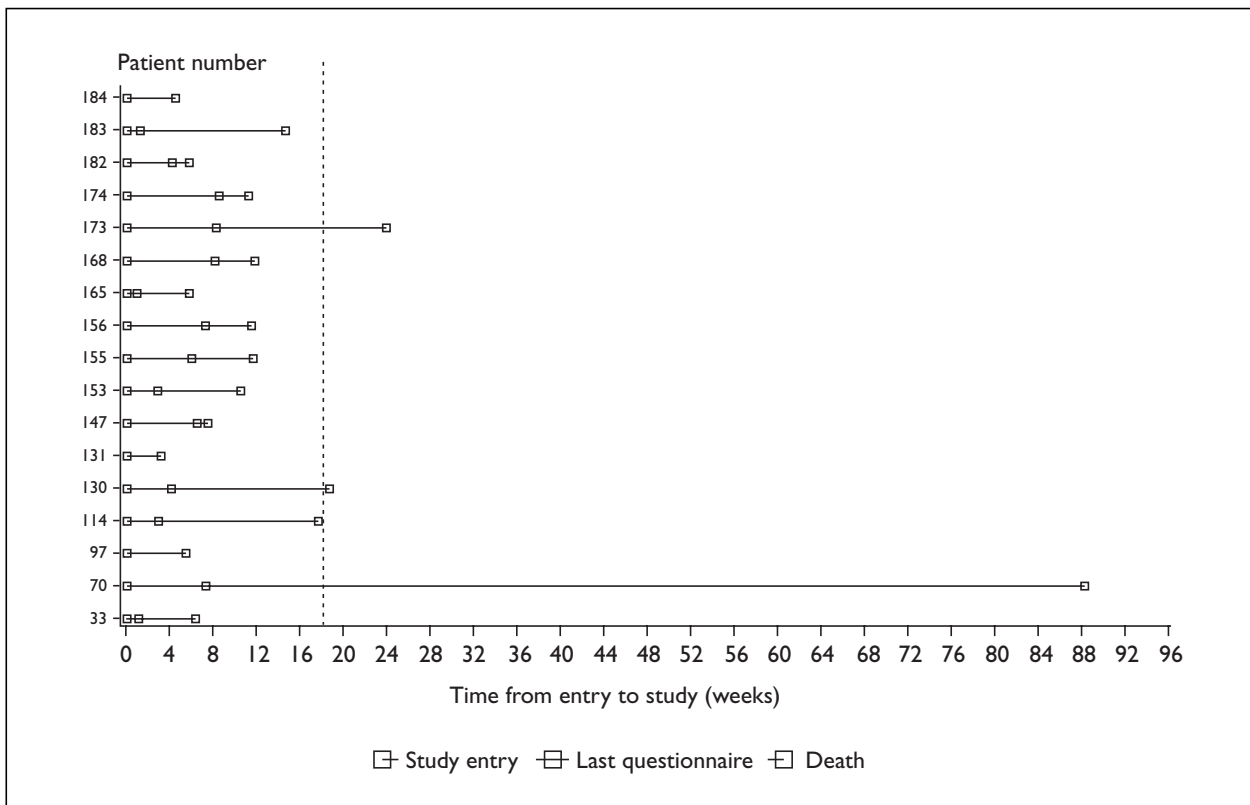
In some analyses of the MIC data, values for missing data caused by drop-out were imputed. Several approaches to imputing this missing data were considered (see *Table 8*). Each have their disadvantages and it was decided, in situations where imputation was necessary, to use the **last value carried forward** approach. The impact of different methods for imputation should be investigated in a sensitivity analysis.

### References

1. Cullen MH, Billingham LJ, Woodroffe CM, Chetiyawardana AD, Gower NH, Joshi R, *et al.* Mitomycin, ifosfamide and cisplatin in unresectable non-small cell lung cancer: effects on survival and quality of life. *J Clin Oncol*. Submitted 1999.
2. Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;**85**:365–76.



**FIGURE 2** Drop-outs on CT arm of the MIC study: timing of last questionnaire and death in relation to study entry time (NB: patient 106 was still alive when last seen, 186 weeks after study entry)



**FIGURE 3** Drop-outs on PAL arm of the MIC study: timing of last questionnaire and death in relation to study entry time (NB: patient 97 died on same day as last questionnaire; patients 131 and 184 completed only one questionnaire, on day of study entry)

**TABLE 8** Possible approaches to imputing missing data in the MIC study

Approach	Description	Problems
Last value carried forward	Assume drop-outs stay in the same health state, as measured at the last assessment, until date of death or censor date, whichever comes first.	<ul style="list-style-type: none"> <li>• If last recorded measure of QoL was 'good', it may be invalid to assume patient remains in a 'good' state until death.</li> <li>• If dropped out early, assumption of steady state needs to be applied to long period.</li> </ul>
Worst value carried forward	Assume drop-out caused by ill health and thus drop-outs move into the poorest health state (if they are not already in that state) from the time of drop-out onwards.	<ul style="list-style-type: none"> <li>• Choice of time at which patient is moved to 'poor' health state may not be obvious – time of next planned assessment could be used.</li> <li>• If drop-out is not caused by ill health, as may well be true for those patients who die a considerable time after dropping out, they may be allocated to 'poor' state too early.</li> <li>• Bias may result since assumption only applicable to those who dropped out while in 'good' state.</li> </ul>
Linear decrease over time	Assume QoL decreases linearly from time of drop-out until death.	<ul style="list-style-type: none"> <li>• Bias may result since assumption only applicable to those patients who dropped out while in 'good' state.</li> <li>• Assumption of linearity may be questioned.</li> <li>• Method difficult to apply to individuals with censored survival times.</li> </ul>

## Chapter 5

# Quality-of-life analysis in the presence of informative drop-out: introduction

In some studies, the main criterion by which health technologies are assessed is quality of life. When the disease under study is potentially fatal, such as cancer, survival will also be an issue, whereas with long-term chronic diseases, such as arthritis, this will not be the case. This review considers situations where both quality of life and survival are of interest. Quality of life is usually the main endpoint of interest in cancer clinical trials investigating treatments which are purely palliative in nature and, within this context, the relief of symptoms is weighed against the adverse effects of treatment. It will also be the main endpoint in trials of new treatments which have been developed to reduce treatment toxicity compared with standard treatments, without any detrimental effect on survival. In these types of studies, survival is an issue but mainly through the problems it causes in interpreting the quality-of-life endpoint.

In studies of seriously ill patients who are near to death, the collection of quality-of-life data over time is problematic. Patients drop out of the study because of illness or death and this results in patient attrition, that is, a reduction in the number of subjects in the quality-of-life study over time. The drop-out process is related to the health status of the patient and is therefore likely to be informative (see page 9); thus the missing data mechanism in such studies is non-ignorable. Ironically, the situations in which quality of life is most likely to be the main endpoint, such as clinical trials of palliative treatments for cancer patients, are the situations in which the problem may well be most serious.

The aim of this part of the report (chapters 5–9) is to discuss methods for analysing longitudinal quality-of-life data. Although the methods discussed here relate to longitudinal data collected over three or more time points, which may be fixed or varying, some could be satisfactorily and usefully

applied to data sets with measurements at just two time points. However, in situations where data is collected at two time points, the analysis is usually in terms of change in quality of life and, as with data collected at just one time point, standard non-longitudinal statistical methods can be used. Informative drop-out may still cause problems in these circumstances and cautious interpretation of the results is needed.

The analysis of quality-of-life data should begin descriptively to give the investigator and reader insight into the data. The methods that can be used to describe quality-of-life data over time are discussed in chapter 6, where the problems of interpretation caused by informative drop-out are highlighted and methods that attempt to account for the problems are described.

Quality-of-life data collected over time are often analysed using standard methods of longitudinal data analysis. These methods all assume that the missing data mechanism is ignorable. This could be appropriate in studies in which survival is not an issue but, where there is informative drop-out, analysis using these longitudinal methods could lead to invalid conclusions. Although the focus of this report is on the analysis of quality-of-life data in situations in which survival is also an issue, chapter 7 is dedicated to describing some of the standard methods of longitudinal data analysis. Despite the general inappropriateness of these methods in situations of informative drop-out, the methods are described briefly with a discussion of their potential problems and possible use with quality-of-life data. In chapter 8 methods for modelling longitudinal quality-of-life data are described which attempt to deal with the problem of informative drop-out. A summary and discussion of quality-of-life analysis in the presence of informative drop-out is presented in chapter 9.





## Chapter 6

# Descriptive methods for the analysis of longitudinal quality-of-life data

### Introduction

The interpretation of longitudinal quality-of-life data can be difficult and an initial exploratory analysis often gives an insight into the data before any formal testing or modelling is carried out. Descriptive methods of analysis do not have the problems associated with multiple testing and, hence, the quality-of-life data can be explored as extensively as desired. The variety of ways of exploring the data both at the patient level and the treatment group level are discussed below. In general, the clearest way to describe data is graphically but, at the group level, it may be preferable to tabulate the results. Descriptive methods for analysing quality-of-life data have been reviewed elsewhere.<sup>1,2</sup>

### Patient profiles

The examination of each patient's quality-of-life data over time individually can be very helpful. It may reveal a consistent pattern across patients and will highlight errors, outliers and patterns of missing data.

Patient profiles can be examined by plotting individual patient scores over time. This technique is most useful for continuous data but may also be used for ordinal data. Carlens and colleagues<sup>3</sup> and Nou and Aberg<sup>4</sup> plotted each patient's quality of life over time and called them **vitagrams**. Profiles for each patient can be plotted as a set of mini-graphs or they can be overlaid on one graph. Individual profiles can be overlaid for each treatment group separately and then used to compare patterns within each treatment group.

Individual patient profiles of MQSs over time are plotted and overlaid for each treatment group in the MIC study (see *Figures 4 and 5*). These show that the data vary widely at all time points and no obvious patterns, in terms of change in quality of life over time, are apparent.

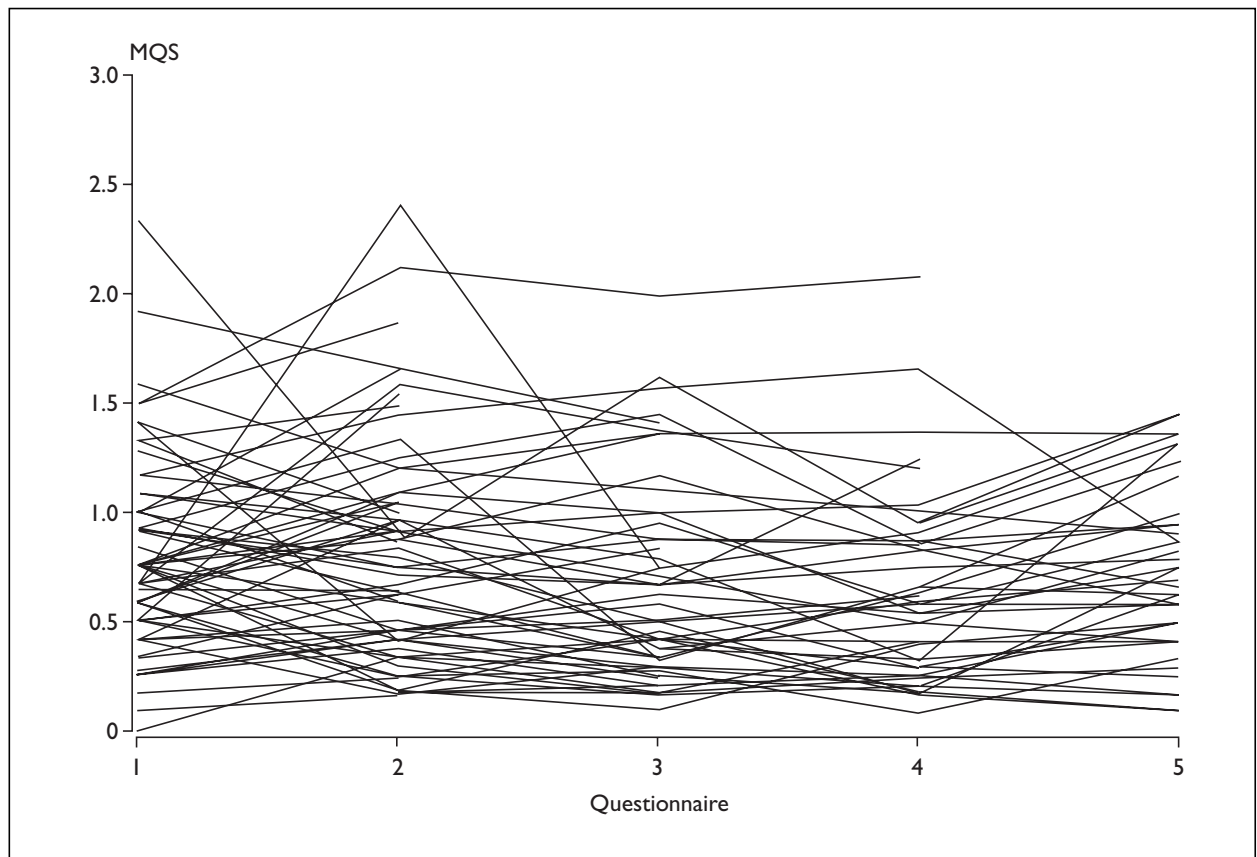
The main problem with patient profile data is that it is often impractical to display data from

large numbers of patients. One solution is to plot a simple random sample of the patients in a study,<sup>2</sup> another may be to categorise and plot the data on a Lexis diagram, as described below.

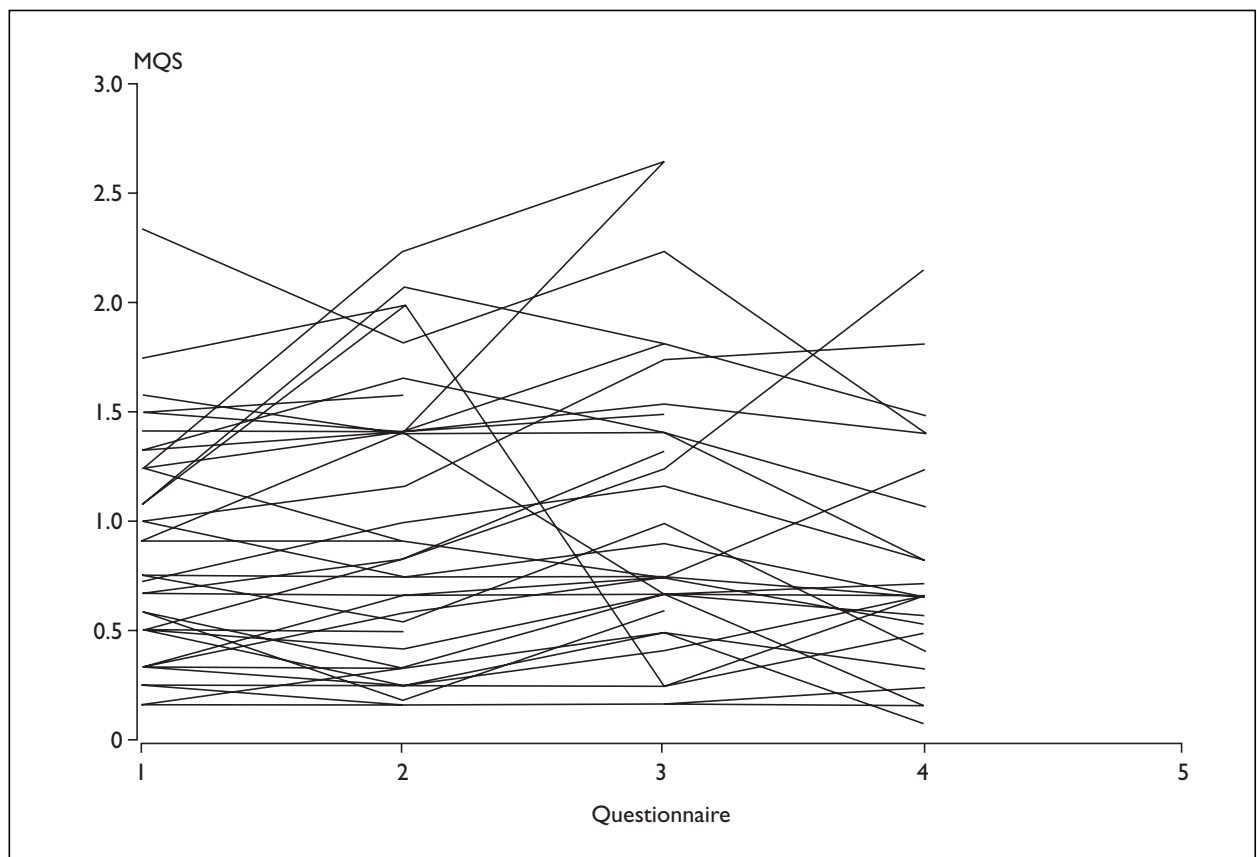
Lexis diagrams, or variations of such (as used for the MIC data in *Figures 6 and 7*), describe individual patient data and allow data for reasonably large numbers of patients to be displayed at once.<sup>2</sup> Each patient profile is represented by a line over time which is solid or dotted depending on the value of a binary quality-of-life variable at that time. The patient profiles are usually ordered by date of entry to study. If desired, a Lexis diagram can be extended to incorporate a higher level categorical variable by using several different types of line but the diagram becomes difficult to interpret as the number of categories increases. The technique is most appropriate for binary data but can also be used for ordinal data with few categories. Continuous and ordinal level data can be adapted for a Lexis diagram by grouping the values into, ideally, two levels.

Patient profiles showing time spent with malaise (MAL = 1, 2 or 3) or without malaise (MAL = 0) were plotted for the MIC study (see *Figures 6 and 7*) by assuming the value of MAL measured at each questionnaire point carries over to the next assessment. Values at the last questionnaire (fifth for CT group and fourth for PAL group) were carried over to the time when the next assessment would have taken place. The graphs do not show any clear patterns in terms of change in quality of life over time.

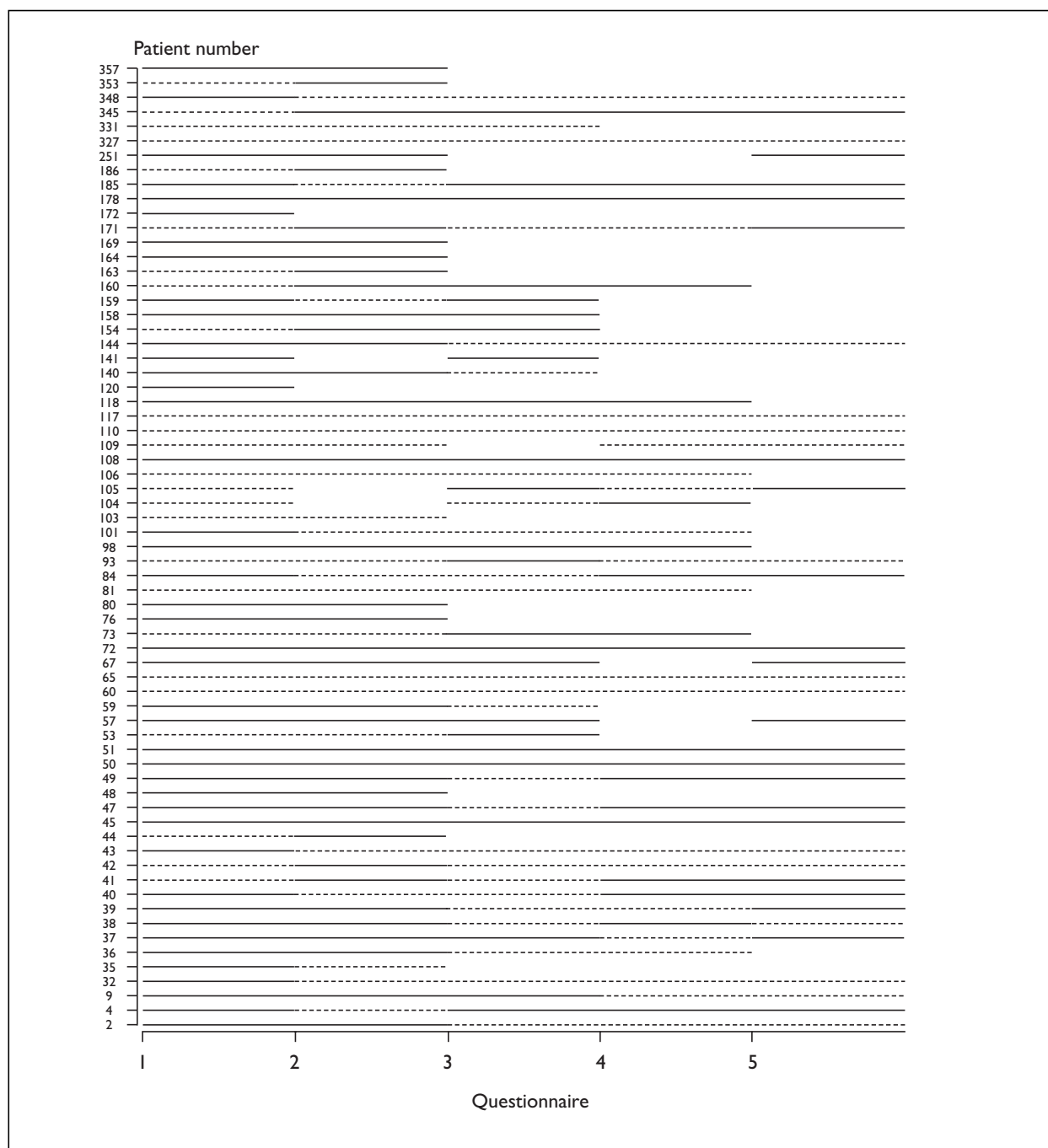
Plots of individual patient profiles may reveal patterns of missing data. Intermittent missing data will be apparent as gaps in the lines, while a profile ending early will signify a patient drop-out. If a patient dies during the course of the quality-of-life study then this can be plotted as a symbol at the appropriate time. Patient profiles can be grouped and overlaid according to differing lengths of follow-up and also, possibly, differing reasons for shortened follow-up, giving a possible insight into the association of the drop-out process with previous quality of life (see page 25 for further details on this approach relating to group profiles).



**FIGURE 4** Patient profiles of MQS over time for the CT arm of the MIC study



**FIGURE 5** Patient profiles of MQS over time for the PAL arm of the MIC study



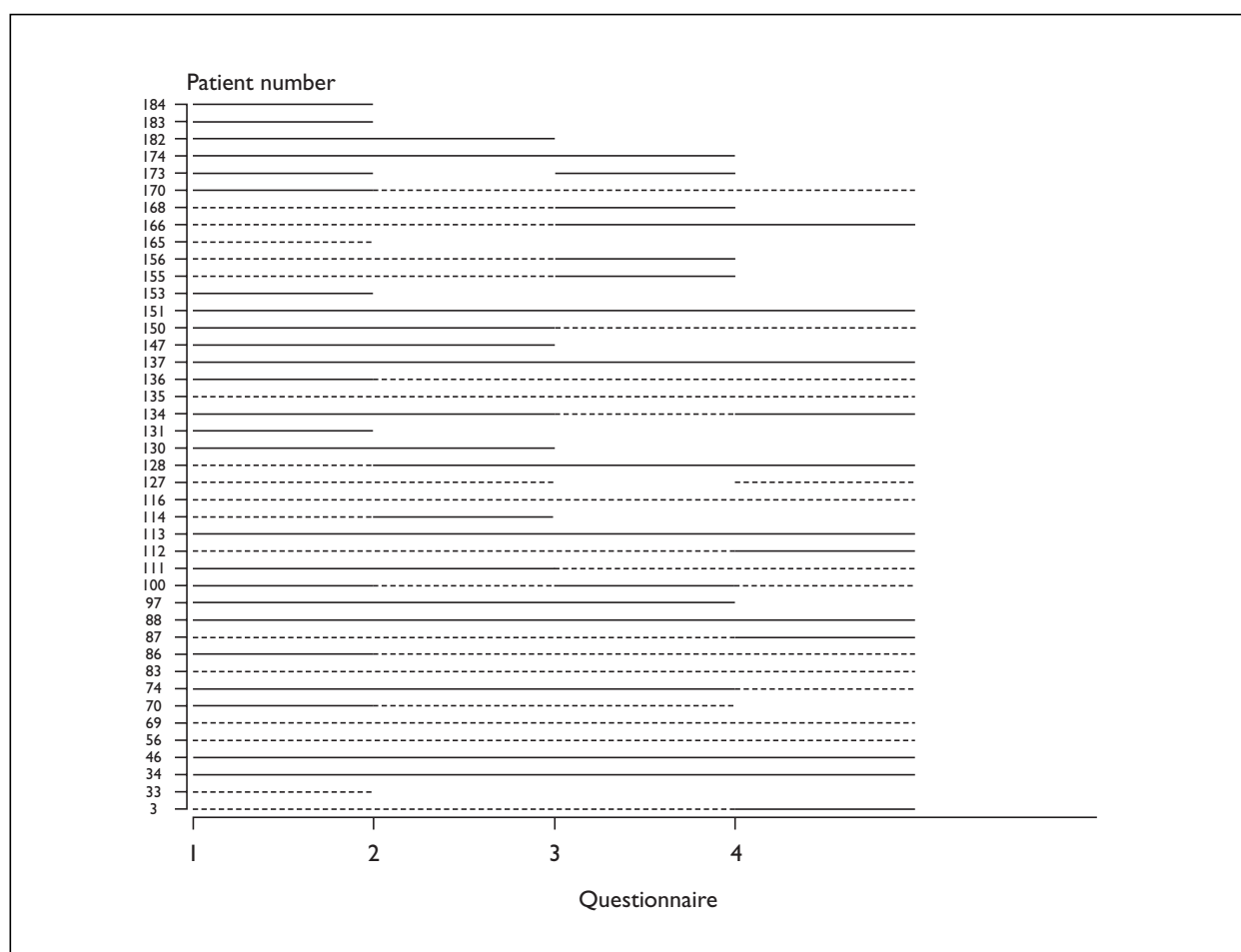
**FIGURE 6** Patient profiles of malaise over time for the CT arm of the MIC study (—, malaise; - - -, no malaise)

## Group profiles

After examining individual patient profiles of quality of life, it is necessary to summarise the experience of patients in each treatment group to enable the treatments to be compared more clearly in relation to their effect on quality of life over time. Plots of group profiles over time, overlaid on the same graph, enable a clear comparison of groups. Informative drop-out can cause problems in interpretation and possible approaches for analysis are discussed below.

## Summary measures to plot

There are two main types of group summary statistic that can be plotted over time, an average measure of quality of life or a proportion with a certain level of quality of life. The choice depends on the type of quality-of-life measure being summarised. In situations where the population is not stable over time, there are several choices regarding which subjects to include in the calculation of the summary measure at each time point and this is discussed below, followed by examples from the MIC study.



**FIGURE 7** Patient profiles of malaise over time for the PAL arm of the MIC study (—, malaise; - - -, no malaise)

### Averages

The mean or median quality-of-life scores in each group can be plotted over time (examples of means are given in O'Brien, *et al.*, Figure 3<sup>5</sup> and in Anderson, *et al.*, Figure 1;<sup>6</sup> examples of medians are given by the MRC Lung Cancer Working Party, 1996, Figure 2<sup>7</sup>). Bars at each time point representing 95% confidence intervals (CIs) for the mean or median should be included. This is the most useful way to represent continuous data. For ordinal data, the use of means is not theoretically correct but may be considered if the ordinal scale is long.

### Proportions

If the data are binary then the proportion with a symptom or side-effect can be plotted over time. Bars at each time point representing 95% CIs for the proportion should be included.

If the data are ordinal or continuous then the proportion reaching or exceeding a certain level of quality of life over time can be plotted (for example, see Figures 2 and 3 in the report of the MRC Lung Cancer Working Party, 1992,<sup>8</sup>

Figure 2 in their 1991 report,<sup>9</sup> and Figure 3 in their 1996 report<sup>7</sup>). Again, bars representing 95% CIs for the proportion should be included at each time point. Detail on the severity of the symptom or side-effect will be lost.

### Inclusion criteria for plot

Summarising the quality of life of patients in each treatment group over time is complicated by the problem of missing data, especially when informative drop-out occurs. Cautious interpretation of the data is needed in these situations. There are several choices regarding which subjects to include in the calculation of the summary measure over time. Summary measures can be calculated and plotted either for the whole sample of patients in the study, or for the subgroup of patients who completed all assessments, or for subgroups of patients according to different lengths of follow-up and reasons for drop-out.<sup>10,11</sup>

### All patients

If the whole sample of patients is used then, because of drop-out due to death and illness,

the sample reduces over time and this should be made clear in any graph or table by specifying the number contributing to the descriptive measure at each time point. Interpretation of such data needs caution because a comparison of results at different time points compares different groups of patients. Subjects contributing to summary measures at later time points are likely to be the 'healthier' members of the original group. In situations in which quality of life is measured on a 0–1 scale, where 0 represents quality of life equivalent to death, a more realistic description may be obtained by including patients who drop out because of death, which gives them a quality-of-life score of 0.<sup>6</sup>

### Complete case analysis

If the analysis is restricted to just those patients with complete data, then not only will it greatly reduce the sample size but it will also result in biased estimates, unless missing data is missing completely at random. Those included are not likely to be a representative subset of the overall sample, because patients who complete all assessments must be survivors and compliers and, hence, are probably a 'healthy' subgroup of the overall sample. The longer the follow-up period, the greater the level of attrition and the more reduced, and possibly more biased, the sample for analysis will be.

### Subgroups according to differing lengths of follow-up and reasons for drop-out

Patients can be split into subgroups according to differing lengths of follow-up. Subgroups can be formed according to number of completed assessments,<sup>10</sup> or a more detailed break-down based on different reasons for drop-out, such as censoring, death or lost to follow-up, could be used.<sup>11</sup> Quality of life over time can be compared across subgroups to establish its association with the drop-out process. If there are no obvious differences between the subgroups then it may be valid to combine the data for analysis.<sup>10</sup> A quality-of-life study needs a large number of participants for subgroups to contain adequate numbers of subjects for this type of analysis. This type of approach is used by de Stavola and Christensen<sup>12</sup> in their method for dealing with informative drop-out in longitudinal data modelling (see chapter 8).

### Group profiles for the MIC data

#### All patients

Summary measures for MQS and MAL in the MIC study were calculated for all patients in the study. The numbers of patients contributing to the summary measure at each time point have been specified in the graphs to highlight the

problems of informative drop-out and aid cautious interpretation.

Treatment group means together with their associated 95% CIs are plotted over time for MQS (see Figure 8). A possibly naïve interpretation of these plots, ignoring the effect of missing data, may be as follows. MQS reduces slightly (i.e. quality of life improves) over the first three time points in the CT group and increases slightly (i.e. quality of life deteriorates) in the PAL group, with 95% CIs suggesting that the groups significantly differ at the third time point. After this point, the curves move back towards each other again and quality of life in both treatment groups seems comparable.

The proportions of patients with malaise (i.e. MAL = 1, 2 or 3) are plotted over time for each treatment group in the MIC study (see Figure 9). Bars representing 95% CIs for the proportions are included. This shows the extent of malaise in the two groups is fairly high (at about the 60% level) and reasonably comparable over time. In both groups, the extent of malaise reduces initially but then increases again. As with MQS, this interpretation could be questionable since it ignores the effect of patient drop-out.

### Complete case analysis

Plots of MQS and malaise were plotted over time only for those patients who completed the first four questionnaires (see Figures 10 and 11). For MQS, only 43 patients on the CT arm had complete data, with 21 dropping out before completing the fourth

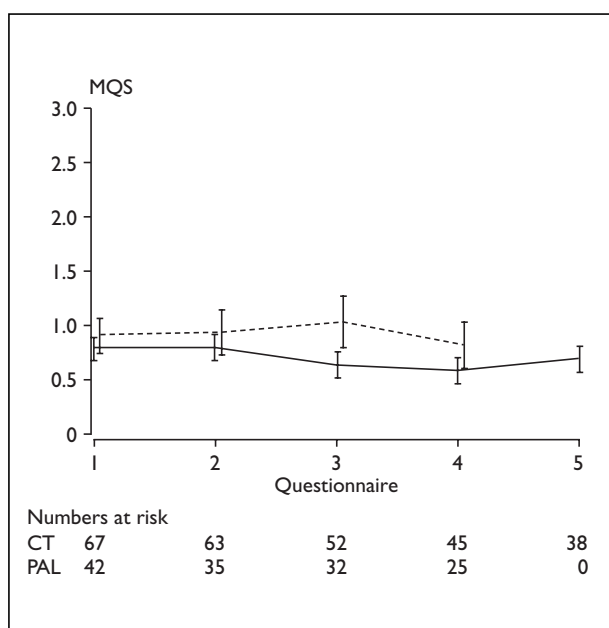
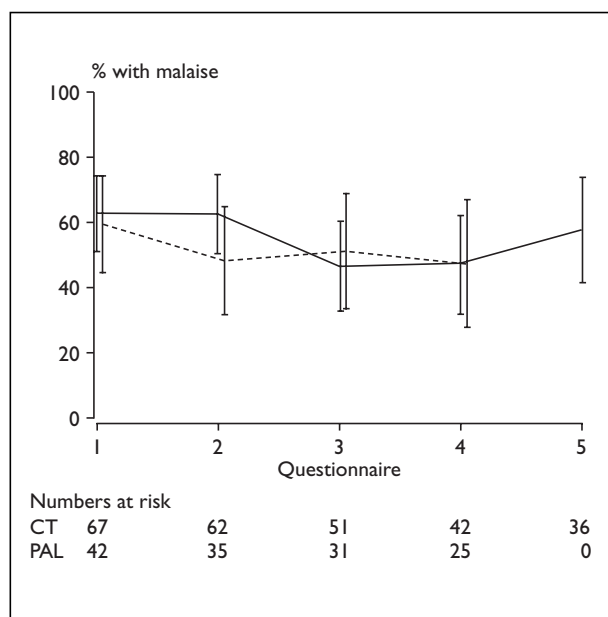


FIGURE 8 Treatment group means (and 95% CIs) over time for MQS in the MIC study (—, CT; - - -, PAL)

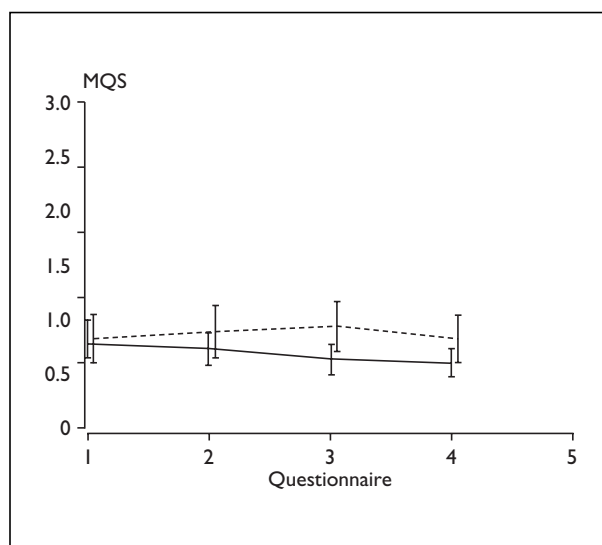


**FIGURE 9** Percentage of patients with malaise (and 95% CIs) over time for each treatment group in the MIC study (—, CT; ---, PAL)

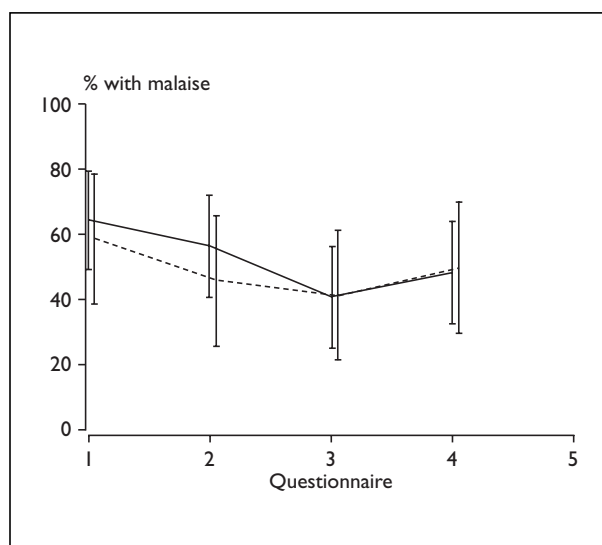
questionnaire and three having intermittent missing questionnaires, while on the PAL arm, 25 had complete data, with 17 dropping out. For malaise, a further four patients on the CT arm and one on the PAL arm were excluded because they had missing values for malaise on one of the first four questionnaires. These plots contain data for those patients who did not drop out of the study and are thus likely to represent the ‘healthier’ patients from the MIC study. The plots, however, are not dissimilar to those for all patients (see *Figures 8 and 9*).

#### Subgroups according to differing lengths of follow-up

The patients in the MIC study (from both treatment arms) were split into subgroups according to the number of the last completed questionnaire (see *Table 9*). The mean MQS for each subgroup was plotted over time (see *Figure 12*). This shows that quality of life in those who dropped out after one questionnaire was worst at baseline and, in those who dropped out after two or three questionnaires, quality of life over time deteriorated slightly before dropping out. Quality of life in those who did not drop out during the first four questionnaires appears to be reasonably stable. Although the plot needs to be interpreted with caution, because the numbers in some of the subgroups are small, it suggests that there may be some association between drop-out and quality of life. This means that for a valid analysis of quality-of-life data, the drop-out process should probably not be ignored.



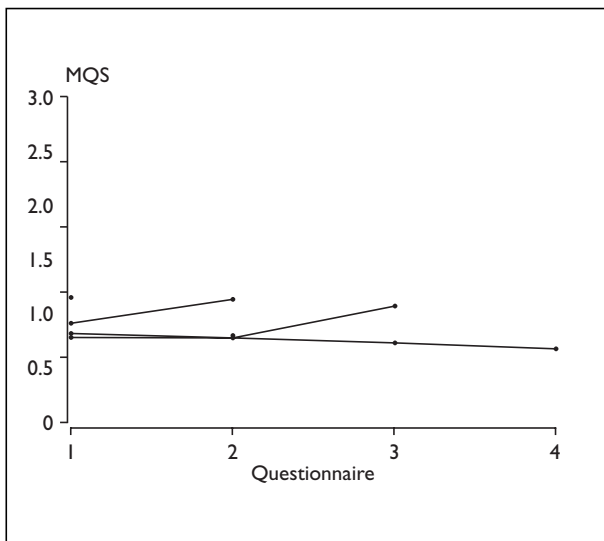
**FIGURE 10** Treatment group means (and 95% CIs) over time for MQS in MIC study patients with complete data (—, CT (n = 43); ---, PAL (n = 25))



**FIGURE 11** Percentage of patients with malaise (and 95% CIs) over time for patients with complete data in each treatment group in the MIC study (—, CT (n = 39); ---, PAL (n = 24))

**TABLE 9** Number of patients in the MIC study in each follow-up subgroup

Number of last completed questionnaire	Number of patients
1	7
2	17
3	14
4 or 5	71



**FIGURE 12** MQS for all patients in the MIC study: means over time for subgroups with differing periods of follow-up

## References

1. Fayers PM, Jones DR. Measuring and analysing quality of life in cancer clinical trials: a review. *Stat Med* 1983;**2**:429–46.
2. Beacon HJ. The statistical analysis of self assessed quality of life data in cancer clinical trials [PhD thesis]. London School of Hygiene and Tropical Medicine, University of London; 1996.
3. Carlens E, Dahlstrom G, Nou E. Comparative measurements of quality of survival of lung cancer patients after diagnosis. *Scand J Respir Dis* 1970;**51**:268–75.
4. Nou E, Aberg T. Quality of survival in patients with surgically treated bronchial carcinoma. *Thorax* 1980;**35**:255–63.
5. O'Brien BJ, Buxton MJ, Ferguson BA. Measuring the effectiveness of heart-transplant programs – quality of life data and their relationship to survival analysis. *J Chronic Dis* 1987;**40** (suppl 1):S137–53.
6. Anderson JP, Kaplan RM, Schneiderman LJ. Effects of offering advance directives on quality-adjusted life expectancy and psychological well-being among ill adults. *J Clin Epidemiol* 1994;**47**:761–72.
7. MRC Lung Cancer Working Party. Randomized trial of palliative two-fraction versus more intensive 13-fraction radiotherapy for patients with inoperable non-small cell lung cancer and good performance status. *Clin Oncol* 1996;**8**:167–75.
8. MRC Lung Cancer Working Party. A Medical Research Council (MRC) randomised trial of palliative radiotherapy with two fractions or a single fraction in patients with inoperable non-small cell lung cancer (NSCLC) and poor performance status. *Br J Cancer* 1992;**65**:934–41.
9. MRC Lung Cancer Working Party. Assessment of quality of life in small-cell lung cancer using a daily diary card developed by the Medical Research Council Lung Cancer Working Party. *Br J Cancer* 1991;**64**:299–306.
10. Hopwood P, Stephens RJ, Machin D. Approaches to the analysis of quality of life data: experiences gained from a Medical Research Council Lung Cancer Working Party palliative chemotherapy trial. *Qual Life Res* 1994;**3**:339–52.
11. Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Spiegelhalter DJ, Jones DR. Quality-of-life assessment: can we keep it simple? *J R Stat Soc A* 1992;**155**:353–93.
12. de Stavola BL, Christensen E. Multilevel models for longitudinal variables prognostic for survival. *Lifetime Data Anal* 1996;**2**:329–47.





## Chapter 7

# Standard methods for the analysis of longitudinal data

### Introduction

The methods discussed in this chapter are standard approaches to analysing longitudinal data and all assume that the mechanism giving rise to missing data is ignorable. The application of these methods to longitudinal quality-of-life data is problematic because, in this situation, the missing data are generally a result of informative drop-out and, hence, the missing data mechanism is non-ignorable. Methods discussed here should therefore be applied with caution to quality-of-life data, as they may result in invalid conclusions.

For continuous measures, such as those derived from a linear analogue scale or an aggregated global score, the appropriateness of parametric or non-parametric methods depends on the normality of the data distribution. It could be argued that since the continuous measures have truncated distributions, parametric methods will always be inappropriate.<sup>1</sup> It may be possible to normalise a distribution using a transformation, such as log or square root, and then compare treatments using a parametric approach on the transformed variable.<sup>2,3</sup>

### Analysis of each time point separately

If quality of life is assessed at fixed time points during the study then the data can be analysed by considering each time point separately. At each time point, there may be three different research questions of interest.

1. To what extent do treatment groups differ in terms of quality of life at that time point?
2. To what extent do treatment groups differ in terms of change in quality of life relative to a previous time point (usually baseline)?
3. Is the change in quality of life between the given time point and a previous time point statistically significant?

All these research questions can be tackled using simple, standard statistical methods. The method of analysis depends on the type of variable used to measure quality of life (binary, ordinal, continuous non-normal or continuous normal) and the number of treatment groups being compared. Examples of quality-of-life studies, in which the data are analysed at each time point can be found in the literature.<sup>1,2,4</sup>

There are three major problems with this simple approach to longitudinal data analysis:

- fixed time points are required
- the longitudinal nature of the data is ignored
- multiple analyses are involved, which may result in differences being regarded as significant by chance.

In addition to these problems, the approach is not generally suitable for quality-of-life data since it will not overcome the problem of informative drop-out. The patients at each time point will be a subset of the original sample of patients recruited to the study and, the further from baseline the time point, the more biased the sample will be, in that they will be the most 'well' patients from the original group.

### Summary measures

Using summary measures, as promoted by Matthews and colleagues,<sup>5,6</sup> is the simplest method for analysing longitudinal data. It reduces the repeated measures over time on an individual to a single summary measure, which can then be analysed using standard statistical methods, for example, for treatment group comparisons. The choice of summary measure needs to be clinically meaningful and will depend on the nature of the measure under investigation, together with the disease and treatment under study. In quality-of-life studies, the calculation of summary measures in patients who drop out may be difficult and if the drop-out process is informative, the analysis of the summary measure may give biased results.

### **Simple summary measures for longitudinal quality-of-life data**

A wide range of summary measures could be chosen to represent longitudinal quality-of-life data. For example, the maximum score reached over time may be appropriate, or change in quality of life between two time points. Alternatively, the slope representing change over time for each individual could be computed.<sup>7</sup>

Summary measures may be difficult to calculate when informative drop-out is present in the data; however, it may be possible to accommodate drop-outs into the analysis by imputing appropriate quality-of-life values. For example, in situations where quality of life is measured on a 0–1 scale, with 0 representing quality of life equivalent to death, time points with missing quality-of-life data caused by death may be allocated values of 0. This enables patients who drop out of the study because of death to be included in the analysis<sup>7</sup> and may reduce bias caused by informative drop-out.

A standard summary measure used for longitudinal data is **area under the curve**. For each individual, the area under the curve of their repeated quality-of-life measures over time can be calculated and used as a summary measure.<sup>7,8</sup> If patients do not have full follow-up then their summary measure will be censored, suggesting that methods of survival analysis may be appropriate. Censoring, however, will be informative and standard methods will give biased results. An alternative estimator has been proposed but is still biased.<sup>9</sup> An outcome measure that combines quality-of-life and survival data, the QALY, is a special form of this summary measure and is discussed in detail later (see chapters 15 and 16).

Simple binary indicator summary measures can be used to define palliation.<sup>10–12</sup> For example, an indicator may be set to show whether an individual experienced a decrease in the level of severity of a particular symptom or in overall quality of life at any time compared with baseline or, alternatively, whether the individual experienced total disappearance of a symptom at any time. Other summary measures include duration of palliation and percentage of patient survival time during which there was palliation.<sup>11,12</sup> These summary measures will only be valid if all patients have died, otherwise they should be restricted to a set follow-up time for which all patients in the study have been followed. Again, calculation of these summary measures may be problematic in the presence of informative drop-out.

### **Time to the occurrence of a quality-of-life related event**

Longitudinal quality-of-life data can be summarised by a single value representing the time to the occurrence of a quality-of-life related event. The most widely used clinical endpoint, which in some situations will be quality-of-life related, is relapse-free survival, that is, the time from study entry to disease relapse. Nabholz and colleagues<sup>13</sup> used time to first occurrence of an important clinical adverse event or disease progression as a quality-of-life oriented endpoint. Rosenman and Choi<sup>14</sup> used the Karnofsky index as a global measure of quality of life and used the time until first occurrence of a Karnofsky index of less than 60 as the summary measure on which to compare treatments. Hopwood and colleagues<sup>15</sup> suggest time to first improvement or time to first worsening of quality of life from baseline as a summary measure, with times for patients not achieving such targets being censored. Time to palliation of various symptoms, in those patients with the symptom present pre-treatment, has also been used as a summary measure for treatment comparison.<sup>16</sup>

Once a quality-of-life oriented endpoint has been defined, standard survival analysis techniques (see chapter 11) can be used to analyse the data. One advantage of this method is that patients who do not achieve the endpoint because of death or censoring would still be included in the analysis as a censored data point, thus dealing with the problem of informative drop-out. However, because the censoring mechanism may be related to the 'survival' time, standard survival analysis techniques may be invalidated by informative censoring (see page 11). Another advantage of this summary measure is that it can be used in situations where quality of life has been assessed at varying time points. However, if quality of life is assessed at only a few widely spaced time points, then the summary measure will be very crude.

A problem with this approach is that the potential for change, whether in terms of improvement or worsening, depends on the baseline value. For example, the worse a patient is at baseline, the greater the potential for improvement, so patients who do not experience a symptom at baseline are not able to improve and are thus excluded from the analysis. This will only be a problem in the comparison of treatments if the treatment groups differ with respect to baseline symptoms. The fact that a treatment may prevent a symptom from starting may be important and would not be assessed by this type of endpoint.

## Repeated measures analysis of variance

A conventional repeated measures analysis of variance<sup>17</sup> may be an appropriate statistical model to use in situations where quality of life is measured at fixed time points and will be most appropriate when the number of time points is reasonably small. The method assumes the data to be normally distributed, with the covariance matrix the same for all treatment groups. Generalisations on this simple model are possible.<sup>18</sup> Using the model it can be determined if quality of life changes over time and if treatment groups vary in terms of changes in quality of life, while taking account of within-patient correlations.

Patient drop-out in quality-of-life studies results in missing measurements and causes the data to be unbalanced. One strategy to enable the use of repeated measures analysis of variance in this situation is to undertake a **complete case analysis**, in which treatment comparisons are made only in patients with complete data. Problems of reduced numbers for analysis and potential bias make this approach problematic and so an alternative strategy, which allows all available data to be included, may be preferable. This latter approach is based on the assumption that missing data are missing at random and, if this is not the case, then a method of analysis that allows for informative drop-out needs to be considered (see chapter 8).

### Complete case analysis

A complete case analysis will not generally be applicable to quality-of-life data since it is only valid if missing data are missing completely at random. If there are very few subjects with complete data then a complete case analysis would be based on a very small sample. Even if the reduced sample is reasonably large, it is not likely to be representative of the whole sample. Those patients who survived for the duration of the study and were well enough throughout to complete questionnaires would generally be the 'healthier' participants and thus estimates of mean quality of life over time are likely to be biased (upwards). If the change in quality of life over time is independent of the level of quality of life, then, although estimates at each time will be biased, the estimates of the change may be unbiased.

In situations where quality of life is measured on a 0–1 scale, where 0 represents quality of life equivalent to death, patients who drop out because of death can be included in the analysis by giving

them a quality-of-life score of 0.<sup>19</sup> This will minimise the problem of reduced numbers and may provide more realistic conclusions.

### Analysis based on assumption of missing data being missing at random

The analysis suggested by Zwinderman<sup>20</sup> for quality-of-life data with missing measurements is based on the assumption that the missing data are missing at random, that is, non-response depends only on past measurements and not on future ones. This assumption enables the mean change pattern to be estimated using just the observed data.

The mean at time point,  $t + 1$ , is estimated as:

$$m_{t+1} = m_t + d_{t,t+1} \quad (2)$$

where  $m_t$  is the estimated mean at time point  $t$  and  $d_{t,t+1}$  is the mean difference in measurements at time points  $t$  and  $t + 1$  for those patients who have measurements at both time points.

This method assumes that the change in quality of life over time is independent of the level of quality of life. Allen-Mersh and colleagues<sup>21</sup> assessed longitudinal quality-of-life data using a repeated measures analysis of covariance based on Zwinderman's approach.

Zwinderman<sup>20</sup> discusses various computer programs in relation to this type of analysis. If the data are in the format of one record per patient containing all repeated measurements then BMDP5V<sup>22</sup> automatically uses all available data. The GLM procedure in SAS/STAT,<sup>23</sup> however, cannot deal with the missing data and performs a complete case analysis only, unless the data are transformed so that they consist of one record per patient per repeated measure. Brooks and colleagues<sup>3</sup> used BMDP5V to carry out a repeated measures analysis of variance on quality-of-life data and thus, intentionally or unintentionally, assumed missing data were missing at random.

## More complex modelling techniques for longitudinal data

Models for longitudinal quality-of-life data examine the patterns of change over time and the factors, such as treatment, that may have some influence on these patterns. The modelling needs to account for the correlation among values within each subject which is present in longitudinal data. There is much literature on modelling longitudinal data<sup>18,24,25</sup> and such

methods are considered here in terms of their application to quality-of-life data.

Longitudinal quality-of-life studies usually give rise to unbalanced data. This may be a result of assessment times varying from subject to subject or because of missing data. The impact of different types of missing data mechanisms on longitudinal data analysis has been discussed.<sup>26</sup> Analysis of longitudinal quality-of-life data requires modelling techniques that both capture the dynamic nature of the data and cope with the unbalanced structure. Techniques also need to be able to handle all types of data, since quality-of-life measures could be binary, ordinal or continuous.

Some of the more complex modelling approaches that satisfy these requirements fall into two main categories: **marginal models** (sometimes called population-averaged models) and **random effects models** (sometimes called subject-specific models). Both approaches are based on the theory of generalised linear models<sup>27</sup> and allow the analysis of binary, ordinal or continuous response variables. As such, the expected response is related to the explanatory variables in the model via a **link function**. The choice of link function depends on type of response variable; logit or probit link for binary responses, cumulative logit link for ordinal responses, log link for counts and identity link for continuous data.

In a marginal model, the regression coefficients estimate 'population-averaged' effects, that is, they give the estimated effect of explanatory variables on the response of the population as a whole. Random effects models allow each individual to have their own regression coefficients; these subject-specific effects are assumed to have been drawn from an overall population distribution, which is often the primary focus of the analysis.

Another difference between the two classes of model is the way in which they incorporate within-subject correlation. In the marginal modelling approach, the effect of explanatory variables on the marginal distributions of response and the within-subject correlation are modelled separately, with the latter treated as a nuisance factor. In the random effects approach, within-subject correlation is accommodated by incorporating a subject-specific random effect into the model.

For normally distributed data, parameter estimates can be obtained by using maximum likelihood methods. Laird and Ware<sup>28</sup> computed maximum

likelihood estimates in random effects models using the EM algorithm. An alternative for parameter estimation is to use a Bayesian approach using Gibbs sampling.<sup>28</sup>

For categorical responses, full likelihood estimation of regression parameters is generally not feasible. One alternative is to use a weighted least-squares approach.<sup>29</sup> There are two main limitations to this approach: large numbers of subjects are required and continuous explanatory variables cannot be accommodated. Another alternative is the generalised estimating equations method, developed for marginal models and based on the concept of quasi-likelihood.<sup>30,31</sup> Generalised estimating equations offer one way of estimating regression parameters together with their variances, while taking within-subject correlation into account, and can be used for both marginal and random effects models.<sup>32</sup> Bayesian analysis can also be used to estimate parameters.<sup>33</sup>

Hierarchical or multilevel models<sup>34</sup> are becoming a standard method in the analysis of longitudinal data where the lowest level of the hierarchy consists of the observations over time within each subject and second level units are the subjects. They have been applied to quality-of-life data.<sup>35,36</sup>

Modelling techniques for longitudinal data have been discussed in general,<sup>37</sup> in relation to normally distributed responses,<sup>38</sup> in relation to non-normally distributed responses,<sup>39</sup> and in relation to categorical responses.<sup>40-47</sup> Software for implementing the sort of modelling techniques discussed here are available, such as PROC MIXED in SAS/STAT<sup>48</sup> and MLN.<sup>49</sup> Examples of their application to quality-of-life data is limited,<sup>35,36,50</sup> possibly because of their complexity. The main problem with applying these models to quality-of-life data is that they assume the missing data mechanism is ignorable. In particular, the generalised estimating equations approach requires missing data to be missing completely at random. When there is informative drop-out, standard modelling techniques will give biased parameter estimates and modelling techniques that allow for informative drop-out should be considered (see chapter 8).

## References

1. O'Brien BJ, Buxton MJ, Ferguson BA. Measuring the effectiveness of heart transplant programs: quality of life data and their relationship to survival analysis. *J Chronic Dis* 1987;**40** (suppl 1):S137-53.

2. Hurny C, Bernhard J, Coates AS, Castiglione-Gertsch M, Peterson HF, Gelber RD, *et al.* Impact of adjuvant therapy on quality of life in women with node-positive operable breast cancer. *Lancet* 1996;**347**:1279–84.
3. Brooks MM, Gorkin L, Schron EB, Wiklund I, Campion J, Ledingham RB. Moricizine and quality of life in the Cardiac Arrhythmia Suppression Trial II (CAST II). *Control Clin Trials* 1994;**15**:437–49.
4. O'Brien BJ, Banner NR, Gibson S, Yacoub MH. The Nottingham Health Profile as a measure of quality of life following combined heart and lung transplantation. *J Epidemiol Community Health* 1988;**42**:232–4.
5. Matthews JNS, Altman DG, Campbell MJ, Royston P. Analysis of serial measurements in medical research. *BMJ* 1990;**300**:230–5.
6. Matthews JNS. A refinement to the analysis of serial data using summary measures. *Stat Med* 1993;**12**:27–37.
7. Hollen PJ, Gralla RJ, Cox C, Eberly SW, Kris MG. A dilemma in analysis: issues in the serial measurement of quality of life in patients with advanced lung cancer. *Lung Cancer* 1997;**18**:119–36.
8. Lydick E, Epstein RS, Himmelberger D, White CJ. Area under the curve: a metric for patient subjective responses in episodic diseases. *Qual Life Res* 1995;**4**:41–5.
9. Korn EL. On estimating the distribution function for quality of life in cancer clinical trials. *Biometrika* 1993;**80**:535–42.
10. Stephens RJ, Hopwood P. Defining and analysing the palliation of physical symptoms: a deceptively difficult exercise. *Palliat Med* 1995;**9**:63–75.
11. MRC Lung Cancer Working Party. Inoperable non-small cell lung cancer (NSCLC): a Medical Research Council randomised trial of palliative radiotherapy with two fractions or ten fractions. *Br J Cancer* 1991;**63**:265–70.
12. MRC Lung Cancer Working Party. A Medical Research Council (MRC) randomised trial of palliative radiotherapy with two fractions or a single fraction in patients with inoperable non-small cell lung cancer (NSCLC) and poor performance status. *Br J Cancer* 1992;**65**:934–41.
13. Nabholz JM, Gelmon K, Bontenbal M, Spielmann M, Catimel G, Conte P, *et al.* Multicenter, randomized comparative study of 2 doses of paclitaxel in patients with metastatic breast cancer. *J Clin Oncol* 1996;**14**:1858–67.
14. Rosenman J, Choi NC. Improved quality of life of patients with small-cell carcinoma of the lung by elective irradiation of the brain. *Int J Radiat Oncol Biol Phys* 1982;**8**:1041–3.
15. Hopwood P, Stephens RJ, Machin D. Approaches to the analysis of quality of life data: experiences gained from a Medical Research Council Lung Cancer Working Party palliative chemotherapy trial. *Qual Life Res* 1994;**3**:339–52.
16. MRC Lung Cancer Working Party. Randomized trial of palliative two-fraction versus more intensive 13-fraction radiotherapy for patients with inoperable non-small cell lung cancer and good performance status. *Clin Oncol* 1996;**8**:167–75.
17. Winer BJ. Statistical principles in experimental design. New York: McGraw-Hill; 1971.
18. Diggle PJ, Liang K, Zeger SL. Analysis of longitudinal data. Oxford: Clarendon Press; 1994.
19. Anderson JP, Kaplan RM, Schneiderman LJ. Effects of offering advance directives on quality adjusted life expectancy and psychological well-being among ill adults. *J Clin Epidemiol* 1994;**47**:761–72.
20. Zwiderman AH. Statistical analysis of longitudinal quality of life data with missing measurements. *Qual Life Res* 1992;**1**:219–24.
21. Allen-Mersh TG, Earlam S, Fordy C, Abrams K, Houghton J. Quality of life and survival with continuous hepatic-artery floxuridine infusion for colorectal liver metastases. *Lancet* 1994;**344**:1255–60.
22. BMDP Statistical Software. BMDP statistical software manual, vol. 2. California, USA: University of California Press; 1988.
23. SAS Institute. SAS/STAT users guide, v. 6, vol. 2. 4th ed. Cary, NC: SAS Institute; 1989.
24. Hand D, Crowder M. Practical longitudinal data analysis. London: Chapman and Hall; 1996.
25. Lindsey JK. Models for repeated measurements. New York: Oxford University Press; 1994.
26. Laird NM. Missing data in longitudinal studies. *Stat Med* 1988;**7**:305–15.
27. Dobson AJ. An introduction to generalized linear models. London: Chapman and Hall; 1990.
28. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982;**38**:963–74.
29. Koch GG, Landis JR, Freeman JL, Freeman DH, Lehnert RG. A general methodology for the analysis of experiments with repeated measurement of categorical data. *Biometrics* 1977;**33**:133–58.
30. Liang K, Zeger SL. Longitudinal data analysis using generalised linear models. *Biometrika* 1986;**73**:13–22.
31. Zeger SL, Liang K. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics* 1986;**42**:121–30.
32. Zeger SL, Liang K, Albert PS. Models for longitudinal data: a generalised estimating equation approach. *Biometrics* 1988;**44**:1049–60.

33. Best NG, Spiegelhalter DJ, Thomas A, Brayne CEG. Bayesian analysis of realistically complex models. *J R Stat Soc A* 1996;**159**:323–42.
34. Goldstein H. Multilevel statistical models. 2nd ed. London: Edward Arnold; 1995.
35. Beacon HJ, Thompson S. Multi-level models for repeated measurement data: application to quality of life data in clinical trials. *Stat Med* 1996;**15**:2717–32.
36. Beacon HJ. The statistical analysis of self assessed quality of life data in cancer clinical trials [PhD thesis]. London School of Hygiene and Tropical Medicine, University of London; 1996.
37. Zeger SL, Liang K. An overview of methods for the analysis of longitudinal data. *Stat Med* 1992;**11**:1825–39.
38. Louis TA. General methods for analysing repeated measures. *Stat Med* 1988;**7**:29–45.
39. Davis CS. Semi-parametric and non-parametric methods for the analysis of repeated measurements with applications to clinical trials. *Stat Med* 1991;**10**:1959–80.
40. Agresti A. Analysing repeated categorical response data. In: Agresti A. Categorical data analysis. New York: Wiley; 1990.
41. Agresti A. A survey of models for repeated ordered categorical response data. *Stat Med* 1989;**8**:1209–24.
42. Landis JR, Miller ME, Davis CS, Koch GG. Some general methods for the analysis of categorical data in longitudinal studies. *Stat Med* 1988;**7**:109–37.
43. Ware JH, Lipsitz S, Speizer FE. Issues in the analysis of repeated categorical outcomes. *Stat Med* 1988;**7**:95–107.
44. Agresti A. Distribution-free fitting of logit models with random effects for repeated categorical responses. *Stat Med* 1993;**12**:1969–87.
45. Francom SF, Chuang-Stein C, Landis JR. A log-linear model for ordinal data to characterize differential change among treatments. *Stat Med* 1989;**8**:571–82.
46. Wei LJ, Stram DO. Analysing repeated measurements with possibly missing observations by modelling marginal distributions. *Stat Med* 1988;**7**:139–48.
47. Ashby M, Neuhaus JM, Hauck WW, Bacchetti P, Heilbron DC, Jewell NP, *et al.* An annotated bibliography of methods for analysing correlated categorical data. *Stat Med* 1992;**11**:67–99.
48. SAS Institute. SAS/STAT software: changes and enhancements; release 6.07. SAS technical report P-229. Cary, NC: SAS Institute; 1992.
49. Rasbash J, Woodhouse G. MLn command reference. London: Institute of Education; 1995.
50. Zwinderman AH. The measurement of change of quality of life in clinical trials. *Stat Med* 1990;**9**:931–42.

## Chapter 8

# Longitudinal modelling methods that deal with informative drop-out

### Introduction

In longitudinal studies of quality of life with informative drop-out, the standard methods of longitudinal data analysis discussed in chapter 7 are frequently invalid, since they assume that the missing data mechanism is ignorable. Using standard methods in situations of informative drop-out may give biased results. A variety of methods for analysing longitudinal data in the presence of informative drop-out have been proposed but there is little evidence of their application to quality-of-life data. This chapter provides a brief overview of such methods.

### Methods

A variety of methods have been proposed for modelling the change in a continuous variable over time while accounting for informative drop-out, and Little<sup>1</sup> provides a review of such methods. Methods are generally based on random effects models, although an approach based on marginal models has also been considered.<sup>2</sup> The techniques focus on a classical approach but a Bayesian approach to longitudinal data models with informative drop-out has been discussed.<sup>3</sup>

Most methods are developed from a linear random effects model,<sup>4-8</sup> in which it is assumed that each subject's measurements follow a linear regression with random intercept and slope. The average rate of change can be estimated using likelihood-based approaches or using either weighted or unweighted averages of the individual least-squares slope estimates. In the presence of informative drop-out, likelihood-based approaches and weighted averages have been shown to be seriously biased and, although unweighted averages were generally found to be unbiased, they are too inefficient to merit consideration.<sup>4,6,9</sup> Thus, informative drop-out needs to be accounted for when modelling longitudinal quality-of-life data with such data missing.

Assuming a linear random effects model for the response variable over time, Wu and Carroll<sup>4</sup>

model the relationship between the probability of drop-out and the path of the response variable, as described by the intercept and slope parameters, using a probit model. Parameters were estimated using pseudo-maximum likelihood. They suggest that Cox proportional hazards and logistic models could also be appropriate.

Wu and Bailey<sup>5,6</sup> developed methods that account for informative drop-out without actually modelling the drop-out process. They proposed a conditional linear model for the individual least-squares estimated slopes, assuming them to be linear functions of total follow-up time. Two methods for estimating the population rate of change, based on a weighted least-squares approach were suggested and compared to pseudo-maximum likelihood estimates; linear minimum variance unbiased estimates and linear minimum mean squared error estimates. Mori and colleagues<sup>7</sup> developed the method further by using empirical Bayes methodology to estimate rate of change, enabling estimation of individual as well as population slopes.

The problem with the probit drop-out model approach and the conditional linear model approach is that at least two measurements per subject are needed to enable least-squares estimation of individual slopes. A further limitation of the conditional linear model approach is that the method assumes the potential follow-up time is the same for all subjects and thus application of such methods to studies with staggered patient entry times may be invalid.

Schluchter<sup>8</sup> developed the conditional linear model approach to allow for staggered patient entry and individuals with single measurements. The method assumes that the individual intercept, slope and log survival time follow a trivariate normal distribution and the maximum likelihood estimates of model parameters are calculated using the EM algorithm. Disadvantages of this approach are that it is computationally complex and, hence, not easily accessible and that large amounts of data may be required. A multilevel version of Schluchter's model has been applied

to quality-of-life data.<sup>10</sup> Related methods based on general rather than linear random effects models have also been proposed<sup>11,12</sup> but these suffer from the same limitations.

The method of multilevel modelling has been adapted to deal with informative drop-out by analysing sequentially overlapping portions of the follow-up information.<sup>13</sup> This is comparable to the descriptive method discussed earlier in which patients are split into subgroups according to differing lengths of follow-up (see page 25). A set of sequential time points,  $T_1, T_2, \dots, T_k, \dots, T_K$ , which are meaningful to the study are determined. All those patients who are still in the trial at time  $T_k$  form the risk set for that time point and the follow-up information for each of those patients consists just of the data up to time  $T_k$ . In this way, overlapping risk sets are created with patients with long overall follow-up contributing to most risk sets while patients who dropped out early contribute to the first few only. Multilevel models are fitted to each of the  $K$  risk sets. An overall multilevel model, including a third level representing the last risk set to which a patient belonged, can be fitted to evaluate the extent of informative drop-out in the data and provide a direct comparison of time profiles of patients with different follow-up times.

## Discussion

Longitudinal data models that deal with informative drop-out may be appropriate methods for analysing quality-of-life data but there is little evidence of their application in this field. The proposed methods have focused on continuous responses and thus may not be applicable to quality-of-life measures that are categorical or highly skewed continuous variables. De Stavola and Christensen<sup>13</sup> claim that their multilevel model approach can include categorical variables. Beacon<sup>10</sup> suggests that the conditional linear models proposed by Wu and Bailey<sup>5,6</sup> and, in particular, the trivariate normal model proposed by Schluchter<sup>8</sup> are the most suitable for application to quality-of-life data. Further work is needed to fully investigate the use of these methods in quality-of-life assessment, especially in relation to the application to categorical responses. Although the methods discussed here allow for informative drop-out, simultaneous modelling

methods, discussed later in chapter 19, in which both the quality-of-life and the drop-out process are explicitly modelled, may provide a potentially more powerful approach.

## References

1. Little RJA. Modelling the drop-out mechanism in repeated measures studies. *J Am Stat Assoc* 1995;**90**:1112–21.
2. Diggle P, Kenward MG. Informative drop-out in longitudinal data analysis. *Appl Stat* 1994;**43**:49–93.
3. Best NG, Spiegelhalter DJ, Thomas A, Brayne CEG. Bayesian analysis of realistically complex models. *J R Stat Soc A* 1996;**159**:323–42.
4. Wu MC, Carroll RJ. Estimation and comparison of changes in the presence of informative right censoring by modelling the censoring process. *Biometrics* 1988;**44**:175–88.
5. Wu MC, Bailey KR. Estimation and comparison of changes in the presence of informative right censoring: conditional linear model. *Biometrics* 1989;**45**:939–55.
6. Wu MC, Bailey K. Analysing changes in the presence of informative right censoring caused by death and withdrawal. *Stat Med* 1988;**7**:337–46.
7. Mori M, Woodworth GG, Woolson RF. Application of empirical Bayes inference to estimation of rate of change in the presence of informative right censoring. *Stat Med* 1992;**11**:621–31.
8. Schluchter MD. Methods for the analysis of informatively censored longitudinal data. *Stat Med* 1992;**11**:1861–70.
9. Wang-Clow F, Lange M, Laird NM, Ware JH. A simulation study of estimators for rates of change in longitudinal studies with attrition. *Stat Med* 1995;**14**:283–97.
10. Beacon HJ. The statistical analysis of self assessed quality of life data in cancer clinical trials [PhD thesis]. London School of Hygiene and Tropical Medicine, University of London; 1996.
11. de Gruttola V, Tu XM. Modelling progression of CD4-lymphocyte count and its relationship to survival time. *Biometrics* 1994;**50**:1003–14.
12. Pawitan Y, Self S. Modelling disease marker processes in AIDS. *J Am Stat Assoc* 1993;**88**:719–26.
13. de Stavola BL, Christensen E. Multilevel models for longitudinal variables prognostic for survival. *Lifetime Data Anal* 1996;**2**:329–47.



## Chapter 9

# Quality-of-life analysis in the presence of informative drop-out: summary and discussion

In quality-of-life studies where survival is also an issue, patients are often severely ill, and drop-out caused by illness or death is a common occurrence. This drop-out process may be informative but the situation may be complex, involving drop-outs of several types. Standard analytical methods for longitudinal data assume that the missing data mechanism is ignorable, and thus use of these methods for longitudinal quality-of-life data may give biased results and invalid conclusions. However, if the drop-out rates and reasons for drop-out are balanced across treatment arms then between-treatment comparisons may remain unbiased.<sup>1,2</sup>

The analysis of longitudinal quality-of-life data should begin descriptively to give insight to the data before any formal testing or modelling is carried out. Graphs of individual patient profiles allow data to be examined at the most basic level, while treatment group profiles enable a clearer comparison of treatments in relation to their effect on quality of life over time. Interpretation of group profiles needs caution when informative drop-out is present, as group summary measures may well be biased.

An initial approach to the formal analysis is to use summary measures, in which treatment comparison is based on a single value summarising the quality-of-life data over time. The method enables the use of simple standard statistical techniques and thus has the advantage of being easy to apply. It is also reasonably flexible in terms of the type of data to which it can be applied and gives results that are clinically meaningful. In replacing a set of repeated measures with a single summary measure, however, the method does not fully capture the dynamic nature of quality-of-life data. Furthermore, calculation of summary measures may be difficult and analysis may be biased if informative drop-out is present.

If quality-of-life data are collected at a reasonably small number of fixed time points and have an underlying normal distribution, then a simple repeated measures analysis of variance may be appropriate. Strategies that enable application of the method to quality-of-life data with missing measurements are available<sup>3</sup> but these methods assume the

missing data mechanism is ignorable, which will not of course be the case if drop-out is informative.

More complex modelling techniques, such as random coefficient and marginal models, allow greater flexibility in terms of being able to cope with unbalanced data, varying time points and non-normality, but their application to quality-of-life data may be difficult and problematic because of the complexity of the models.<sup>4</sup> Examples of the application of more complex modelling techniques to quality-of-life data is limited.<sup>5-7</sup> Although these techniques model change over time, they do not explicitly model the missing data mechanism and thus are not valid when informative drop-out is present.

Methods which assume data are missing at random, rather than missing completely at random,<sup>3,7</sup> go some way to dealing with the missing data problem in longitudinal quality-of-life studies but will still give biased results in situations of informative drop-out. There are no easy ways to test if missing data are missing at random or are informative, and assumptions have to be based on subjective judgement. The assumptions should be assessed as part of a sensitivity analysis.

Modelling techniques that deal with informative drop-out have been developed but are complex and focus on continuous response data. The practical application of such methods may be difficult since appropriate software is not readily available. The application of such methods to quality-of-life data has, as yet, not been fully investigated and further work is needed to assess their use in this context. Simultaneous modelling methods (discussed later in chapter 19) which explicitly model both the quality-of-life and the drop-out process may provide an approach that is potentially more powerful.

## References

1. MRC Lung Cancer Working Party. Assessment of quality of life in small-cell lung cancer using a daily diary card developed by the Medical Research Council Lung Cancer Working Party. *Br J Cancer* 1991;**64**:299-306.

2. MRC Lung Cancer Working Party. A Medical Research Council (MRC) randomised trial of palliative radiotherapy with two fractions or a single fraction in patients with inoperable non-small cell lung cancer (NSCLC) and poor performance status. *Br J Cancer* 1992;**65**:934–41.
3. Zwinderman AH. Statistical analysis of longitudinal quality of life data with missing measurements. *Qual Life Res* 1992;**1**:219–24.
4. Hopwood P, Stephens RJ, Machin D. Approaches to the analysis of quality of life data: experiences gained from a Medical Research Council Lung Cancer Working Party palliative chemotherapy trial. *Qual Life Res* 1994;**3**:339–52.
5. Beacon HJ, Thompson S. Multi-level models for repeated measurement data: application to quality of life data in clinical trials. *Stat Med* 1996;**15**:2717–32.
6. Beacon HJ. The statistical analysis of self assessed quality of life data in cancer clinical trials [PhD thesis]. London School of Hygiene and Tropical Medicine, University of London; 1996.
7. Zwinderman AH. The measurement of change of quality of life in clinical trials. *Stat Med* 1990;**9**:931–42.

## Chapter 10

# Analysis of survival data adjusting for quality of life: introduction

In many studies, especially cancer clinical trials, treatments are assessed primarily in terms of survival, usually defined as time from entry to trial until death. Survival is the main endpoint in studies evaluating potentially curative treatments or treatments which are expected to prolong life significantly. There may also be an interest in quality of life as a secondary endpoint, especially in situations where treatments have side-effects.

In comparing treatments in terms of survival, it is often necessary to adjust for patient-related factors that could potentially affect the survival time of a patient. These covariates may include demographic variables such as age and sex, physiological variables such as white blood cell count, or disease-related variables such as type of tumour or stage of disease at entry to the study. One such patient-related factor to consider as a covariate may be quality of life, either in terms of a baseline measure or in terms of its changing

values over time. In chapters 11–13 methods of analysing quality-of-life and survival data are dealt with, in which the aim of the study is to compare treatments in terms of survival while adjusting for quality of life. The analysis also allows the effect of quality of life on survival (that is, as a predictor) to be explored.

In chapter 11, some background to the analysis of survival data is given and the standard methods of analysis used to compare treatments in terms of survival without any covariate adjustment are outlined. In chapter 12, methods of analysis are discussed which enable a comparison of treatments while adjusting for the effects of covariates. Inclusion of both fixed and time-dependent covariates are considered and the particular application of these methods to situations with quality of life as a covariate is also discussed. The findings of this section of the review are summarised and discussed in chapter 13.



# Chapter 11

## Comparison of treatments in terms of survival

### Introduction to survival analysis

Survival data differs from other types of continuous data in that the actual survival time may not necessarily be observed for all subjects in the study, thus giving rise to censored data (see chapter 3). Any methods used to analyse survival data must be able to deal with censored data.

There are many books and papers that discuss statistical methods for survival analysis. Parmar and Machin<sup>1</sup> provide non-technical coverage, while Collett<sup>2</sup> and especially Kalbfleisch and Prentice<sup>3</sup> provide the reader with more theoretical detail.

### Survivor and hazard functions

Survival data is generally described and modelled in terms of two related functions, the survivor function and the hazard function. The survivor function,  $S(t)$ , represents the probability that an individual survives from the time origin to some time beyond  $t$ , and is given by:

$$S(t) = p(T > t) = 1 - F(t) = 1 - \int_{u=0}^t f(u) du \quad (3)$$

where  $T$  is the random variable representing survival time, and the distribution of survival times is described by  $f(t)$ , the probability density function, and  $F(t)$ , the associated cumulative distribution function.

The hazard function,  $h(t)$ , is the probability that an individual dies at time  $t$ , given that they have survived up to that time. It represents the instantaneous death rate for an individual surviving to time  $t$ . It is linked to the probability density function and survivor function by:

$$h(t) = \frac{f(t)}{S(t)} = - \frac{d}{dt} [\log S(t)] \quad (4)$$

The survivor function and hazard function can be estimated from observed data. If the form of  $f(t)$  is not specified then non-parametric procedures can be used, otherwise parametric models can be fitted to the data.

### Some parametric forms for survival data

In some situations it may be appropriate to assume a distribution for  $f(t)$ , the probability density function of the survival time. The most common distributions used to model survival data are the exponential and Weibull distributions, the exponential being a special form of Weibull distribution. Only these two distributions are considered in this report but, in general, other distributions such as log-normal, log-logistic and gamma may be more appropriate.<sup>2</sup>

If survival times have an exponential distribution then

$$f(t) = \lambda \exp(-\lambda t) \quad (5)$$

the survivor function is given by

$$S(t) = \exp(-\lambda t) \quad (6)$$

and the hazard function is given by

$$h(t) = \lambda \quad (7)$$

Thus an exponential distribution assumes that the hazard rate is constant over time.

If the constancy of the hazard rate is not a valid assumption, then a Weibull distribution may be a more appropriate distribution for survival times. In this case,

$$f(t) = \lambda \gamma t^{\gamma-1} \exp(-\lambda t^\gamma) \quad (8)$$

the survivor function is given by

$$S(t) = \exp(-\lambda t^\gamma) \quad (9)$$

and the hazard function is given by

$$h(t) = \lambda \gamma t^{\gamma-1} \quad (10)$$

The parameters  $\gamma$  and  $\lambda$  determine the shape and scale of the hazard function and are thus called the shape and scale parameters, respectively. In the special case of  $\gamma = 1$ , the distribution of survival times is exponential and the hazard is constant. If  $\gamma > 1$ , the hazard increases with time and if  $0 < \gamma < 1$  then the hazard decreases monotonically.

The suitability of these parametric models can be assessed using a log-cumulative hazard plot. This is a graph of  $\log(-\log [S(t)])$  against  $\log t$  where  $S(t)$  is the survivor function estimated by the Kaplan–Meier method (see below). If a Weibull model is appropriate, the line will be approximately straight, with the slope giving an estimate of the shape parameter for the distribution. If an exponential model is appropriate, the slope of the line will be approximately 1.

The parameters for a model are estimated from the survival data using the method of maximum likelihood. In SAS/STAT, the LIFEREG procedure<sup>4</sup> fits parametric models to survival data.

### Comparison of treatments using Kaplan–Meier estimates

The survivor function can be estimated non-parametrically from observed data, both censored and uncensored, using the Kaplan–Meier method. This method is also called the product-limit method and is based on maximum likelihood estimation.<sup>5</sup> Suppose deaths occur at times  $t_1 < t_2 < \dots < t_j < \dots < t_n$ , then the Kaplan–Meier estimate of the survivor function is given by

$$\hat{S}(t) = \prod_{j=1}^k \left(1 - \frac{d_j}{n_j}\right) \text{ for all } t_k \leq t \quad (11)$$

where  $n_j$  is the number of individuals alive just before time  $t_j$  and  $d_j$  is the number of deaths at time  $t_j$ . Survival times censored at time  $t_j$  are assumed to occur immediately after the death time when computing values of  $n_j$ . CIs for the survivor function can be calculated using a variety of different methods.<sup>1,2</sup>

The calculation of Kaplan–Meier estimates is based on the assumption that the deaths of individuals in the sample occur independently of one another. This allows the probabilities of surviving from one interval to the next to be multiplied together to give the survivor function. It should also be noted that the Kaplan–Meier method gives the maximum likelihood estimate of the survivor function only if deaths and censoring are independent.<sup>5</sup> Thus, for unbiased Kaplan–Meier estimates, it is necessary for the censoring mechanism to be non-informative.

Kaplan–Meier estimates of the survivor function  $S(t)$  can be plotted against time  $t$  as a survival

curve. The survival curve is a stepped plot with the survivor function dropping instantaneously at each time of death and remaining level between successive death times. These provide a useful summary of the data and can be used to determine summary statistics such as median survival time. Survival curves for different treatment groups can be used to compare treatments descriptively in terms of survival. A more formal comparison of survival curves can be made using various non-parametric tests (see page 43). The SAS/STAT LIFETEST procedure<sup>4</sup> calculates Kaplan–Meier estimates and plots survival curves.

#### Survival analysis of the MIC data

Kaplan–Meier estimates of the survivor function for each treatment group in the MIC study were calculated for just those patients in the quality-of-life study and plotted as survival curves (Figure 13). Summary statistics for each treatment group were calculated from these survival curves (Table 10).

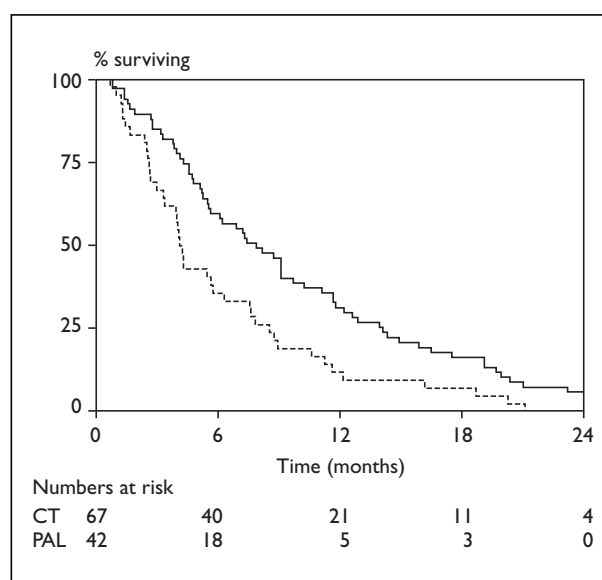


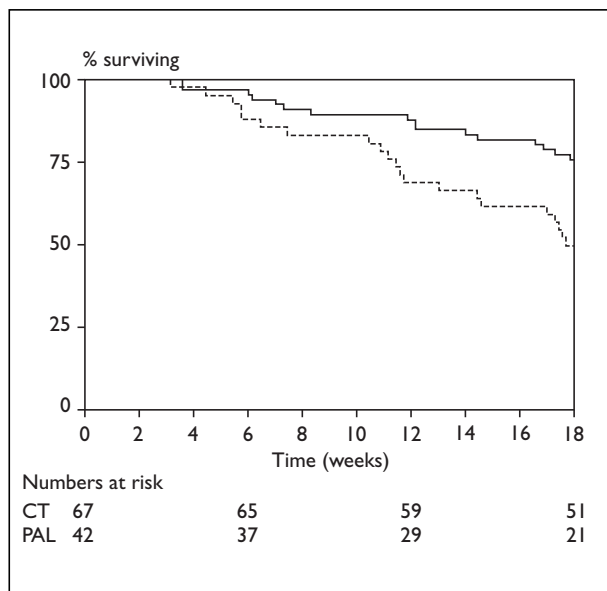
FIGURE 13 Overall survival for patients in the MIC quality-of-life study (—, CT; - - -, PAL)

TABLE 10 Summary statistics of survival for the MIC study

	CT arm (n = 67)	PAL arm (n = 42)
Number of deaths	65	42
Median (95% CI)	7.9 months (5.5–10.2)	4.2 months (3.3–6.3)
% surviving 1 year (95% CI)	31% (20–42%)	12% (2–22%)

For those individuals in the quality-of-life study, survival appeared to be better in the CT group than in the PAL group.

The 18-week period from entry to trial has been selected as the time on which the analysis will focus, since it was during this time that quality-of-life data was collected (see page 14). Survival within 18 weeks of entry to the trial can be examined using Kaplan–Meier estimates. The survival curves of each treatment group were plotted (Figure 14). These are just the overall survival curves (from Figure 13) magnified and cut-off at 18 weeks to highlight that period. The summary statistics for survival were taken from the curves (Table 11). Within the first 18 weeks of entry to trial, the CT group appeared to have a better survival than the PAL group.



**FIGURE 14** Survival within 18 weeks for patients in the MIC quality-of-life study (—, CT; - - -, PAL)

**TABLE 11** Summary statistics of survival within 18 weeks for the MIC study

	CT arm (n = 67)	PAL arm (n = 42)
Number of deaths	16	21
% surviving 18 weeks <sup>a</sup> (95% CI)	76% (66–86%)	50% (35–65%)
Note: medians were not reached in this analysis		
<sup>a</sup> Since there were no censored data before 18 weeks, the Kaplan–Meier estimates are the same as the crude proportions		

## Non-parametric tests for comparing survival

Survival in two or more groups of patients can be compared using a non-parametric test such as the log-rank test, also called the Mantel–Cox test.<sup>1,2</sup> This is the most widely used method of comparing survival curves.

The method essentially calculates at each death time, for each treatment group, the expected number of deaths under the null hypothesis of no difference between groups. These are then summed to give the total expected number of deaths in each treatment group, say  $E_i$  for treatment group  $i$ . The log-rank test compares the observed number of deaths in each treatment group, say  $O_i$  for treatment group  $i$ , to the expected number by calculating the test statistic

$$\chi^2 = \sum_{i=1}^g \frac{(O_i - E_i)^2}{E_i} \quad (12)$$

and comparing it to a chi-square distribution with  $(g - 1)$  degrees of freedom, where  $g$  is the number of treatment groups.

In the situation where two groups are being compared, the log-rank test is testing the null hypothesis that the ratio of the hazard rates in the two groups is equal to 1. The hazard ratio is a measure of the relative survival experience in the two groups and is estimated by

$$HR = \frac{O_1 / E_1}{O_2 / E_2} \quad (13)$$

where  $O_i / E_i$  is the estimated hazard rate in group  $i$ . A CI for the hazard ratio can be calculated.<sup>1</sup>

Other non-parametric tests which are sometimes used to compare groups in terms of survival are the Mantel–Haenszel test and the Wilcoxon test. These are more suitable than the log-rank test when the assumption of proportional hazards is not valid for the alternative hypothesis.<sup>1</sup> The assumption of proportional hazards can be assessed by means of a log-cumulative hazard plot (see page 46).

### Survival analysis of the MIC data

Hazard ratios were calculated and log-rank tests performed to compare the two treatment groups in the MIC study in terms of both survival and survival within 18 weeks (Table 12). There was strong evidence to suggest that chemotherapy

**TABLE 12** Hazard ratios and log-rank test for survival and survival within 18 weeks for the MIC study

	<b>CT vs PAL hazard ratio (95% CI)</b>	<b>Results from log-rank test</b>
Survival	0.56 (0.37, 0.86)	$\chi^2 = 9.06, p = 0.0026$
Survival within 18 weeks	0.40 (0.20, 0.79)	$\chi^2 = 8.07, p = 0.0045$

reduced the risk of death compared with standard palliative treatment both overall and within 18 weeks of entry to trial.

## References

1. Parmar MKB, Machin D. Survival analysis: a practical approach. England: John Wiley; 1995.
2. Collett D. Modelling survival data in medical research. London: Chapman and Hall; 1994.
3. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. New York: John Wiley; 1980.
4. SAS Institute Inc. SAS/STAT users guide, v. 6, vol. 2. 4th ed. Cary, NC: SAS Institute Inc; 1989.
5. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;**53**:457–81.



## Chapter 12

# Comparison of treatments in terms of survival adjusting for covariates

### Introduction

In comparing treatments in terms of survival, it is often necessary to adjust for patient-related factors, known as covariates, that could potentially affect the survival time of a patient. Covariates may be continuous measures, such as tumour size, or they may be ordinal variables, such as, for example, performance status rated as poor, fair and good. They may also be binary measures, such as the presence/absence of a symptom.

Covariates that keep the same value for the duration of a study are called fixed or time-independent, while those with possibly changing values over time are called time-dependent. Fixed covariates can be adjusted for in a survival analysis either using stratified survival analysis or using a Cox proportional hazards model. Time-dependent covariates can also be adjusted for in a Cox model.

### Stratified survival analysis

The simplest way to incorporate covariates into a survival analysis is to use a stratified survival analysis. This compares survival between treatment groups within each level of a covariate. The log-rank test methodology can be used to test for any differences between treatments, and comparisons of treatments within each stratum are combined to give an overall comparison of treatments that has been adjusted for the effect of the covariate.

Stratified survival analysis has two main limitations. Firstly, the method is only applicable to fixed

covariates and, secondly, the covariate has to be in a categorical format with, in practice, very few levels. Thus it may be preferable, especially when several variables are involved, to use a Cox model (see page 46).

For situations with quality of life as a covariate, a stratified survival analysis could be used if the quality-of-life measures are available at the start of the analysis and remain unchanged throughout, such as, for example, baseline measures. The measure of quality of life may need to be categorised so that it consists of a few levels containing a reasonable number of subjects.

### Stratified survival analysis of the MIC data

The overall survival and survival within 18 weeks were compared for both treatment groups in the MIC study using a stratified survival analysis. Stratification according to two different measures of baseline quality of life taken from the first questionnaire were considered. Using MAL, baseline quality of life was categorised as no malaise (MAL = 0) and malaise (MAL = 1, 2 or 3), and using MQS, baseline quality of life was categorised as above and below the median value of MQS (i.e. MQS  $\geq$  0.75 and MQS  $<$  0.75).

Although baseline malaise does not appear to affect survival (see *Table 13*), there is evidence of a relationship between the baseline values of MQS and overall survival (see *Table 14*). The hazard rates suggest that the risk of death is reduced if the baseline value of MQS is below the median, that is, if the baseline quality-of-life measure is 'good'. In a comparison of treatments,

**TABLE 13** Survival analysis of MIC data by baseline malaise

	<i>n</i>	Observed deaths	Expected deaths	Hazard rate (observed/expected)	$\chi^2$	<i>p</i> -value
<b>Survival</b>						
No malaise at baseline	42	41	41.48	0.99	0.009	0.92
Malaise at baseline	67	66	65.52	1.01		
<b>Survival within 18 weeks</b>						
No malaise at baseline	42	16	14.51	1.10	0.25	0.62
Malaise at baseline	67	21	22.49	0.93		

**TABLE 14** Survival analysis of MIC data by baseline MQS

	<i>n</i>	Observed deaths	Expected deaths	Hazard rate (observed/expected)	$\chi^2$	<i>p</i> -value
<b>Survival</b>						
MQS < 0.75	50	48	59.74	0.80	5.38	0.02
MQS ≥ 0.75	59	59	47.26	1.25		
<b>Survival within 18 weeks</b>						
MQS < 0.75	50	14	18.35	0.76	2.05	0.15
MQS ≥ 0.75	59	23	18.65	1.23		

stratification by these covariates does not greatly affect the results of the log-rank tests, with *p*-values from the stratified log-rank tests being similar to those from the unstratified test (see Table 15).

## Cox proportional hazards model with fixed covariates

### Background

Survival data can be modelled using the Cox proportional hazards model.<sup>1</sup> This is a regression model suitable for survival data, which allows the variation in survival, as expressed by the hazard function, to be explained by certain factors. By incorporating treatment with other covariates in the model, the differences in survival between treatment groups can be investigated while adjusting for the other covariates. Here the model is considered when covariates are fixed but the model can be extended to include time-dependent covariates (see page 47). The variables in the model can be continuous or categorical – in the latter case they take the form of dummy variables.

In the Cox proportional hazards model, the hazard function  $h(t)$  is modelled as:

$$h(t) = h_0(t) \exp(\beta^T x) \quad (13)$$

**TABLE 15** Stratified and unstratified survival analysis for treatment comparison in the MIC study (new results in bold)

	$\chi^2$	<i>p</i> -value
<b>Survival</b>		
Unstratified	9.06	0.0026
<b>Stratified by baseline malaise</b>	<b>8.70</b>	<b>0.0032</b>
<b>Stratified by baseline MQS</b>	<b>8.54</b>	<b>0.0035</b>
<b>Survival within 18 weeks</b>		
Unstratified	8.07	0.0045
<b>Stratified by baseline malaise</b>	<b>8.09</b>	<b>0.0044</b>
<b>Stratified by baseline MQS</b>	<b>8.25</b>	<b>0.0041</b>

where  $h_0(t)$  is the underlying baseline hazard function,  $x$  is a vector of fixed covariates and  $\beta$  is a vector of regression coefficients. Parameters are estimated by maximising a partial likelihood function.<sup>2</sup>

The model does not assume any particular form of probability distribution for the survival times and thus the underlying baseline hazard is allowed to be arbitrary. It does, however, have one underlying assumption, that of proportionality of hazards and, because of this restriction, the model is usually referred to as semi-parametric. If the proportional hazards assumption is valid, the ratio of the hazards in the subgroups defined by the covariate values remains approximately constant over time. The assumption of proportional hazards can be assessed by means of a log-cumulative hazard plot where, using Kaplan–Meier estimates of the survivor function,  $\log(-\log S(t))$  is plotted against  $\log t$  for each subgroup. If the assumption of proportional hazards is valid, the graphs will be parallel.<sup>3</sup>

Nested models can be compared by checking the change in the value of  $-2 \log$  (likelihood) against a chi-square distribution with degrees of freedom equal to the difference in the number of parameters being estimated.<sup>3</sup> Most standard statistical software packages, including the PHREG procedure in SAS/STAT software,<sup>4</sup> can be used to fit Cox regression models with fixed covariates.

### Application to quality-of-life data

The quality of life of a patient on entry to a study may be related to the survival of that patient and could be included in a survival model as a fixed covariate. Including baseline quality of life, together with treatment as covariates in the model, enables the difference in survival between treatment groups to be assessed, while accounting for the baseline effect of quality of life. Measures of quality of life may be continuous or categorical.

### Cox regression analysis of MIC data with quality of life as fixed covariate

A Cox model was fitted to the MIC survival data with two fixed covariates, treatment and baseline quality of life. Treatment was included as a binary variable with 1 representing chemotherapy and 0 representing standard palliative care. As with the stratified survival analysis (see page 45), two different measures of baseline quality of life taken from the first questionnaire were included. Using MAL as before, quality of life was included as a binary variable with 0 representing no malaise (MAL = 0) and 1 representing malaise (MAL = 1, 2 or 3). MQS was used as an alternative measure of quality of life and was included in the model as a continuous variable, rather than as a categorical variable in the stratified analysis.

The regression coefficients and hazard ratios for the quality-of-life covariates suggest that there is no relationship between baseline malaise and survival (see Table 16) but there is evidence of a relationship between baseline values of MQS and overall survival (see Table 17). Hazard rates suggest that increasing values of baseline MQS (i.e. worsening quality of life) are associated with an increased hazard of death. These results are comparable to the stratified analysis (see page 45). Regression coefficients and hazard ratios (see Tables 16 and 17) together with likelihood ratio tests (see Table 18) show that after adjusting for baseline values of malaise and MQS, there is still a significant treatment effect on both survival and survival within 18 weeks. The results show that chemotherapy reduces the hazard of death.

## Cox model with time-dependent covariates

### Background

A time-dependent covariate is a variable which may explain survival differences; its value for any individual may vary over time. Unlike fixed covariates, which remain constant over the study period, time-dependent covariates are assessed throughout a patient's follow-up period. The assessment times are not necessarily at regular intervals and often differ between patients. There are often problems in collecting complete data on time-dependent covariates and, in some circumstances, the data may be censored. Time-dependent covariates could be continuous, ordinal or binary. In particular, a time-dependent binary covariate may represent the occurrence of an event, that is, an indicator variable whose value remains at zero until the event occurs, at which time it becomes 1.

**TABLE 16** Cox regression analysis with baseline malaise as fixed covariate

	$\beta$ (95% CI)	Hazard ratio (95% CI)
<b>Survival</b>		
Malaise	0.04 (-0.35, 0.43)	1.04 (0.70, 1.54)
Treatment (adjusted for malaise)	-0.60 (-1.00, -0.20)	0.55 (0.37, 0.82)
<b>Survival within 18 weeks</b>		
Malaise	-0.15 (-0.80, 0.50)	0.86 (0.45, 1.65)
Treatment (adjusted for malaise)	-0.91 (-1.56, -0.26)	0.40 (0.21, 0.77)

**TABLE 17** Cox regression analysis with baseline MQS as fixed covariate

	$\beta$ (95% CI)	Hazard ratio (95% CI)
<b>Survival</b>		
MQS	0.79 (0.36, 1.22)	2.20 (1.43, 3.40)
Treatment (adjusted for MQS)	-0.55 (-0.95, -0.15)	0.58 (0.39, 0.86)
<b>Survival within 18 weeks</b>		
MQS	0.45 (-0.15, 1.05)	1.56 (0.86, 2.85)
Treatment (adjusted for MQS)	-0.85 (-1.51, -0.19)	0.43 (0.22, 0.83)

**TABLE 18** Cox regression analysis with fixed covariates compared with stratified and unstratified survival analysis for treatment comparison in the MIC study (new results in bold)

	$\chi^2$	p-value
<b>Survival</b>		
Unstratified	9.06	0.0026
Stratified by baseline malaise	8.70	0.0032
Stratified by baseline MQS	8.54	0.0035
<b>Adjusted for baseline malaise</b>	<b>8.40</b>	<b>0.0038</b>
<b>Adjusted for baseline MQS</b>	<b>6.86</b>	<b>0.0088</b>
<b>Survival within 18 weeks</b>		
Unstratified	8.07	0.0045
Stratified by baseline malaise	8.09	0.0044
Stratified by baseline MQS	8.25	0.0041
<b>Adjusted for baseline malaise</b>	<b>7.53</b>	<b>0.0061</b>
<b>Adjusted for baseline MQS</b>	<b>6.41</b>	<b>0.0113</b>

The Cox proportional hazards model has been discussed above in relation to fixed covariates but the model can be extended to incorporate time-dependent covariates. The Cox model for the hazard function can be written as

$$h(t) = h_0(t) \exp(\beta^T \mathbf{x} + \delta^T \mathbf{z}(t)) \quad (15)$$

where  $h_0(t)$  is the underlying baseline hazard,  $\mathbf{x}$  is a vector of fixed covariates with associated regression coefficients vector  $\beta$  and  $\mathbf{z}(t)$  is a vector of time-dependent covariates with associated regression coefficients vector  $\gamma$ . The variables in  $\mathbf{z}(t)$  represent successive measures of a binary, ordinal or continuous covariate.

When time-dependent variables are included in the model the relative hazard,  $h(t) / h_0(t)$  becomes time-dependent and so the model ceases to be a proportional hazards model.

#### Estimating model parameters

Suppose the death times of the  $n$  individuals in a study are  $t_1 < t_2 < \dots < t_i < \dots < t_n$ , where some of these death times may be censored, and suppose  $R(t_i)$  is the set of individuals at time  $t_i$  who are at risk of death. To fit the Cox model with a time-dependent covariate  $z(t)$ , the value of the covariate  $z(t_i)$  at each uncensored death time  $t_i$  is needed for all individuals in the risk set  $R(t_i)$ . The value of the covariate at any time  $t$  is usually taken to be the last recorded value prior to  $t$ . In situations where there are recorded values either side of the time of interest, then it may be preferable to use either the value at the closest time or, for continuous variables, a linearly interpolated value.

Nested models can be compared by checking the change in value of  $-2 \log$  (likelihood) against a chi-square distribution with degrees of freedom equal to the difference in the number of parameters being estimated. In SAS/STAT,<sup>4</sup> the PHREG procedure can be used to fit survival models with time-dependent covariates, although the `agreg` function in S-PLUS<sup>5</sup> may be more straightforward to use.

#### Application to quality-of-life data

During a study, a patient may experience changes in quality of life as time passes. This could be described either by a changing quality-of-life score or by the movement in and out of various quality-of-life health states. The change in score or pattern of movement between states may help to explain survival differences and should be considered for inclusion in any survival model as a covariate. The change in quality of life over time is clearly a time-dependent covariate.

Consideration of models that include a time-dependent covariate representing quality of life and a treatment term may provide useful information on the effectiveness of the treatment. If the model includes a quality-of-life term but the estimate of treatment effect is not clearly different from zero, this could indicate that the treatment affects quality of life, which in turn affects survival, but the treatment has no additional effect.

#### Cox regression analysis of MIC data with quality of life as time-dependent covariate

In a survival analysis of the MIC study, quality of life was included as a time-dependent covariate in two different ways. Patients were successively assessed for quality of life over time for 18 weeks from entry to study and the values of both MQS and MAL were included as time-dependent covariates in a Cox model of survival within 18 weeks. MQS was included in terms of the changing value over time while malaise was included in terms of the movement between two different health states: no malaise (MAL = 0) and malaise (MAL = 1, 2 or 3). Time spent with no malaise was allocated the value 0 and with malaise was allocated the value 1. Treatment was included in the model as a binary variable, with 1 representing chemotherapy and 0 representing standard palliative care.

The analysis required, for the whole survival time of a patient, a value of quality of life to be available continuously over the 18-week analysis period. To enable this, two assumptions regarding the quality-of-life data were made (see page 58 for further details of these assumptions).

- Changes in quality of life were assumed to occur midway between assessments and patients were assumed to remain in steady state between changes.
- Patients were assumed to remain in steady state from time of last assessment to either death or 18 weeks (whichever came first).

The S-PLUS<sup>5</sup> program was used to perform the analysis. The data were transformed so that it consisted of multiple lines per patient, with each line representing the time a patient spent either in different states of malaise or with different values of MQS. Each line of data consisted of the times of entry to and exit from the state relative to date of entry to study, an indicator variable representing whether the patient had died or not at the end of the interval, and a variable representing treatment.

The regression coefficients and hazard ratios for the quality-of-life covariates suggest that there is

**TABLE 19** Cox regression analysis with malaise as time-dependent covariate

	$\beta$ (95% CI)	Hazard ratio (95% CI)
<b>Survival within 18 weeks</b>		
Malaise	0.72 (-0.01, 1.45)	2.05 (0.99, 4.25)
Treatment (adjusted for malaise)	-0.94 (-1.59, -0.29)	0.39 (0.20, 0.75)

**TABLE 20** Cox regression analysis with MQS as time-dependent covariate

	$\beta$ (95% CI)	Hazard ratio (95% CI)
<b>Survival within 18 weeks</b>		
MQS	0.99 (0.44, 1.54)	2.68 (1.55, 4.66)
Treatment (adjusted for MQS)	-0.70 (-1.37, -0.03)	0.50 (0.25, 0.97)

no relationship between the changing values of malaise and survival within 18 weeks (see *Table 19*) but there is evidence of a relationship between the changing values of MQS and survival within 18 weeks (see *Table 20*). Hazard rates suggest that increasing values of MQS (i.e. worsening quality of life) are associated with an increased hazard of death. Regression coefficients and hazard ratios (see *Tables 19* and *20*), together with likelihood ratio tests (see *Table 21*), show that after adjusting

**TABLE 21** Cox regression analysis with time-dependent covariates compared with fixed covariate, stratified and unstratified survival analysis for treatment comparison in the MIC study (new results in bold)

	$\chi^2$	p-value
<b>Survival within 18 weeks</b>		
Unstratified	8.07	0.0045
Stratified by baseline malaise	8.09	0.0044
Stratified by baseline MQS	8.25	0.0041
Adjusted for baseline malaise	7.53	0.0061
Adjusted for baseline MQS	6.41	0.0113
<b>Adjusted for changing malaise</b>	<b>8.02</b>	<b>0.0046</b>
<b>Adjusted for changing MQS</b>	<b>4.30</b>	<b>0.0381</b>

for the effect of changing values of both malaise and MQS over time, survival within 18 weeks is still significantly different between treatments.

## References

1. Cox DR. Regression models and life tables. *J R Stat Soc B* 1972;**34**:187-220.
2. Cox DR. Partial likelihood. *Biometrika* 1975; **62**:269-76.
3. Collett D. Modelling survival data in medical research. London: Chapman and Hall; 1994.
4. SAS Institute Inc. SAS/STAT software: changes and enhancements; release 6.07. SAS technical report P-229. Cary, NC: SAS Institute Inc; 1992.
5. Statistical Science Inc. S-PLUS user's manual. Seattle: Statistical Sciences Inc; 1991.



## Chapter 13

# Analysis of survival data adjusting for quality of life: summary and discussion

Treatment comparison in terms of survival may, in some circumstances, need to adjust for the effects of quality of life. Standard survival analysis techniques can be used for this purpose. Quality of life may be included as a covariate, either in terms of baseline values or in terms of changing values over time. The quality-of-life measure may be binary, ordinal or continuous, and survival times may have a specified or unspecified underlying distribution.

If adjustment for baseline quality-of-life measures is necessary, then a stratified survival analysis can be used. Although the method is simple, it has limited use; the stratification variable must be categorical and have a small number of levels, and the method is unable to cope with more than a few covariates. Categorisation of a continuous quality-of-life measure for use in a stratified survival analysis may result in a loss of information; a modelling approach, which can incorporate any number and type of covariates, may be more appropriate.

Cox proportional hazards regression models can be used to adjust for baseline measures of quality of life in a survival analysis, by incorporating the values as a fixed covariate. Cox models can also be used to adjust for quality of life in terms of change over time by including the changing values as a time-dependent covariate. Modelling in this way also enables investigation of possible interactions between treatment and quality of life, which may be particularly important if a treatment, for example, causes early toxicity.

If assessments of quality of life are infrequent or data are missing for reasons other than death, then it may be difficult to adjust for changing quality of life with any degree of accuracy. The analysis may be improved by modelling quality of life and survival as two simultaneous processes, as discussed later in chapter 19. In this approach, values of quality of life incorporated into a survival analysis as a covariate are estimated from the model for quality-of-life data over time fitted to all patients.





## Chapter 14

# Simultaneous analysis of quality-of-life and survival data: introduction

Health technology assessment often requires evaluation in terms of a trade-off between quality and length of life. In comparing the effectiveness of a new treatment against a standard treatment in a randomised clinical trial, benefits in terms of survival often have to be weighed up against the extent of undesirable side-effects. Alternatively, a new treatment may have no extra benefit in terms of survival but it may result in improved quality of life through improved relief of symptoms or reduced toxicity.

When quality of life and survival are analysed as separate endpoints, it is often difficult to assess the balance between the two in selecting the optimal therapy for patients, especially when the two endpoints indicate conflicting treatment preferences. It may therefore be preferable to assess the effect of a treatment on quality of life and survival simultaneously, and the statistical methods appropriate for that purpose are discussed in chapters 15–20.

The concept of QALYs is introduced in chapter 15. This is a composite outcome measure in which the survival time of a patient is scaled down according to the quality of life they experience. The quality-adjusted survival time is then used to compare treatments in a quality-adjusted survival analysis,

discussed in chapter 16, which includes the TWiST (time spent without symptoms of disease and toxicity of treatment) concept and Q-TWiST (quality-adjusted TWiST) methodology.

An alternative approach to the simultaneous analysis of quality-of-life and survival data, discussed in chapter 17, is multistate survival analysis. Rather than analysing quality-adjusted survival times, this method models the movement of patients between a finite number of health states defined in terms of quality of life. Some relatively new methodology is outlined in chapter 18, which combines this approach with the previous quality-adjusted survival analysis approach by incorporating the results from a multistate model into a Q-TWiST analysis.

The direct modelling of quality of life and survival is considered as two simultaneous processes in chapter 19, in particular, how the two interrelate. Such modelling allows for the comparison of quality of life conditional upon survival. A more general approach, which considers modelling the drop-out process rather than survival, is also discussed.

The findings of this part of the study are summarised and discussed in chapter 20.



# Chapter 15

## Quality-adjusted life years

### Introduction

QALYs are a measure of health status and can be used as an alternative to survival time in some areas of health technology assessment. They account for both quality and duration of survival in a single outcome measure and express health status in terms of well years of life.

The QALY methodology (see page 57) consists of down-weighting periods of survival time for which patients experience sub-optimal levels of quality of life. A special type of weight, often used in QALY calculations, is the utility, which was mentioned briefly in chapter 3 as a means of measuring quality of life. An introduction to utilities follows below. The QALY methodology is illustrated with data from the MIC study.

QALYs were first proposed in the USA by Fanshel and Bush<sup>1</sup> and have been developed in the UK by Alan Williams and colleagues at York University.<sup>2</sup> They are extensively used in health economics, forming the basis for cost-utility analysis, and Torrance<sup>3</sup> was at the forefront of the methodology in this field. They are also widely used in the field of decision analysis. These QALY applications are discussed below. In this report, the main application of the QALY concept is in chapter 16, where the assessment of health technologies in terms of quality-adjusted survival analysis is discussed. The advantage of QALYs lies in their simplicity but they have been the subject of much criticism (see page 61) and a number of alternatives have been proposed.

### Utilities

Each health state experienced by an individual during the course of their disease and treatment is associated with a quality of life. Each can be assigned a value that measures the preference of the individual for a health state relative to other states. The value reflects the quality of life in that health state. The value generally lies between 0, representing death, and 1, representing perfect health, although it can be negative, representing health states judged to be worse than death. Technically, these values are utilities when they are

measured under conditions of uncertainty but they may approximate to utilities when measured under other conditions.

Once the health states experienced by patients in a study have been defined, then the utilities for these states need to be determined. Methods of measuring health state utilities have been discussed<sup>4-6</sup> and Torrance,<sup>7</sup> in particular, provides a comprehensive review. There are three broad methods available to researchers for determining utilities for a study; using subjective judgement, using values in the literature or measuring the values. The values can be measured by using a **holistic approach** in which the health state being valued is considered as a complete entity or by using a **decomposed approach** which values a health state indirectly by considering particular aspects of a health state.

### Determining utilities using subjective judgement

The researchers themselves can make subjective assessments as to the utilities of various health states in their study or they can call on clinicians or other experts for their opinion on the values. This is a quick and inexpensive way to determine utilities but will not necessarily provide realistic measures. Sensitivity analysis should be carried out to determine the robustness of the conclusions to the choice of utility values.

### Taking utilities from the literature

The literature can be searched for utility values from previously conducted studies that could be used in the present study. The subjects and health states for which utilities were calculated in a published study should be appropriate for the current need. One very commonly used study is that by Rosser and Kind,<sup>8</sup> whose valuations are at the forefront of the work on QALYs undertaken in the UK.

### Rosser and Kind valuation matrix

The matrix of utilities derived by Rosser and Kind<sup>8</sup> is based on a classification system derived by Rosser and Watts<sup>9</sup> and values paired combinations of eight categories of disability (no disability to unconscious) and four categories of distress (none, mild, moderate and severe). The matrix represents the average valuation by 70 people (20 patients,

20 nurses, 20 healthy volunteers and ten doctors), given on a scale of 0, representing death, to 1, representing perfect health, with two negative values representing states worse than death.

If patients in a study are categorised into the same health states as the Rosser and Kind valuation matrix,<sup>8</sup> then the utilities from the matrix can be used in QALY calculations. One approach is to use clinical judgement to place patients into the categories,<sup>10-13</sup> while another approach is to use the Health Measurement Questionnaire, as developed at the University of York.<sup>14</sup> The questionnaire has been specially devised so that responses given by individuals can be used to place them into one of the Rosser and Kind categories.<sup>15,16</sup> A further approach may be to reprocess quality-of-life data collected by other means to categorise patients.<sup>17-21</sup>

There has been some discussion regarding the appropriateness of using the Rosser and Kind valuation matrix.<sup>21-24</sup> Criticisms relate to the study being out-of-date, the sample on which valuations were measured being small and unrepresentative of the general population, and the scale being insensitive. There has also been criticism regarding the use of the Health Measurement Questionnaire before proper testing and validation of the instrument has been carried out.<sup>25</sup> The use of the Rosser and Kind valuation matrix has been largely superseded by others such as the EuroQol.<sup>26</sup>

### Measuring utilities: the holistic approach

The most accurate way to measure utility values is by questioning a sample of individuals and eliciting their relative preferences for health states. The holistic approach considers a health state as a complete entity. The subjects could be the patients themselves, health professionals or the general public. The choice of subjects may not be important since large differences in utilities between different groups of people are unusual,<sup>7</sup> although some differences have been observed.<sup>27-29</sup>

In a clinical trial scenario, patients are the most appropriate subjects when measuring the utility of their condition. If, for example, health professionals or the general public are asked to measure the utility of a health state then a clear description needs to be given, since, unlike a patient, they are not actually in that health state and have not experienced it. The description should be abbreviated in terms of the physical, emotional and social functioning of the condition and should include age at onset, duration of the state, exact prognosis for what follows the state and whether it applies to them or someone else.<sup>7</sup> The general

public may need a different level of detail to that given to health professionals.

Having selected the individuals on whom utilities will be measured, an appropriate approach to measuring utilities needs to be chosen. There are three main methods that will be discussed here, although other methods such as ratio scaling<sup>7</sup> may also be used. The **rating scale** is a direct method for eliciting subject preferences while the **standard gamble** and **time trade-off** (TTO) are indirect methods.

#### Rating scale

The rating scale method involves presenting subjects with a line on a page, with one end of the line clearly defined as the most preferred health state and the other end as the least preferred. The subjects are asked to place health states on the line, such that the order and spacing represent preferences and differences in preferences as perceived by the subject.

#### Standard gamble

This is the classical method of measuring cardinal preferences and is the technique used in economic theory to measure utilities. It is based directly on the utility theory presented by von Neumann and Morgenstern.<sup>30</sup>

The subject is presented with two scenarios that they must choose between:

- (i) treatment resulting in either normal healthy life for  $t$  years with probability  $p$  or immediate death with a probability of  $1 - p$
- (ii) a certain health state for life ( $t$  years).

The probability  $p$  is varied until the respondent is indifferent to the two choices, thus giving the utility value,  $p$ , for that state. There is evidence to suggest that the standard gamble method overestimates utilities for health states and an improved version has been proposed.<sup>31</sup>

#### TTO

This method was developed specially for use in health care by Torrance and colleagues.<sup>32</sup> The subject is asked how much time,  $x$ , in a state of perfect health he/she considers equivalent to a period,  $t$ , in his/her current health state (worse than perfect health). The choice of the length of period  $t$  is problematic. In practice, the subject chooses between:

- (i) a certain health state for time  $t$  followed by death
- (ii) healthy for time  $x < t$  followed by death.

Time  $x$  is varied until the respondent is indifferent to the choices and  $x/t$ , the TTO-score, ranging from 0 to 1, provides an estimate of the utility for that state.

When the TTO-score assessed for time  $t$  is applied to other periods then **constant proportional trade-off** is assumed. For example, if an individual considers 16 years in perfect health equivalent to 20 years in a reduced health state, then it is assumed they would also consider 4 years in perfect health to be equivalent to 5 years in the reduced health state. Studies have shown that this assumption may not always be valid.<sup>27,33,34</sup> A violation of the proportionality assumption invalidates the calculation of a TTO-score and subsequent application of this score to periods of length other than  $t$  in QALY calculations (see page 62). For further discussion regarding the interpretation of the TTO score as a measure of utility, see the paper by Stalmeier and colleagues.<sup>35</sup>

### Comparison of methods

The standard gamble method is the gold standard for measuring utilities but is impractical, especially when a large number of health states need to be measured, and can be expensive because of the need for interviewers. The TTO is similar to the standard gamble but is easier to use, and the rating scale method is the simplest of all. The TTO and rating scale methods, however, do not in general measure utility directly since they do not involve probabilities, and utilities are preferences measured under conditions of uncertainty. Preference values,  $v$ , from methods such as TTO and rating scale, can be converted to utilities,  $u$ , using a power function:<sup>36,37</sup>

$$(1 - u) = (1 - v)^\alpha \quad (16)$$

where  $\alpha$  reflects the risk attitude of the individuals ( $0 < \alpha < 1$  reflects risk-averse,  $\alpha > 1$  reflects risk-seeking, and  $\alpha = 1$  reflects risk-neutral) and may be estimated by measuring values and utilities on the same health states. The TTO method only measures the true utility when people are risk-neutral ( $\alpha = 1$ ) and the utility function is linear in time. Studies have shown that the TTO method tends to underestimate utilities since individuals are generally risk-averse rather than risk-neutral.<sup>38</sup>

The rating scale, standard gamble and TTO methods have been compared with each other and with other methods.<sup>29,39-43</sup> The different methods can give very different results and this can then affect any analysis that incorporates them.<sup>29,40</sup> Torrance<sup>41</sup> recommended researchers to use

the TTO technique if they could afford it and, otherwise, to use the rating scale with power curve correction (equation 16).

### Measuring utilities: the decomposed approach

The decomposed approach to measuring utilities is based on multi-attribute utility theory,<sup>37</sup> in which health status is described in terms of a set of core attributes, each broken down into a series of levels describing the range of functioning (as shown by Barr and colleagues,<sup>44</sup> for example). Each different combination of levels, one from each attribute, represents a unique health state.

Utility values for each subject in the study are estimated by measuring the health status of a patient using a questionnaire, and converting their responses into a utility by combining them in a predetermined formulation. A variety of questionnaires have been devised to measure preferences using a decomposed approach, including the Health Utility Index, the Quality of Well-being Scale, the Rosser Index and the EuroQoL<sup>26</sup> (now called the EQ-5D). A comparison of these instruments is reported in another issue of *Health Technology Assessment*.<sup>45</sup>

## QALY methodology

### The QALY model

To calculate QALYs, years of life are multiplied by a fraction, the quality-adjustment fraction, which expresses the impairment in quality of life experienced during this time. The quality-adjustment fraction, ranging from 0, representing quality of life equivalent to death, to 1, representing perfect health, may represent the utility and can be derived in several ways (see above).

If the patient experiences or is expected to experience a series of health states  $s_i$  ( $i = 1$  to  $n$ ), with different levels of quality of life measured by utilities  $u_i$  ( $i = 1$  to  $n$ ), and the time spent in each state  $s_i$  is given by  $t_i$  ( $i = 1$  to  $n$ ), then the conventional approach to calculating QALYs is to sum the weighted times spent in the different states. This gives the following standard form for the QALY model:

$$\text{QALY} = \sum_{i=1}^n u_i t_i \quad (17)$$

In graphical terms, if the quality of life of an individual, represented by utility values, is plotted over time, then the QALYs for that individual are

calculated as the area under this curve (see below and also *Figures 16 and 17* later for examples from the MIC data).

### Discounting

QALYs are sometimes discounted to adjust for the fact that immediate benefits are generally valued more highly than later ones. The benefits of future life years are converted into present values by discounting. A discount rate of 5% per annum is used in most studies. If discounting is at a rate of  $r\%$  per year, then the discount factor for year  $n$  is given by:

$$\text{Discount factor}_n = (1 + r/100)^{-n} \quad (18)$$

For example, suppose an individual lives for 5 years with a quality-adjustment fraction of 0.8, then the total number of non-discounted QALYs will be 4 (i.e.  $5 \times 0.8$ ). If discounting at 5% per year then the total number of discounted QALYs will be 3.46 (i.e.  $(1.05^{-1} \times 0.8) + (1.05^{-2} \times 0.8) + \dots + (1.05^{-5} \times 0.8)$ ). Gudex and Kind<sup>14</sup> expand further on discounting in their explanation of QALY methodology using Rosser and Kind valuations.

### Model assumptions

Glaziou and colleagues<sup>46</sup> discuss the assumptions on which the QALY model is based, which are as follows:

- **utility independence:** the utility value for a health state does not depend on the time spent in that state
- **context independence:** the utility value assigned to a health state is independent of previous or future quality of life or the amount of remaining life
- **risk neutrality:** all life years are valued equivalently.

The risk neutrality assumption means that time is included in the model as a linear term,  $t$ , rather than as a non-linear function of time, that is,  $f(t)$ . More general models, that include some sort of discounting or risk-adjustment, have also been suggested.<sup>46–48</sup>

### Practical issues

Given longitudinal quality-of-life data and survival times for a group of subjects, there are two approaches to calculating QALYs. The first is to use the serial measurements as utility values over time and thus calculate QALYs for each individual. This is straightforward if the instrument for measuring quality of life yields a utility-type value, otherwise the quality-of-life scores need to

be transformed so that they are on a 0–1 scale, with 0 representing death and 1 representing perfect health. The second approach is to use the quality-of-life data to categorise a patient at any time into one of a number of health states. The utility values for these health states are determined separately and can be combined with the time spent in each health state to form QALYs for each patient.

If the quality of life for an individual changes between two consecutive assessments then, for both approaches to QALY calculation, a technique for determining the exact time of the change is required. Three possible techniques are as follows (see *Figure 15*):

- assume that quality of life changes linearly between assessments
- assume that quality of life is maintained from one assessment to the next
- assume that quality of life changes at the mid-point between assessments.

The appropriate choice may depend on the type of quality-of-life measurement, the timing of the assessments and the period to which the questions relate. Ganiats and colleagues<sup>49</sup> compared the first two techniques and concluded that either the appropriate choice for a study should be decided at the protocol stage or a sensitivity analysis should be carried out to establish the extent to which the conclusions are affected by the technique chosen.

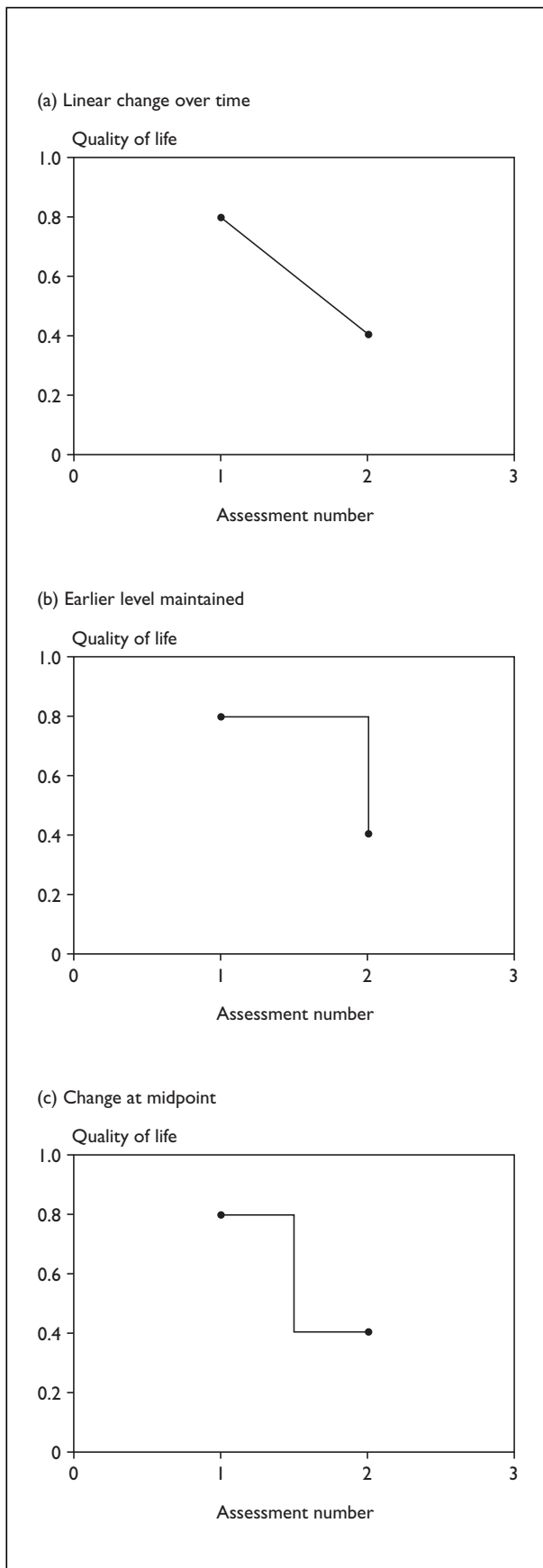
### Example: calculating QALYs in the MIC study

In the MIC study, quality-of-life data is available only for the first 18 weeks from study entry (see chapter 4) and so only the survival time for this limited period can be adjusted for quality of life. Assessments were made at several distinct time points during this period and various assumptions are needed in order to translate the quality-of-life information so that it is available over continuous time (see below). Sensitivity analysis could be used to assess the impact of the assumptions on the conclusions of any analysis. The two approaches to calculating QALYs, based on these assumptions, are described below.

#### Assumptions regarding quality of life over time

##### **Time between study entry and first assessment**

Quality of life during the time between study entry and first assessment was assumed to be that measured at the first assessment, that is, the first



**FIGURE 15** Handling the change in quality of life between assessments

value was carried backwards to date of study entry. This is not an unreasonable assumption, firstly because the questionnaire related to the previous 3-week period and, secondly, because the time between study entry and first assessment was reasonably short (median, 1 week; range, 0–6.7 weeks).

#### **Time between last assessment and either death or 18 weeks**

Quality of life during the time between last assessment and either death or 18 weeks (whichever came first) was assumed to be that measured at the last assessment, that is, the last value was carried forward. This is not an unreasonable assumption if the period is short but, for those who drop out of the study for reasons other than death, this assumption may be problematic (see *Table 8*). Other approaches, such as worst value carried forward and linear decrease over time were considered. Each has its own problems, especially when dealing with drop-outs (see *Table 8*).

#### **Times of changes in quality of life**

If quality of life changes between two consecutive assessments then the change is assumed to occur at the midpoint between the two assessment times. This is assumption (c) discussed above under 'Practical issues'. Quality of life is assumed to remain constant between changes. Intermittent missing quality-of-life values are handled by considering consecutive **non-missing** assessments when determining changes in quality of life. Thus, if there was a change in quality of life between assessment  $m$  taken at  $t_m$  and assessment  $n$  taken at  $t_n$ , where  $m < n$ , then the time of change was taken as:

$$\text{Time of change} = t_m + 0.5 (t_n - t_m) \quad (19)$$

Alternatives for estimating the exact times of change were considered. The actual assessment date when the change in quality of life was measured could have been used, that is, assumption (b) discussed above under 'Practical issues'. However, this option takes no account of the fact that the questionnaire was retrospective. To accommodate this, an alternative was considered in which time of change was chosen as the date 3 weeks before the date when the change in quality of life was measured. For various reasons, this was problematic and it was considered that, in reality, the patient would probably not recall their quality of life that far back. The midpoint option therefore seemed the most realistic and the effect of this choice on the analysis could be investigated in a sensitivity analysis.

## Calculating QALYs

Two different approaches, as discussed earlier, were considered for calculating QALYs in the MIC study. One approach was to use MQSs to estimate utility over time and calculate QALYs. The other approach was to use MAL to divide survival time into periods of differing quality of life and use selected utility values to calculate QALYs.

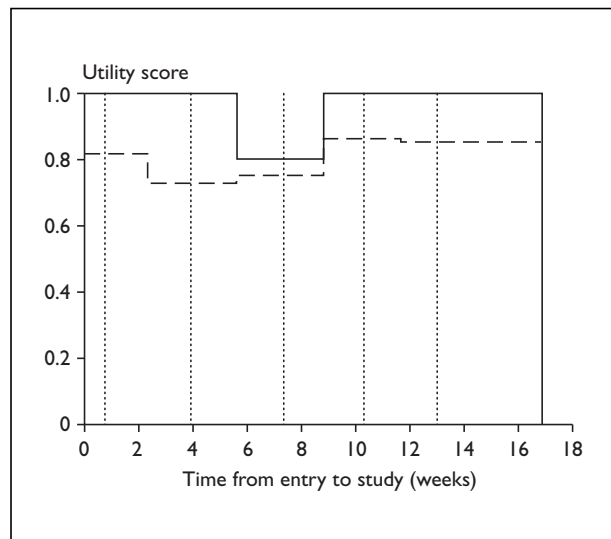
In order to use the MQS as a utility, transformation of the score was required. The maximum possible score of 3, occurring when all symptoms and side-effects were severe, was not deemed to be equivalent to death and so a score of 4 was allocated to death. The minimum possible score was 0, which occurred when there were no symptoms or side-effects, reflecting perfect health. The following transformation reverses the score and places it on a 0 to 1 scale:

$$MQS' = 1 - MQS / 4 \quad (20)$$

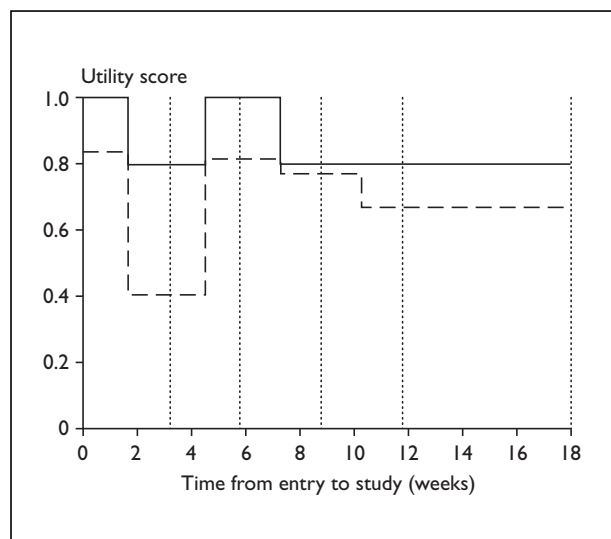
with  $MQS' = 0$  assumed to be the quality-of-life score for death. The problem with this approach in the MIC study is that quality of life was not measured as a utility. In particular, the responses were categorical and quality of life was not assessed relative to death. Thus the use of the transformed score as a utility is subjective and controversial, and should be viewed only as an illustrative example.

As an alternative, MAL was used to divide each patient's survival time within 18 weeks into periods of 'good' (MAL = 0) and 'poor' (MAL = 1, 2 or 3) quality of life. To illustrate this approach, utility values for these two health states were chosen arbitrarily as 1 and 0.8, respectively. The choice of utility values is subjective but alternative values can be considered as part of a sensitivity analysis. This approach to QALY calculation was chosen as the preferred option for the MIC study, since it was more flexible with regard to utilities.

As an illustration, the QALYs calculated from each approach were compared for two individuals in the MIC study, one who died within the 18-week period (see *Figure 16*) and one who died after 18 weeks (see *Figure 17*). The values of  $MQS'$  were plotted over time, together with the division of survival time based on MAL. The area under each of these curves gives the QALYs within 18 weeks for those individuals. The plots in *Figure 16* gave QALYs of 13.58 for  $MQS'$  and 16.21 for MAL, while those in *Figure 17* gave QALYs of 12.21 for  $MQS'$



**FIGURE 16** Quality-of-life profile for an individual in the MIC study who died within 18 weeks (—, MAL; ---, MQS'; vertical dotted lines represent QoL assessment times)



**FIGURE 17** Quality-of-life profile for an individual in the MIC study who died after 18 weeks (—, MAL; ---, MQS'; vertical dotted lines represent QoL assessment times)

and 15.27 for MAL. The QALYs based on MAL are used in chapter 16 to illustrate a quality-adjusted survival analysis.

## Applications of QALYs

QALYs are widely used in the fields of health economics and decision analysis and a brief outline of such applications follows. Examples are given and further examples can be found in the bibliography (see page 109). QALYs can also be used as the outcome measure in an assessment of treatments in a clinical trial. This forms the basis



for quality-adjusted survival analysis, discussed in chapter 16.

### The use of QALYs in health economics

Cost–utility analysis assesses health technologies in terms of cost per QALY. For example, a hip replacement costs £750 per QALY gained and a heart transplant costs £5000 per QALY gained.<sup>2</sup> Different strategies for treating a group of patients can be compared in terms of this outcome measure, to aid determination of the optimal treatment.<sup>50–56</sup> The use of QALYs in this context has been discussed.<sup>57–59</sup> Cost–utility analysis can also be used to assess the benefits of health interventions such as disease prevention programmes,<sup>60</sup> screening programmes<sup>61,62</sup> and healthcare guidelines.<sup>63</sup> Further examples can be found in the decision-analysis field (see below).

A major use of cost–utility analysis is in the prioritisation of strategies for treating different groups of patients, aiding decisions regarding resource allocation. Gudex<sup>21</sup> provides a detailed example of the use of QALYs for this purpose. The use of QALYs in resource allocation has been discussed,<sup>2,64–68</sup> although there is much controversy regarding such applications (see below).

### The use of QALYs in decision analysis

Decision analysis is a technique used to aid clinical decision making and is sometimes seen as an alternative to other approaches, such as clinical trials and meta-analysis, for health technology assessment. It provides a means of synthesising all existing knowledge on an intervention and quantifying the net benefits. Decision analysis is not only useful for evaluating optimal treatment for patients, but also for identifying the most important variables in decision-making processes.<sup>69</sup>

Life expectancy has been the most popular outcome measurement for decision analyses but QALYs are also now being widely used as an alternative.<sup>70</sup> Using QALYs as an outcome measure requires estimates of life expectancy, together with appropriate quality-of-life adjustments. Cost–utility analysis is a specialised form of decision analysis using cost per QALY as an outcome measure. There are many examples in the literature in which QALYs have been used in a decision analysis to evaluate various health technologies.<sup>71–113</sup>

Weinstein and Feinberg<sup>69</sup> provide a definitive guide to the methodology of clinical decision analysis, while others discuss the use of Markov models in this context.<sup>114,115</sup> The basic model for a decision

analysis, which defines the clinical situation for which a decision is needed, is usually represented by a decision tree. For each treatment alternative, the sequence of possible events are represented by a series of decision nodes (events over which clinician has some control) and chance nodes (events over which there is no control). Transition probabilities are allocated to each chance event, using existing knowledge from previous studies, routinely collected data or subjective expert opinion. In the same way, utility values need to be assigned to each health state to reflect the quality of life in that state.

If the outcome measures for the decision analysis are QALYs, then estimates of life expectancy are combined with appropriate utility values. Life expectancy can be estimated using methods such as declining exponential approximation to life expectancy (or DEALE) and Gompertz models,<sup>116,117</sup> or methods based on Markov models, such as Markov cohort simulation and Monte Carlo simulation.<sup>114,115</sup> Quality-adjusted life expectancy can be calculated for each treatment option, with the treatment with the maximum value assumed to be optimal.

Sensitivity analyses are usually carried out to explore how possible ranges of values in terms of transition probabilities, utilities and costs may affect the conclusions. Considering the effects of changing more than one variable at a time, in a two- or three-way sensitivity analysis, for example,<sup>89</sup> may be more realistic but can become complex. Another complex approach, probabilistic sensitivity analysis, uses probability density functions for each variable to account for uncertainty.<sup>89</sup> Receiver–operator-characteristic curves can also be used in a sensitivity analysis to identify the most important variables for a decision problem.<sup>112</sup>

## Critical appraisal of QALYs

There has been much discussion about QALYs relating to both the theoretical and practical difficulties associated with their calculation, and controversy remains regarding their usage. A selection of references discussing QALYs are given here and further references can be found in the bibliography (page 109).

### Theoretical and practical difficulties

A major difficulty with QALYs is the derivation of appropriate quality-adjustment fractions. QALYs need to be theoretically and methodologically correct but they also need to be based on ‘good’

quality (i.e. accurate) data.<sup>6</sup> The basic concept that quality of life can be measured in cardinal numbers on a ratio scale could itself be questioned<sup>118</sup> but, assuming that quality of life is measurable, there is still the difficulty that quality-adjustment fractions need to account for the many different ways in which quality of life can be impaired.<sup>119</sup>

Quality-adjustment fractions need to be derived from an appropriate and reliable source. The extensive use of the Rosser valuation matrix<sup>8</sup> as quality-adjustment fractions in QALYs is questioned,<sup>119–121</sup> the quality ratings were based on judgements from a small sample of arbitrarily chosen respondents and were never intended to be used for calculating QALYs. A further difficulty in the derivation of quality-adjustment fractions is the requirement for them to satisfy ‘reciprocal commensurability’ between duration and quality of survival (called constant proportional trade-off on page 57), that is,  $x$  years of life at  $x^{-1}$  quality should be equivalent to 1 year of life at unimpaired quality.<sup>119</sup> There is some empirical evidence that validates this concept.<sup>122</sup>

There has been much discussion regarding the validity of the assumptions underlying the QALY model (see page 58), especially by authors advocating an alternative measure (see below). Carr-Hill<sup>123</sup> examines in detail the theoretical assumptions underlying the QALY procedure. There is some empirical evidence to show that none of the assumptions hold.<sup>124</sup> In particular, empirical evidence from Richardson and colleagues<sup>125</sup> casts doubt on the validity of the additive assumption in the usual QALY model. The advantage of the model assumptions is that they greatly simplify the measurement requirements<sup>126</sup> and alternatives, such as the healthy-years equivalent (HYE), which overcome the restriction of the assumptions are difficult to implement practically (see below). QALYs are based on utility theory and, since this is the source of some weaknesses of the methodology, regret theory and prospect theory have been suggested as alternatives.<sup>127</sup>

### Ethics and criticisms regarding the use of QALYs

There is much debate about the ethics of using QALYs.<sup>119,120,127–143</sup> When QALYs are used to determine the optimal treatment for patients, there appear to be no ethical arguments against them. However, when QALYs are used as an aid to resource allocation, that is, to prioritise which patients to treat, then there are many ethical arguments against them. Harris<sup>138</sup> believes in this context that they are “positively dangerous

and morally indefensible” and discusses the unfairness of using QALYs for resource allocation, with particular reference to the fact that they are ageist.

Another concern regarding the use of QALYs for resource allocation is that like should be compared with like;<sup>144</sup> if, in the calculation of QALYs for different health technologies, different types of valuational measures have been used, then it is not valid to make a direct comparison.<sup>65</sup> Furthermore, if QALYs are to be used in a social context, such as for resource allocation decisions, then the derivation of utilities for input into QALY calculations needs to be based on a social context rather than on an individual basis.<sup>145,146</sup>

## Alternative and related methods to QALYs

### Healthy-years equivalents

HYE were originally proposed by Mehrez and Gafni<sup>147</sup> as an alternative measure to QALYs. Like QALYs they combine two outcomes of interest: quality and quantity of life. QALYs were criticised for only partially incorporating patient preferences, with the utility approach used to obtain quality-adjustment fractions for each state separately in a patient’s health profile. HYE were proposed as superior since they obtain the utility for the **whole** health profile and, therefore, fully represent the individual’s preferences.

The advantages of using HYE are that they do not rely on the restrictive assumptions of the QALY model (see page 58) and they allow attitudes towards risk to be incorporated.<sup>124,147</sup> There has been much debate regarding these claims, together with comparisons between HYE and QALYs<sup>126,148–157</sup> (see Bibliography for further references).

HYE are calculated using a two-stage scheme involving two standard gamble questions (see page 56 for information on the standard gamble procedure). It has been argued<sup>126,149</sup> that this two-stage procedure is unnecessary since it is equivalent to a straightforward TTO (see page 56 for information on the TTO technique). One of the major problems with HYE, admitted by Mehrez and Gafni in their original paper,<sup>147</sup> is their feasibility. Individuals have to participate in much longer and more complex interviews, and the costs associated with this, together with the willingness of participants, may be prohibitive. The completeness, validity and reliability of the answers given in such a situation are questionable.

### Saved young life equivalent

Nord<sup>158</sup> proposed an alternative to QALYs for comparing health technologies that would overcome some of the problems associated with them. He suggested comparisons in terms of units of saving a young life and restoring the young life to full health, referred to as the saved young life equivalent (SAVE). It was chosen in the belief that most people would regard it as the maximum benefit that a single individual could obtain. Outcomes from interventions, described as fully as desired, could then be valued in terms of numbers of SAVES. Nord proposed that SAVES, which measure social value, should be used in addition to QALYs rather than to replace them.

This proposed measure was criticised for not being adequately defined, particularly in terms of what was meant by 'young' and also for disadvantaging older or disabled patients.<sup>159-161</sup> Nord defended his measure, saying that SAVES allow society to decide on the importance of age in valuing interventions.<sup>162</sup> The SAVE procedure is a specific form of a more general approach called the 'person-trade-off approach', proposed by Nord as an alternative to the QALY.<sup>163</sup>

### Quality-adjusted lives

Another alternative to the QALY is the quality-adjusted life (QAL), where treatments are assessed in terms of numbers of lives saved rather than length of life. Stevenson and colleagues<sup>164</sup> assessed neonatal intensive care for low birthweight babies using QALs as well as QALYs but only presented results based on QALs, claiming that the two measures gave similar results in this context. The method assessed neonatal intensive care purely in terms of numbers of surviving babies but each life was weighted according to the level of disablement of the child.

### Healthy life expectancy

Healthy life expectancy (or health expectancy) is an indicator of the health status of a **population**, combining mortality and morbidity into a single index, and is used in an epidemiological context. It has been defined as the number of further years of life in good health that someone of a specified age can, on average, expect to enjoy, given the age-specific rates of mortality and morbidity prevailing in the population.<sup>165</sup>

Healthy life expectancy is known under various different names depending on the definition of 'healthy'. Examples include disease-free life expectancy, disability-free life expectancy and quality-adjusted life expectancy. There has been

some discussion regarding terminology used in the field of healthy life expectancy.<sup>166,167</sup>

There are three different methods for calculating healthy life expectancy: the Sullivan method, the double decrement method and the multistate method. These are described and compared by Barendregt and colleagues.<sup>168</sup> All three methods are variations on the standard procedure for calculating life expectancy using life tables and each has different data requirements. The Sullivan method is the most widely used because it is the simplest and is the least demanding in terms of data requirements but it can be unreliable.<sup>168</sup>

Healthy life expectancy can be used to assess the impact of chronic diseases on the health of a population,<sup>169,170</sup> to assess the impact of an intervention on the health of a population,<sup>171</sup> or to examine the underlying trend in the health of a population.<sup>172</sup> If healthy life expectancy is to be used as a measure by which to compare populations then standardisation, in terms of definitions and instruments measuring morbidity, is required.<sup>165,173</sup> Conceptual and ethical issues regarding the use of healthy life expectancy have also been raised.<sup>174,175</sup>

### References

1. Fanshel S, Bush JW. A health-status index and its applications to health services outcomes. *Operat Res* 1970;**18**:1021-66.
2. Williams A. Economics of coronary artery bypass grafting. *BMJ* 1985;**291**:326-9.
3. Torrance GW. Social preferences for health states: an empirical evaluation of three measurement techniques. *Socio-Econ Plan Sci* 1976;**10**:129-36.
4. Yin DP, Forman HP, Langlotz CP. Evaluating health services: the importance of patients' preferences and quality of life. *Am J Roentgenol* 1995;**165**:1323-8.
5. Bakker C, van der Linden S. Health-related utility measurement: an introduction. *J Rheumatol* 1995;**22**:1197-9.
6. Coast J. Developing the QALY concept: exploring the problems of data acquisition. *Pharmacoeconomics* 1993;**4**:240-6.
7. Torrance GW. Measurement of health state utilities for economic appraisal. *J Health Econ* 1986;**5**:1-30.
8. Rosser R, Kind P. A scale of valuations of states of illness: is there a social consensus? *Int J Epidemiol* 1978;**7**:347-58.
9. Rosser RM, Watts VC. The measurement of hospital output. *Int J Epidemiol* 1972;**1**:361-8.

10. Beard MEJ, Inder AB, Allen JR, Hart DNJ, Heaton DC, Spearing RL. The costs and benefits of bone marrow transplantation. *N Z Med J* 1991;**104**:303–5.
11. Parker MJ, Myles JW, Anand JK, Drewett R. Cost–benefit analysis of hip fracture treatment. *J Bone Joint Surg Br* 1992;**74**:261–4.
12. Rawles J, Light J, Watt M. Quality-of-life in the 1st 100 days after suspected acute myocardial infarction: a suitable trial end-point. *J Epidemiol Community Health* 1992;**46**:612–16.
13. Rawles J, Light J, Watt M. Loss of quality-adjusted days as a trial end point: effect of early thrombolytic treatment in suspected myocardial infarction. *J Epidemiol Community Health* 1993;**47**:377–81.
14. Gudex C, Kind P. The QALY toolkit [discussion paper]. York: Centre for Health Economics, University of York; 1987.
15. Cole RP, Shakespeare V, Shakespeare P, Hobby JAE. Measuring outcome in low priority plastic surgery patients using quality-of-life indexes. *Br J Plast Surg* 1994;**47**:117–21.
16. Kerridge RK, Glasziou PP, Hillman KM. The use of quality-adjusted life years (QALYs) to evaluate treatment in intensive care. *Anaesth Intensive Care* 1995;**23**:322–31.
17. Coast J. Reprocessing data to form QALYs. *BMJ* 1992;**305**:87–90.
18. Ranaboldo CJ, Chant ADB. Reprocessing data to form QALYs [letter]. *BMJ* 1992;**305**:424.
19. Harwood R. Reprocessing data to form QALYs [letter]. *BMJ* 1992;**305**:424.
20. Drewett RF, Minns RJ, Sibly TF. Measuring outcome of total knee replacement using quality-of-life indexes. *Ann R Coll Surg Engl* 1992;**74**:286–90.
21. Gudex C. QALYs and their use by the Health Service. Discussion Paper 20. York: Centre for Health Economics, University of York; 1986.
22. Parker MJ. Hip fractures and QALYs [reply]. *J Bone Joint Surg Br* 1993;**75**:164.
23. Fordham R. Hip fractures and QALYs. *J Bone Joint Surg Br* 1993;**75**:163–4.
24. Donaldson C, Atkinson A, Bond J, Wright K. QALYs and long-term care for elderly people in the UK: scales for assessment of quality of life. *Age Ageing* 1988;**17**:379–87.
25. Carr-Hill RA, Morris J. Current practice in obtaining the Q in QALYs: a cautionary note. *BMJ* 1991;**303**:699–701.
26. Williams A. The role of the EuroQol instrument in QALY calculations. Discussion Paper 130. York: Centre for Health Economics, University of York; 1995.
27. Sackett DL, Torrance GW. The utility of different health states as perceived by the general public. *J Chron Dis* 1978;**31**:697–704.
28. Ashby J, O’Hanlon M, Buxton MJ. The time trade-off technique: how do the valuations of breast cancer patients compare to those of other groups? *Qual Life Res* 1994;**3**:257–65.
29. van der Donk J, Levendag PC, Kuijpers AJ, Roest FHJ, Habbema JDF, Meeuwis CA, *et al.* Patient participation in clinical decision making for treatment of T3 laryngeal cancer: a comparison of state and process utilities. *J Clin Oncol* 1995;**13**:2369–78.
30. von Neumann J, Morgenstern O. Theory of games and economic behaviour. 3rd ed. New York: Wiley; 1953.
31. Wakker P, Stiggelbout A. Explaining distortions in utility elicitation through the rank-dependent model for risky choices. *Med Decis Making* 1995;**15**:180–6.
32. Torrance GW, Thomas WH, Sackett DL. A utility maximisation model for the evaluations of health care programs. *Health Serv Res* 1972;**7**:118–33.
33. Stiggelbout AM, Kiebert GM, Kievit J, Leer JWH, Habbema JDF, de Haes JCJM. The utility of the time trade-off method in cancer patients: feasibility and proportional trade-off. *J Clin Epidemiol* 1995;**48**:1207–14.
34. Sutherland HJ, Llewellyn-Thomas H, Boyd NF, Till JE. Attitudes toward quality of survival: the concept of ‘maximum endurable time’. *Med Decis Making* 1982;**2**:299–309.
35. Stalmeier PFM, Bezembinder TGG, Unic IJ. Proportional heuristics in time tradeoff and conjoint measurement. *Med Decis Making* 1996;**16**:36–44.
36. Mauldin PD, Weintraub WS. Determining patient utility following coronary revascularization. *J Invasive Cardiol* 1993;**5**:171–8.
37. Torrance GW, Furlong W, Feeny D, Boyle M. Multiattribute preference functions: health utilities index. *Pharmacoeconomics* 1995;**7**:503–20.
38. Stiggelbout AM, Kiebert GM, Kievit J, Leer JWH, Stoter G, de Haes JCJM. Utility assessment in cancer patients: adjustment or time tradeoff scores for the utility of life years and comparison with standard gamble scores. *Med Decis Making* 1994;**14**:82–90.
39. Kaplan RM, Feeny D, Revicki DA. Methods for assessing relative importance in preference based outcome measures. *Qual Life Res* 1993;**2**:467–75.
40. Hornberger JC, Redelmeier DA, Petersen J. Variability among methods to assess patients well-being and consequent effect on a cost-effectiveness analysis. *J Clin Epidemiol* 1992;**45**:505–12.

41. Torrance GW. Utility approach to measuring health-related quality of life. *J Chronic Dis* 1987;**40**:593–600.
42. Daly E, Gray A, Barlow D, McPherson K, Roche M, Vessey M. Measuring the impact of menopausal symptoms on quality-of-life. *BMJ* 1993;**307**:836–40.
43. Nease RF Jr, Tsai R, Hynes LM, Littenberg B. Automated utility assessment of global health. *Qual Life Res* 1996;**5**:175–82.
44. Barr RD, Feeny D, Furlong W, Weitzman S, Torrance GW. A preference-based approach to health-related quality-of-life for children with cancer. *Int J Pediatr Hematol/Oncol* 1995;**2**:305–15.
45. Brazier J, Deverill M, Harper R, Booth A. A review of the use of health status measures in economic evaluation. *Health Technol Assess* 1999;**3**(9).
46. Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. *Stat Med* 1990;**9**:1259–76.
47. Eraker SA, Baker S, Miyamoto JM. Parameter estimates for a quality adjusted life year utility model. *Clin Res* 1984;**32**:A294.
48. Johannesson M, Pliskin JS, Weinstein MC. A note on QALYs, time tradeoff, and discounting. *Med Decis Making* 1994;**14**:188–93.
49. Ganiats TG, Browner DK, Kaplan RM. Comparison of two methods of calculating quality-adjusted life years. *Qual Life Res* 1996;**5**:162–4.
50. Tousignant P, Cosio MG, Levy RD, Groome PA. Quality-adjusted life years added by treatment of obstructive sleep apnea. *Sleep* 1994;**17**:52–60.
51. Goodwin PJ, Field R, Evans WK, Pater J. Cost-effectiveness of cancer chemotherapy: an economic evaluation of a randomised trial in small-cell lung cancer. *J Clin Oncol* 1988;**6**:1537–47.
52. Oldridge N, Furlong W, Feeny D, Torrance G, Guyatt G, Crowe J, *et al*. Economic evaluation of cardiac rehabilitation soon after acute myocardial infarction. *Am J Cardiol* 1993;**72**:154–61.
53. Drummond MF, Heyse J, Cook J, McGuire A. Selection of end-points in economic evaluations of coronary heart disease interventions. *Med Decis Making* 1993;**13**:184–90.
54. Smith RD, Hall J, Gurney H, Harnett PR. A cost–utility approach to the use of 5-fluorouracil and levamisole as adjuvant chemotherapy for Duke C colonic carcinoma. *Med J Aust* 1993;**158**:319–22.
55. Malter AD, Larson EB, Urban N, Deyo RA. Cost-effectiveness of lumbar discectomy for the treatment of herniated intervertebral disc. *Spine* 1996;**21**:1048–54.
56. Ramsey SD, Patrick DL, Albert RK, Larson EB, Wood DE, Raghu G. The cost-effectiveness of lung transplantation: a pilot study. *Chest* 1995;**108**:1594–601.
57. Ziegler F. Cost–utility in CNS drug trials. *Eur Psychiatry* 1996;**11**:159–64.
58. Uther JFB. The automatic implantable defibrillator is the most realistic and cost-effective way of preventing sudden cardiac death. *Aust N Z J Med* 1992;**22**:636–8.
59. Lambert CM, Hurst NP. Health economics as an aspect of health outcome: basic principles and application in rheumatoid arthritis. *Br J Rheumatol* 1995;**34**:774–80.
60. Holtgrave DR, Qualls NL. Threshold analysis and programs for prevention of HIV infection. *Med Decis Making* 1995;**15**:311–17.
61. StLeger AS, Spencely M, McCollum CN, Mossa M. Screening for abdominal aortic aneurysm: a computer assisted cost–utility analysis. *Eur J Vasc Endovasc Surg* 1996;**11**:183–90.
62. de Koning HJ, van Ineveld BM, van Oortmarssen GJ, de Haes JCJM, Collette HJA, Hendriks JHCL, *et al*. Breast cancer screening and cost-effectiveness: policy alternatives, quality-of-life considerations and the possible impact of uncertain factors. *Int J Cancer* 1991;**49**:531–7.
63. Nease RF, Owens DK. A method for estimating the cost effectiveness or incorporating patient preferences into practice guidelines. *Med Decis Making* 1994;**14**:382–92.
64. Kawachi I, Bethwaite P, Bethwaite J. The use of quality-adjusted life years (QALYs) in the economic appraisal of health care. *N Z Med J* 1990;**103**:46–8.
65. Fletcher A. Pressure to treat and pressure to cost: a review of cost-effectiveness analysis. *J Hypertens* 1991;**9**:193–8.
66. Spiegelhalter DJ, Gore SM, Fitzpatrick R, Fletcher AE, Jones DR, Cox DR. Quality-of-life measures in health care. 3. Resource allocation. *BMJ* 1992;**305**:1205–9.
67. Petrou S, Renton A. The QALY – a guide for the public health physician. *Public Health* 1993;**107**:327–36.
68. Robinson R. Economic evaluation and health care. 4. Cost–utility analysis. *BMJ* 1993;**307**:859–62.
69. Weinstein MC, Fineberg HV. Clinical decision analysis. Philadelphia: WB Saunders; 1980.
70. Nettleman MD. Outcome measurements in decision analysis: life versus quality of life. *Infect Control Hosp Epidemiol* 1989;**10**:521–4.
71. Brasel KJ, Weigelt JA. Blunt thoracic aortic trauma: a cost–utility approach for injury detection. *Arch Surg* 1996;**131**:619–25.
72. Bennett CL, Matchar D, McCrory D, McLeod DG, Crawford ED, Hillner BE. Cost-effective models for flutamide for prostate carcinoma patients: are they helpful to policy makers. *Cancer* 1996;**77**:1854–61.

73. Fiscella K, Franks P. Cost-effectiveness of the transdermal nicotine patch as an adjunct to physicians smoking cessation counselling. *JAMA* 1996;**275**:1247–51.
74. Chang RW, Pellissier JM, Hazen GB. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. *JAMA* 1996;**275**:858–65.
75. Hillner BE, Desch CE, Carlson RW, Smith TJ, Esserman L, Bear HD. Trade-offs between survival and breast preservation for 3 initial treatments of ductal carcinoma in-situ of the breast. *J Clin Oncol* 1996;**14**:70–7.
76. Kiberd BA, Jindal KK. Screening to prevent renal failure in insulin-dependent diabetic patients: an economic evaluation. *BMJ* 1995;**311**:1595–9.
77. Javitt JC, Aiello LP. Cost-effectiveness of detecting and treating diabetic retinopathy. *Ann Intern Med* 1996;**124**:164–9.
78. Adams PC, Kertesz AE, Valberg LS. Screening for hemochromatosis in children of homozygotes: prevalence and cost-effectiveness. *Hepatology* 1995;**22**:1720–7.
79. Gage BF, Cardinalli AB, Albers GW, Owens DK. Cost-effectiveness of warfarin and aspirin for prophylaxis of stroke in patients with nonvalvular atrial fibrillation. *JAMA* 1995;**274**:1839–45.
80. Tsevat J, Duke D, Goldman L, Pfeffer MA, Lamas GA, Soukup JR, *et al*. Cost-effectiveness of captopril therapy after myocardial infarction. *J Am Coll Cardiol* 1995;**26**:914–19.
81. Kellett J, Clarke J. Comparison of accelerated tissue plasminogen activator with streptokinase for treatment of suspected myocardial infarction. *Med Decis Making* 1995;**15**:297–310.
82. Velanovich V. Immediate biopsy versus observation for abnormal findings on mammograms: an analysis of potential outcomes and costs. *Am J Surg* 1995;**170**:327–32.
83. Groot CUD, Hagenbeek A, Verdonck LF, Lowenberg B, Rutten FFH. Cost-effectiveness of ABMT in comparison with CHOP chemotherapy in patients with intermediate and high grade malignant non-hodgkins lymphoma (NHL). *Bone Marrow Transplant* 1995;**16**:463–70.
84. Velanovich V. Blunt splenic injury in adults: a decision analysis comparing options for treatment. *Eur J Surg* 1995;**161**:463–70.
85. Dippel DWJ, van Crevel H, Lindsay KW, Hijdra A, Habbema JDF. Management of subarachnoid hemorrhage: a decision analysis. *Cerebrovasc Dis* 1995;**5**:350–7.
86. King JT, Glick HA, Mason TJ, Flamm ES. Elective surgery for asymptomatic, unruptured, intracranial aneurysms: a cost-effectiveness analysis. *J Neurosurg* 1995;**83**:403–12.
87. Rouse DJ, Owen J, Goldenberg RL, Vermund SH. Zidovudine for the prevention of vertical HIV transmission: a decision-analytic approach. *J Acquir Immune Defic Syndr Hum Retrovirol* 1995;**9**:401–7.
88. Cantor SB, Spann SJ, Volk RJ, Cardenas MP, Warren MM. Prostate cancer screening: a decision analysis. *J Fam Pract* 1995;**41**:33–41.
89. Gabriel SE. Economic evaluation using mathematical models: the case of misoprostol prophylaxis. *J Rheumatol* 1995;**22**:1412–14.
90. Hunink MGM, Wong JB, Donaldson MC, Meyerovitz MF, de Vries J, Harrington DP. Revascularization for femoropopliteal disease: a decision and cost-effectiveness analysis. *JAMA* 1995;**274**:165–71.
91. Kellett J, Graham J. Referral for coronary angiography after exercise testing: traditional decision making versus decision analysis. *Mon J Assoc Physicians* 1995;**88**:401–8.
92. Wong JB, Koff RS, Tine F, Pauker SG. Cost-effectiveness of interferon-alpha-2b treatment for hepatitis-B antigen-positive chronic hepatitis-B. *Ann Intern Med* 1995;**122**:664–75.
93. Kamlet MS, Paul N, Greenhouse J, Kupfer D, Frank E, Wade M. Cost-utility analysis of maintenance treatment for recurrent depression. *Control Clin Trials* 1995;**16**:17–40.
94. Etchason J, Petz L, Keeler E, Calhoun L, Kleinman S, Snider C, *et al*. The cost-effectiveness of pre-operative autologous blood donations. *N Engl J Med* 1995;**332**:719–24.
95. Yin DP, Baum RA, Carpenter JP, Langlotz CP, Pentecost MJ. Cost-effectiveness of MR angiography in cases of limb threatening peripheral vascular disease. *Radiology* 1995;**194**:757–64.
96. Patterson RE, Eisner RL, Horowitz SF. Comparison of cost-effectiveness and utility of exercise ECG, single photon emission computed tomography, positron emission tomography, and coronary angiography for diagnosis of coronary artery disease. *Circulation* 1995;**91**:54–65.
97. Provenzale D, Kemp JA, Arora S, Wong JB. A guide for surveillance of patients with Barrett's esophagus. *Am J Gastroenterol* 1994;**89**:670–80.
98. Brown ML, Nayfield SG, Shibley LM. Adjuvant therapy for stage III colon cancer: economics returns to research and cost-effectiveness of treatment. *J Natl Cancer Inst* 1994;**86**:424–30.
99. Jonsson B, Karlsson G. Outcome measurement in economic evaluations of growth factors. *Eur J Cancer* 1993;**29A**:S19–22.
100. Javitt J, Cas RD, Chiang YP. Cost-effectiveness of screening and cryotherapy for threshold retinopathy of prematurity. *Pediatrics* 1993;**91**:859–66.

101. Hillner BE, Smith TJ. Should women with node-negative breast cancer receive adjuvant chemotherapy: insights from a decision analysis model. *Breast Cancer Res Treat* 1992;**23**:17–27.
102. Katz DA, Littenberg B, Cronenwett JL. Management of small abdominal aortic aneurysms: early surgery vs watchful waiting. *JAMA* 1992;**268**:2678–86.
103. Naglie IG, Detsky AS. Treatment of chronic nonvalvular atrial fibrillation in the elderly: a decision analysis. *Med Decis Making* 1992;**12**:239–49.
104. Mold JW, Holtgrave DR, Bisonni RS, Marley DS, Wright RA, Spann SJ. The evaluation and treatment of men with asymptomatic prostate nodules in primary care: a decision analysis. *J Fam Pract* 1992;**34**:561–8.
105. Dippel DWJ, Touwotten F, Habbema JDF. Management of children with acute pharyngitis: a decision analysis. *J Fam Pract* 1992;**34**:149–59.
106. de Haes JCJM, de Koning HJ, van Oortmarssen GJ, van Agt HME, de Bruyn AE, van der Maas PJ. The impact of a breast cancer screening program on quality-adjusted life years. *Int J Cancer* 1991;**49**:538–44.
107. Verhoef LCG, Stalpers LJA, Verbeek ALM, Wobbes T, van Daal WAJ. Breast conserving treatment or mastectomy in early breast cancer: a clinical decision analysis with special reference to the risk of local recurrence. *Eur J Cancer* 1991;**27**:1132–7.
108. Jacobson JJ, Schweitzer SO, Kowalski CJ. Chemoprophylaxis of prosthetic joint patients during dental treatment: a decision utility analysis. *Oral Surg Oral Med Oral Pathol* 1991;**72**:167–77.
109. Krahn M, Naylor CD, Basinski AS, Detsky AS. Comparison of an aggressive (United States) and a less aggressive (Canadian) policy for cholesterol screening and treatment. *Ann Intern Med* 1991;**115**:248–55.
110. Weeks JC, Tierney MR, Weinstein MC. Cost-effectiveness of prophylactic intravenous immune globulin in chronic lymphocytic leukemia. *N Engl J Med* 1991;**325**:81–6.
111. Kawachi I, Purdie G. Should treatment of mild to moderate hypertension be targeted: results from a Markov cohort model incorporating multiple risk factors. *J Hum Hypertens* 1990;**4**:651–8.
112. Downs SM, McNutt RA, Margolis PA. Management of infants at risk for occult bacteremia: a decision analysis. *J Pediatr* 1991;**118**:11–20.
113. Hillner BE, Smith TJ. Efficacy and cost-effectiveness of adjuvant chemotherapy in women with node-negative breast cancer: a decision analysis model. *N Engl J Med* 1991;**324**:160–8.
114. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. *Med Decis Making* 1993;**13**:322–8.
115. Beck JR, Pauker SG. The Markov process in medical prognosis. *Med Decis Making* 1983;**3**:419–58.
116. Keeler E, Bell R. New DEALEs: other approximations of life expectancy. *Med Decis Making* 1992;**12**:307–11.
117. Kuntz KM, Weinstein MC. Life expectancy biases in clinical decision modelling. *Med Decis Making* 1995;**15**:158–69.
118. James KW. The morality of the QALY. *BMJ* 1988;**296**:213.
119. Smith A. Qualms about QALYs. *Lancet* 1987;**i**:1134–6.
120. Rawles J. Castigating QALYs. *J Med Ethics* 1989;**15**:143–7.
121. Naji SA, Sheldon TA. The measurement of patient outcomes. *Proc R Soc Edinb B Biol Sci* 1993;**101**:55–75.
122. Evans RW. Quality-adjusted life-years. *Lancet* 1987;**i**:1372.
123. Carr-Hill RA. Assumptions of the QALY procedure. *Soc Sci Med* 1989;**29**:469–77.
124. Mehrez A, Gafni A. The healthy-years equivalents: how to measure them using the standard gamble approach. *Med Decis Making* 1991;**11**:140–6.
125. Richardson J, Hall J, Salkeld G. The measurement of utility in multiphase health states. *Int J Technol Assess Health Care* 1996;**12**:151–62.
126. Wakker P. A criticism of healthy-years equivalents. *Med Decis Making* 1996;**16**:207–14.
127. Mooney G. QALYs – are they enough? A health economist's perspective. *J Med Ethics* 1989;**15**:148–52.
128. Rawles J, Rawles K. The QALY argument: a physician's and a philosopher's view. *J Med Ethics* 1990;**16**:93–4.
129. la Puma J, Lawlor EF. Quality-adjusted life-years: ethical implications for physicians and policy makers. *JAMA* 1990;**263**:2917–21.
130. Kaplan RM, Ganiats TG. QALYs: their ethical implications [letter]. *JAMA* 1990;**264**:2502–3.
131. Danford DA. QALYs: their ethical implications [letter]. *JAMA* 1990;**264**:2503.
132. la Puma J. QALYs: their ethical implications [reply]. *JAMA* 1990;**264**:2503–4.
133. Williams A. Quality-adjusted life-years. *Lancet* 1987;**i**:1372.
134. Drummond MF. Quality-adjusted life-years. *Lancet* 1987;**i**:1372–3.
135. Smith A. Quality-adjusted life-years. *Lancet* 1987;**ii**:47.
136. Cubbon J. The principle of QALY maximization as the basis for allocating health care resources. *J Med Ethics* 1991;**17**:181–4.

137. Harris J. Unprincipled QALYs: a response to Cubbon. *J Med Ethics* 1991;**17**:185–8.
138. Harris J. QALYfying the value of life. *J Med Ethics* 1987;**13**:117–23.
139. Williams A. QALYfying the value of life [response]. *J Med Ethics* 1987;**13**:123.
140. Jones R. What price QALYs? *J R Coll Gen Pract* 1988;**38**:231.
141. Singer P, McKie J, Kuhse H, Richardson J. Double jeopardy and the use of QALYs in health care allocation. *J Med Ethics* 1995;**21**:144–50.
142. Harris J. Double jeopardy and the veil of ignorance – a reply. *J Med Ethics* 1995;**21**:151–7.
143. Williams A. Economics, QALYs and medical ethics: a health economist's perspective. Discussion Paper 121. York: Centre for Health Economics, University of York; 1994.
144. McTurk L. A methodological quibble about QALYs. *BMJ* 1991;**302**:1601.
145. Nord E. The QALY: a measure of social value rather than individual utility? *Health Econ* 1994;**3**:89–93.
146. Nord E. Health status index models for use in resource allocation decisions: a critical review in the light of observed preferences for social choice. *Int J Technol Assess Health Care* 1996;**12**:31–44.
147. Mehrez A, Gafni A. Quality-adjusted life years, utility theory, and healthy-years equivalents. *Med Decis Making* 1989;**9**:142–9.
148. Fryback DG. QALYs, HYE, and the loss of innocence. *Med Decis Making* 1993;**13**:271–2.
149. Johannesson M, Pliskin JS, Weinstein MC. Are healthy-years equivalents an improvement over quality-adjusted life years? *Med Decis Making* 1993;**13**:281–6.
150. Mehrez A, Gafni A. Healthy-years equivalents versus quality-adjusted life years: in pursuit of progress. *Med Decis Making* 1993;**13**:287–92.
151. Gafni A. HYE: do we need them and can they fulfil the promise? *Med Decis Making* 1996;**16**:215–16.
152. Wakker P. HYE: rejoinder. *Med Decis Making* 1996;**16**:216.
153. Weinstein MC, Pliskin JS. Perspectives on health-years equivalents – HYE: what are the issues? *Med Decis Making* 1996;**16**:205–6.
154. Gafni A. Measuring the adverse effects of unnecessary hypertension drug therapy: QALYs vs HYE. *Clin Invest Med* 1991;**14**:266–70.
155. Culyer AJ, Wagstaff A. QALYs versus HYE: a theoretical exposition. Discussion Paper 99. York: Centre for Health Economics, University of York; 1992.
156. Johannesson M. The ranking properties of healthy-years equivalents and quality-adjusted life-years under certainty and uncertainty. *Int J Technol Assess Health Care* 1995;**11**:40–8.
157. Gafni A, Birch S. Guidelines for the adoption of new technologies – a prescription for uncontrolled growth in expenditures and how to avoid the problem. *Can Med Assoc J* 1993;**148**:913–17.
158. Nord E. An alternative to QALYs: the saved young life equivalent (SAVE). *BMJ* 1992;**305**:875–7.
159. Brown PB. An alternative to QALYs: saved young life equivalent (SAVE) [letter]. *BMJ* 1992;**305**:1365.
160. Glicher SR. Alternative to QALYs: saved young life equivalent (SAVE) [letter]. *BMJ* 1992;**305**:1365.
161. Sayers G. Alternative to QALYs: saved young life equivalent (SAVE) [letter]. *BMJ* 1992;**305**:1365.
162. Nord E. Alternative to QALYs: saved young life equivalent (SAVE) [reply]. *BMJ* 1992;**305**:1365–6.
163. Nord E. The person-trade-off approach to valuing health care programs. *Med Decis Making* 1995;**15**:201–8.
164. Stevenson RC, Pharoah POD, Cooke RWI, Sandhu B. Predicting costs and outcomes of neonatal intensive care for very low birth-weight infants. *Publ Health* 1991;**105**:121–6.
165. Bone MR. International efforts to measure health expectancy. *J Epidemiol Commun Health* 1992;**46**:555–8.
166. Robine JM, Mathers CD, Bucquet D. Distinguishing health expectancies and health-adjusted life expectancies from quality-adjusted life years. *Am J Publ Health* 1993;**83**:797–8.
167. Erickson P, Wilson RW. The term years of healthy life: misunderstood, defended, and challenged. A short term for quality-adjusted life years. *Am J Publ Health* 1994;**84**:866.
168. Barendregt JJ, Bonneux L, van der Maas PJ. Health expectancy: an indicator for change. *J Epidemiol Commun Health* 1994;**48**:482–7.
169. Reynolds DL, Torrance GW, Badley EM, Bennett KJ, Chambers LW, Goldsmith CH, et al. Modeling the population health impact of musculoskeletal diseases: arthritis. *J Rheumatol* 1993;**20**:1037–47.
170. Nusselder WJ, van der Velden K, van Sonsbeek JLA, Lenior ME, van den Bos GAM. The elimination of selected chronic diseases in a population: the compression and expansion of morbidity. *Am J Publ Health* 1996;**86**:187–94.
171. Kawachi I, Malcolm LA. The benefits of treating mild to moderate hypertension. A quantitative estimation of the live expectancy gains from pharmacological reduction of blood pressure. *J Clin Epidemiol* 1989;**42**:905–12.



172. Wilkins R, Adams OB. Changes in the healthfulness of life of the elderly population: an empirical approach. *Rev Epidemiol Sante Publique* 1987;**35**:225–35.
173. Robine JM, Ritchie K. Healthy life expectancy: evaluation of global indicator of change in population health. *BMJ* 1991;**302**:457–60.
174. Robine JM, Michel JP, Branch LG. Measurement and utilization of healthy life expectancy: conceptual issues. *Bull World Health Organ* 1992;**70**:791–800.
175. Morrow RH, Bryant JH. Health policy approaches to measuring and valuing human life: conceptual and ethical issues. *Am J Public Health* 1995;**85**:1356–60.



## Chapter 16

# Quality-adjusted survival analysis

### Introduction

Quality-adjusted survival analysis can be used as an alternative to survival analysis in assessing health technologies in studies in which quality of life and survival are both important endpoints. Adjustments are made to account for the quality of life experienced during the survival time using approaches based on the QALY model (see chapter 15).

There are two main approaches to quality-adjusted survival analysis depending on the level of aggregation of the data. The subject-based approach combines quality-of-life and survival data at the patient level, thus creating a single endpoint for each subject on which to compare health technologies, while the population-based approach aggregates the quality-of-life and survival data at a (pre-defined) group level.

The first part of this chapter covers both subject-based and population-based approaches to quality-adjusted survival analysis using QALYs. A special form of QALY, known as TWiST is described; the analysis of this endpoint is discussed in terms of a subject-based approach. An extension of the TWiST concept, Q-TWiST, is also described; here the analysis takes the population-based approach.

### Quality-adjusted survival analysis using a QALY model

#### Subject-based approach

In any study assessing health technologies, QALYs can be calculated for each subject by weighting periods of survival time according to the quality of life experienced during these periods (see chapter 15). The sum of these weighted survival times creates a single endpoint for each subject for which health technologies can then be compared.

If survival times are known for all the patients in a study (i.e. none are censored), then standard techniques for dealing with continuous data can be used.<sup>1</sup> The sample mean and standard deviation of the QALYs for the patients in each treatment arm could be used to provide an estimate of the mean treatment difference in QALYs, together with an

appropriate CI. The hypothesis of no difference between treatments in terms of QALYs can be tested using a *t* test, for example, if distributional assumptions were satisfied.

If the data for some patients are censored, then a standard survival analysis using QALYs for each individual as an endpoint rather than actual survival time may seem appropriate. However, there will be a problem of informative censoring (see page 11). Survival time for patients with poor quality of life will receive a lower weighting than that for patients with good quality of life. Patients with poor quality of life will thus accumulate QALYs at a slower rate and will, therefore, be censored earlier on the **QALY timescale** than those with good quality of life. This will give biased Kaplan–Meier estimates and invalidate the log-rank test, as well as other standard survival analysis techniques, as a means of comparing treatments.

In the situation of informative censoring, there are two options available for a valid analysis.

1. The censoring date for the analysis could be set as the smallest censored survival time value, thus restricting the analysis to a period during which all subjects have full follow-up and eliminating censoring. This is only feasible if the smallest censored value is quite large, otherwise a considerable number of events may be lost, thus reducing the statistical power of the analysis.
2. A population-based approach such as that based on the QALY model or the Q-TWiST model may be used. Using this approach, survival analysis is applied to unweighted survival times, for which censoring is non-informative, and the quality-of-life weightings are then applied to treatment group averages obtained from this unbiased survival analysis.

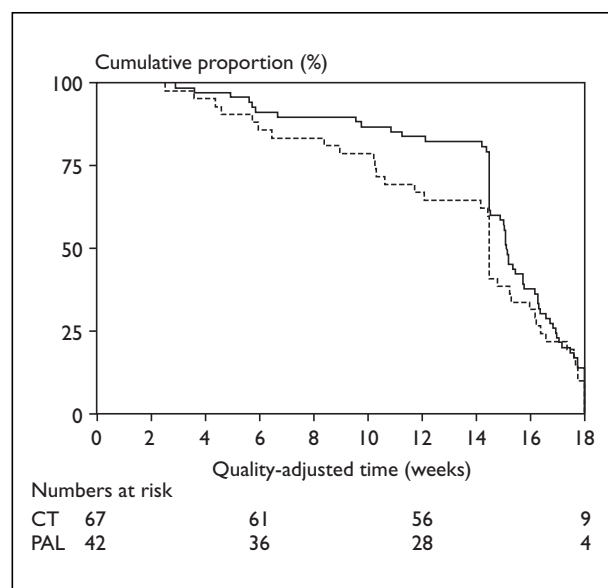
#### Example using the MIC data

Quality-of-life data was restricted to the 18 weeks from entry to the study, hence a quality-adjusted survival analysis was only possible over this period. An unadjusted survival analysis (see page 44) showed a statistically significant difference in survival within 18 weeks of entry to trial (log-rank test  $\chi^2 = 8.07$ ,  $p = 0.005$ ).

QALYs within 18 weeks of study entry were calculated for each individual (see page 58). All were followed-up for at least 18 weeks; hence, there were no censored survival times within this period. Quality-adjusted survival analysis using QALYs for each patient as an endpoint is thus straightforward since, with no censored data, informative censoring is not a problem. Restricting the analysis of the MIC data to 18 weeks from study entry, because of the availability of quality-of-life data, has essentially resulted in overcoming the informative censoring problem in a way equivalent to the first option discussed above.

With no censored data, treatment comparisons can be made using either standard techniques for continuous measures or standard survival techniques without concern about bias. Since the data distribution was not normal, standard survival analysis techniques were used to compare treatments. Kaplan–Meier survival curves for quality-adjusted survival (see *Figure 18*) showed chemotherapy to be superior to palliative therapy but the difference was not of statistical significance (log-rank test  $\chi^2 = 1.33$ ,  $p = 0.25$ ). It should be noted that the steps on these Kaplan–Meier curves do not represent deaths, as they would for a standard survival endpoint, but show reductions in the proportion of patients achieving increasing levels of quality-adjusted survival time.

In conclusion, although treatments differed in terms of survival within 18 weeks, they were not found to differ in terms of quality-adjusted survival. This is based on the use of a value of 0.8 to weight



**FIGURE 18** Quality-adjusted survival analysis of the MIC data using QALYs (—, CT; - - -, PAL)

periods with poor quality of life. Sensitivity analysis should be used to establish the robustness of conclusions to the choice of weighting factor used to calculate the QALY endpoint for each patient.

### Population-based approach

The subject-based approach to quality-adjusted survival analysis using QALYs can be problematic and lead to biased results in situations with censored survival times. It may therefore be preferable to use a population-based approach. This method combines the quality-of-life and survival data at the population or group level rather than at the subject level.

Essentially, a quality-adjusted survival curve for a population is formed by plotting, against time  $t$ , the product of the mean quality of life of patients living at time  $t$  and the probability of surviving to time  $t$ . The area under this quality-adjusted survival curve gives the mean quality-adjusted survival for the population. Beacon<sup>2</sup> refers to this method as the **integrated quality–survival product** and illustrates the method using quality-of-life data.

In situations in which quality of life is measured at  $k$  discrete time points,  $t_1, t_2, \dots, t_k$ , the following function is an estimator of the expected quality-adjusted survival time (QAS).<sup>3</sup>

$$\text{QAS} = \sum_{i=1}^k ((Q_i + Q_{i+1})/2) ((S_i + S_{i+1})/2) (t_{i+1} - t_i) \quad (21)$$

where  $Q_i$  ( $i = 1-k$ ) is an estimate of the mean quality of life at time  $t_i$  and  $S_i$  ( $i = 1-k$ ) is an estimate of the probability of survival to time  $t_i$ . In this model, the quality-of-life and survival probabilities in the interval between two time points are assumed to be the average of the values at the time points at each end of the interval.

Hwang and colleagues<sup>3</sup> suggest that survival probabilities could be estimated using standard methods such as the life table method, Kaplan–Meier estimates, or by using parametric models; they also suggest estimating the mean quality of life at the given times using kernel smoothing methods. Calculation of the standard error of the mean quality-adjusted survival time is mathematically complex and here they suggest using bootstrap methods<sup>4,5</sup> (see page 77).

Ganiats and colleagues<sup>6</sup> identified and compared three methods for estimating the mean quality of life and survival probabilities for input into

equation 21. The simplest method is to use the mean of the quality-of-life data at time  $t_i$  and the proportion of subjects surviving this time as the crude survival estimate. A refinement of this method is to use Kaplan–Meier estimates of the probability of survival to time  $t_i$  instead of the crude estimates. A third suggested method is to use maximum likelihood estimates of the quality of life and the probability of survival at each time point.

Each method has both advantages and disadvantages.<sup>6</sup> In summary, although the maximum likelihood method is theoretically the most appealing, it is technically the most difficult to implement and software is not readily available. Otherwise the method using Kaplan–Meier estimates is generally preferable to the crude method but has the disadvantage that calculation of the CIs for the mean quality-adjusted survival time is not straightforward, requiring bootstrap methods<sup>4,5</sup> (see page 77).

## TWiST

### Introduction

A special QALY endpoint for comparing therapies, which incorporates both length and quality of survival into a single measure, has been developed by Gelber and colleagues<sup>7</sup> in a subject-based approach to quality-adjusted survival analysis. The endpoint which they devised, TWiST, is a measure of the ‘good’ quality time experienced by the patient. It was originally developed to assess treatments for breast cancer,<sup>7–9</sup> and has also been used in the assessment of treatments for ovarian cancer.<sup>10,11</sup>

### Defining TWiST

TWiST is calculated for each patient by subtracting from overall survival those periods during which

treatment or disease reduces their quality of life. This is equivalent to calculating QALYs for a patient using utility values of 0 for times with symptoms (REL) and toxicity (TOX), and utility values of 1 otherwise.

Suppose that for patient  $i$ ,  $TR_i$  is the time from start of treatment to symptomatic disease relapse and  $TOX_i$  is the amount of time spent (not necessarily in consecutive periods) with toxicity prior to relapse (see Figure 19): then

$$TWiST_i = TR_i - TOX_i \quad (22)$$

If the follow-up for patient  $i$  is such that neither  $TR_i$  nor  $TOX_i$  have been observed, then TWiST for this patient will be censored and calculated as

$$TWiST_i = U_i - OTOX_i \quad (23)$$

where  $U_i$  is the length of follow-up for patient  $i$  and  $OTOX_i$  is the value of  $TOX_i$  observed during follow-up. The follow-up for this patient may be such that the actual value of  $TOX_i$  has been observed but  $TR_i$  has not, in which case TWiST for this patient will be censored and can be calculated as

$$TWiST_i = U_i - TOX_i \quad (24)$$

where  $U_i$  is the length of follow-up for patient  $i$ .

The definitions of time with symptoms of disease (REL) and time with toxicity of treatment (TOX) can be adjusted for different clinical situations depending on disease and treatment under study. Defining the untoward events that can occur and determining the importance attached to each one, in terms of the amount of time subtracted from total survival, are of paramount importance in creating a meaningful TWiST measure.<sup>12</sup>

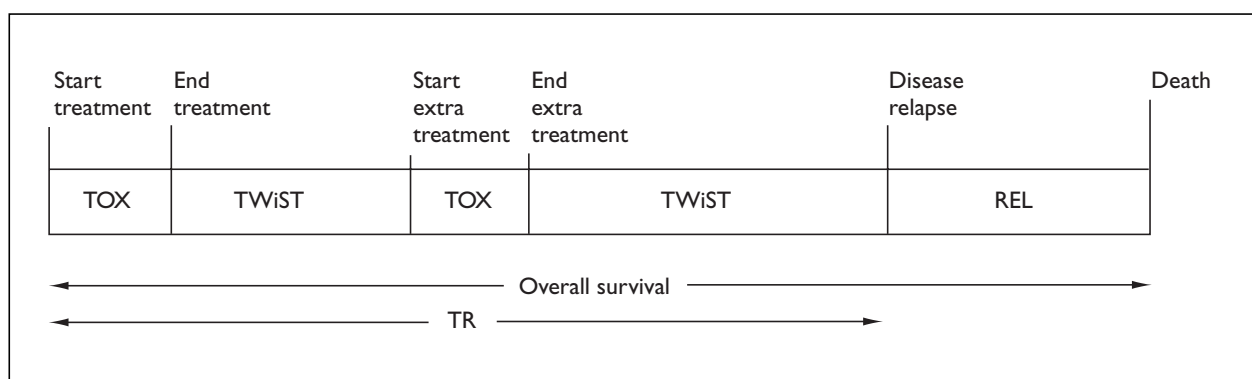


FIGURE 19 An example of TWiST

The definitions of TWiST in the literature have, in general, been based on clinical criteria rather than patient-based measures of quality of life. However, quality-of-life data collected via questionnaires at repeated assessments over time could be used to define TWiST. For example, Glimelius and colleagues<sup>13</sup> used average quality-of-life scores from various parts of their questionnaire to define TWiST for their study. They assigned a quality weighting factor of 1 to the survival time during which a patient had either an unchanged high quality of life and no signs of symptomatically progressive disease or improvement in quality-of-life estimates without being hospitalised; all other survival time was assigned a weighting factor of 0. In further examples,<sup>2,14</sup> survival time spent with 'normal' quality-of-life scores (as defined by the quality-of-life instrument) is used as the definition of TWiST.

### Comparing treatments using TWiST Redefining the endpoint for treatment comparisons

Standard methods of survival analysis, such as Kaplan–Meier, log-rank tests or Cox regression models (see chapters 11 and 12), could be used on the TWiST endpoint to compare treatments. However, as with the subject-based approach to quality-adjusted survival analysis using QALYs (see page 71), standard survival techniques will be invalid because of informative censoring (see page 11). For example, two patients may have been followed-up for the same length of time and both be censored but the one with longer duration of toxicity will be censored earlier than the one with shorter duration. This informative censoring results in the Kaplan–Meier estimates of TWiST being overestimates of the true (uncensored) TWiST.<sup>7</sup>

In order to reduce (or preferably eliminate) the amount of censoring, thus overcoming the problem of informative censoring, Gelber and colleagues<sup>7</sup> suggested using **accumulated TWiST** rather than TWiST as the endpoint for treatment comparisons. Accumulated TWiST, TWiST(L), is defined as the amount of TWiST observed within L time units from the start of treatment. If the cut-off time, L, is chosen such that all patients are followed-up beyond that time (i.e.  $L \leq U_i$  for all  $i$ ), then TWiST(L) is not censored and the standard Kaplan–Meier method could be used to estimate the survivor function of TWiST(L) without bias. In addition to sensitivity analysis on the choice of L, estimating TWiST(L) for various values of L may show over time how the delay in relapse justifies earlier toxic effects.

### Comparison of treatments using means

Treatment comparison may be based on the mean TWiST(L), that is, the mean TWiST achieved within time L from start of treatment. Gelber and colleagues<sup>7</sup> considered estimating mean TWiST(L) using the area under the Kaplan–Meier curve for TWiST(L) but, from simulations, they discovered that even when only 10–20% of TWiST(L) observations were censored, the positive bias of the Kaplan–Meier-derived means was unacceptably large.

A replacement procedure may be preferable to the Kaplan–Meier method when estimating mean TWiST(L).<sup>7</sup> In using replacement procedures, censored values of TWiST(L) are replaced with some value in the interval of possible values and then a simple average can be used as the estimate for mean TWiST(L). Standard parametric or non-parametric methods can then be used to compare the means for two treatments. Simulations showed that of the five possible replacement values considered by Gelber and colleagues,<sup>7</sup> all resulted in reasonable estimates of mean TWiST(L).

### Alternative methods of comparing treatments

As an alternative to comparing treatments using mean TWiST(L), Gelber and colleagues<sup>7</sup> suggested that treatments could be compared with respect to TWiST(L) using quantile distance plots. The quantile distance function, which represents the horizontal distance between the survivor functions for two treatments, is plotted against the population percentile.

Willemsse and colleagues<sup>10</sup> suggested assessing treatments using the ratio between TWiST and progression-free survival, the TWiST index. This measures the proportion of progression-free time with 'good' quality of life and can be used as a kind of cost–benefit ratio. Interpretation may be difficult in situations in which some patients have censored progression-free survival times.

### TWiST analysis of MIC data

In the MIC study, MAL was used to define TWiST. Periods within the first 18 weeks where no malaise was experienced were defined as TWiST. For each patient, these periods were summed to give a total value for TWiST. As with the QALY endpoint, the analysis of the TWiST endpoint in the MIC data is straightforward, since there are no censored survival times within the 18-week analysis period. Restricting the analysis of the MIC data to 18 weeks from study entry, for reasons of quality-of-life data availability, is effectively

equivalent to analysing accumulated TWiST(L) within 18 weeks, i.e. TWiST(18).

With no censored data, treatment comparisons can be made using either standard techniques for continuous measures or standard survival techniques without concern for bias. Since data were not normally distributed, standard survival analysis techniques were used to compare treatments. The area under the Kaplan–Meier survival curves (see *Figure 20*) give estimates of the mean TWiST(18) in each treatment group (mean = 6.6 weeks for CT arm and 6.5 weeks for PAL arm). The treatments were clearly very similar in terms of TWiST(18) (log-rank test  $\chi^2 = 0.04$ ,  $p = 0.85$ ). Thus, although treatments differed significantly in terms of survival within 18 weeks (see page 44), they did not differ with respect to TWiST(18).

### Critical appraisal of TWiST and suggested extensions

The TWiST model is a simplistic way of incorporating quality-of-life into a survival-type endpoint. Brunner<sup>15</sup> criticised the model for many reasons including the fact that the amounts of time deducted from overall survival to create the TWiST endpoint are arbitrary and that the model does not account for the quality-of-life experienced during these times. Brunner considered that the TWiST model obscured rather than clarified the problem of weighing up the gains and losses in quality of life associated with treatment, and advocated subjective judgement in preference.

Gelber and colleagues<sup>7</sup> suggested several generalisations to their TWiST methodology:

- QALYs could be determined for individual patients, based on personalised weightings
- constant proportions of TOX and REL could be added to TWiST to avoid equating these periods to death
- application of a discount factor to adjust the value of future gains relative to the present.

The second suggestion forms a basis for the Q-TWiST methodology (see below). The TWiST model may be preferred to Q-TWiST, since it avoids the subjective quantification of times with symptoms and toxicity.<sup>8</sup> In addition, the TWiST methodology is useful for clinical situations in which toxic therapies are administered to patients who are free of symptomatic disease.<sup>8,9</sup> In particular, TWiST may also be useful for solving the dilemma of treatment selection when disease-free survival differences are statistically significant but overall survival differences are not, since it deals with the fact that extensions to disease-free time may be at the expense of treatment toxicity.<sup>7</sup>

## Q-TWiST

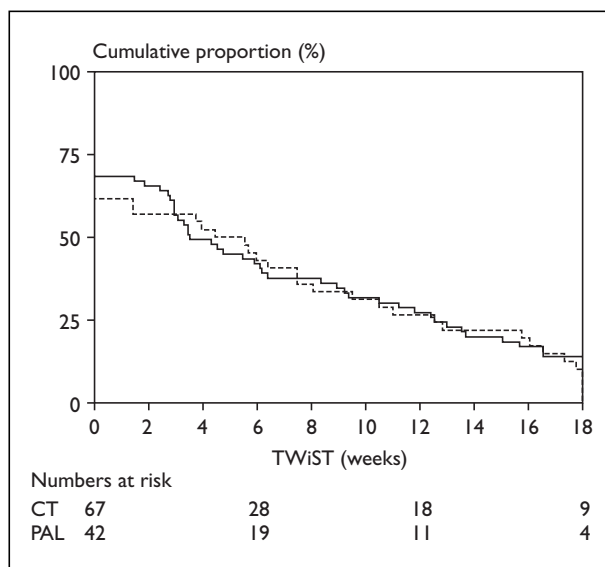
### Introduction

The Q-TWiST endpoint is a natural extension of the quality-of-life oriented endpoint TWiST and an adaptation of the concept of QALYs. TWiST methodology is extended so that periods spent with toxicity or relapse are included in the analysis but are weighted to represent their quality value relative to TWiST. Thus, overall survival is scaled downwards by arbitrarily giving survival during treatment or symptoms a reduced value.

Q-TWiST was originally developed and used to assess the effects of adjuvant therapy in women with breast cancer.<sup>1,16–18</sup> It has since been modified where necessary for use with other cancers, such as lung cancer,<sup>19</sup> lymphoma<sup>20</sup> and rectal cancer,<sup>21</sup> and in other diseases, such as AIDS.<sup>22</sup> Further work is being done to develop Q-TWiST so that it can be applied to neurological diseases such as multiple sclerosis<sup>23</sup> and epilepsy.<sup>24</sup>

Glasziou and colleagues<sup>1</sup> provide a generalised and more mathematical background to the Q-TWiST methodology introduced by Goldhirsch and colleagues.<sup>16</sup> Gelber and colleagues<sup>25</sup> provide an updated overview of the Q-TWiST method that includes all recent developments and extensions to the method.

There are three main steps in applying the Q-TWiST methodology:



**FIGURE 20** Quality-adjusted survival analysis of the MIC data using TWiST (—, CT; ---, PAL)

- (i) defining the Q-TWiST model, which includes defining the quality-of-life oriented health states and their associated weights
- (ii) partitioning the overall survival time into the defined health states in order to calculate the average time spent in each state
- (iii) comparing the treatments using Q-TWiST, calculated from a weighted sum of the average time spent in each state.

## Defining the Q-TWiST model

### Defining health states

During the course of a clinical trial it is assumed that the patient passes through a series of progressive health states that differ in terms of quality of life. The length of time spent in each health state may be affected by the treatment being received by the patient. It is a key assumption that patients progress through the health states in sequence but that any state may be skipped. The first step in a Q-TWiST analysis, therefore, is to define these health states so that they are clinically meaningful and such that they will highlight specific differences between the treatments being compared.

In general, any number and type of health states can be defined, say  $H_1, H_2, \dots, H_{k+1}$  where  $H_{k+1}$  usually represents death. In the original breast cancer application for which the method was developed,<sup>16</sup> the following health states were defined:

TOX	time having subjective toxic side-effects
TWiST	time without symptoms or toxicity
REL	time following systemic relapse (including time spent recovering from treatment for local recurrence).

The clinical criteria which defined those sections of a patient's follow-up time that would fall into these health states were fully specified.

Applications of Q-TWiST tend to use either the same health states as the original breast cancer example, with slightly different definitions or a slightly modified version. Rosenthal and colleagues<sup>19</sup> used the same TOX, TWiST and REL health states in their lung cancer study, as did Gelber and colleagues<sup>21</sup> in their adjuvant study of rectal cancer. Gelber and colleagues<sup>22</sup> used slightly modified states in their Q-TWiST analysis of symptomatic HIV patients. In this study, patients could move from TWiST to a state where first adverse events were experienced and then on to a state of disease progression. Feldstein<sup>12</sup> extended the original breast cancer states, adding

a further 'recovery' state. Gelber and colleagues<sup>17</sup> suggest that in some trials, it may be necessary to define a second period of toxicity in addition to TOX, to represent the late toxic effects of treatments on a patient's quality of life. The Q-TWiST methodology is not always immediately applicable to other diseases, since clear-cut health states may be difficult to define.<sup>23,24</sup> Health states in most Q-TWiST applications are defined using clinical criteria but patient-assessed quality-of-life data has been used.<sup>2</sup>

### The Q-TWiST model

The model used for Q-TWiST is a QALY one (see page 57). In general, weights (which may be utilities) taking values between zero and one, inclusive, are allocated to each health state. These represent the quality values of each health state relative to TWiST, with 0 indicating a state as bad as death and 1 indicating perfect health. Quality-adjusted survival is then defined as the weighted sum of the time spent in each health state. The assumptions for a QALY model (see page 58) are applicable to the Q-TWiST model.

One particular form of the model<sup>16</sup> is represented by

$$\text{Q-TWiST} = u_i \text{TOX} + \text{TWiST} + u_r \text{REL} \quad (25)$$

where *TOX* is the time spent with toxicity resulting from treatment, *REL* is the time spent with symptoms of disease, and  $u_i$  and  $u_r$  are the utilities associated with these periods of survival.

### Discounting in the Q-TWiST model

The Q-TWiST analysis can be adapted so that a greater emphasis is placed on the earlier years. Discounting may be included by apply an appropriate transformation to the times in the data and then using these transformed times in the subsequent partitioned survival analysis.<sup>1</sup>

The transformation used by Goldhirsch and colleagues<sup>16</sup> is such that each time  $t$  is replaced by

$$t_{\text{new}} = 1 - (1 + r)^{-t} \quad (26)$$

where  $r$  is the required discount rate.

### Partitioned survival analysis

It may seem appropriate to calculate Q-TWiST for each individual patient and then use a standard survival analysis on this new quality-of-life oriented endpoint to compare treatments. However, as with TWiST, Q-TWiST is related to the censoring mechanism; patients with poor quality of life are



censored earlier than those with better quality of life. This problem of informative censoring (see chapter 3) will give biased Kaplan–Meier estimates of the survival function and invalidate any standard survival analysis techniques. Partitioned survival analysis, a population-based approach proposed by Glasziou and colleagues,<sup>1</sup> provides a means of handling the problem of informative censoring. Overall survival is partitioned into the time spent in each health state and the mean duration in each state for each group are combined as a weighted sum according to the Q-TWiST model. Weighting the time spent in each health state at the group level, rather than at the individual level, avoids the need to weight censored survival times and thus overcomes the problem of informative censoring.

The date of exiting each successive health state is regarded as an endpoint and Kaplan–Meier estimates are calculated for the time from a fixed origin, such as date of randomisation, to each endpoint. If the exit time from one health state is censored at time  $t$  for a patient, then all subsequent exit times will be censored at time  $t$ . If a state  $H_i$  is skipped, then entering and exiting times are set equal and the exit time for  $H_i$  will be the same as for  $H_{i-1}$ . Kaplan–Meier survival curves corresponding to each transition time can be overlaid on one graph to show the partitioning of overall survival. These are called partitioned survival plots and separate graphs should be produced for each treatment group (see *Figures 21* and *22* relating to the MIC data).

For survival time  $T$ , the area under a survival curve defined by the survivor function,  $S(t)$ , provides an estimate, of the mean survival time,  $E(T)$ , given by

$$E(T) = \int_{t=0}^{\infty} S(t) dt \quad (27)$$

The areas under the survival curves for successive endpoints can therefore be estimated and used to compute the areas between the curves giving estimates of the mean duration of each health state.<sup>26</sup>

If censored data exist, the entire survival curve cannot be estimated; thus these areas can only be calculated if a specified time from randomisation is chosen as the upper time limit for the analysis, that is, the upper limit of the integral will be this time limit rather than infinity. This is usually the upper limit of observation or is based on the follow-up time of the study cohort; it should be chosen so as to reduce or, ideally, eliminate censoring. The median follow-up time, as calculated from the

censoring distribution, is often used as the time limit but will only be appropriate if it adequately reduces the amount of censoring.

Mean times from randomisation to exiting each health state, restricted to the upper time limit, are calculated from the area beneath each estimated survivor function from zero to the chosen finite limit. In practice, the area under a survivor function is estimated by summing the rectangular areas under the Kaplan–Meier curve using the following formula:

$$\text{Area} = \sum_{i=0}^L \hat{S}(t_i) (t_{i+1} - t_i) \quad (28)$$

where each  $t_i$  (from  $i = 1$  to  $i = L - 1$ ) is a death time, with  $t_0$  defined to be zero and  $t_L$  defined to be the chosen upper time limit. The mean survival time given by the SAS Institute in their LIFETEST procedure<sup>27</sup> gives the area under the survival curve.

Differences between successive restricted means for time from randomisation to exiting each health state give the restricted mean duration in each state. The restricted mean quality-adjusted survival is estimated by combining the restricted mean durations as a weighted sum according to the Q-TWiST model. Restricted means based on the product limit method are asymptotically unbiased and normally distributed.<sup>28</sup> Consequently, statistical inferences for quality-adjusted survival can be based on the asymptotic normality of the estimates and require the calculation of standard errors of the estimates. The variance for quality-adjusted survival can be estimated from the vector of utility weights and the variance–covariance matrix for the mean times in each state.<sup>1</sup>

Glasziou and colleagues<sup>1</sup> found no simple expression for the covariance terms when dealing with restricted means and, hence, estimated the variance–covariance matrix using the bootstrap method.<sup>4,5</sup> This means creating a new sample of patients,  $N$ , by repeatedly sampling with replacement from the  $N$  individuals in the trial. This process is repeated thousands of times to obtain a whole series of new data sets. Restricted means for times spent in each state are calculated for each data set to produce an empirical sampling distribution, called a bootstrap sampling distribution, for the statistic. The variances and covariances computed from these values are used as the variance–covariance estimates. Gelber and colleagues<sup>17</sup> used variance–covariance estimates based on a series of 1000 new data sets.

## Comparing treatments in terms of Q-TWiST

In some situations it may be possible to compare Q-TWiST for treatments using specific utility values. These may be chosen arbitrarily if no patient-derived information is available or they could be based on a TTO or standard gamble study, for example. In a study of patients with small-cell lung cancer, Rosenthal and colleagues<sup>19</sup> justified their choice of utility coefficients for their final conclusions ( $u_i = 0.75$  and  $u_r = 0.25$ ) as those which they perceived to most closely resemble the clinical experience of the patients. They compared their Q-TWiST results with those from a study by Goodwin and colleagues,<sup>29</sup> who obtained their utility coefficients ( $u_i = 0.57$  and  $u_r = 0.15$ ) from a proxy group of patients and health professionals.

Other endpoints on which treatments are assessed, such as overall survival, disease-free survival and TWiST, correspond to Q-TWiST with extreme values of  $u_i$  and  $u_r$  (see Table 22).

**TABLE 22** Endpoints corresponding to various utility coefficients in a Q-TWiST analysis

TOX ( $u_i$ )	TWiST	REL ( $u_r$ )	Endpoint
1	1	1	Overall survival
1	1	0	Disease-free survival
0	1	0	TWiST

In most cases, utility weights will be unknown. One approach in this situation is to compare treatments using a spectrum of  $u_i$  and  $u_r$  values. Rosenthal and colleagues<sup>19</sup> examined quality-adjusted survival for a range of utility coefficients and found that the utility coefficient for TOX was far more influential than that for REL.

Another approach is to carry out a threshold utility analysis, a form of sensitivity analysis, in which trial data are used to determine the utility values which would give no difference between treatments, that is, when restricted mean quality-adjusted survival times for both treatments are equal. When there are two unknown utility coefficients, the set of values that give equal quality-adjusted survival is described by a straight line on a two-dimensional plot. CIs can be calculated for this 'threshold line' by finding pairs of utility coefficient values for which the lower and upper bounds of the CI for the treatment effect equals zero.<sup>1,25</sup>

## Q-TWiST analysis of MIC data

The main problem with carrying out a Q-TWiST analysis on the MIC data is defining progressive

health states. The MIC data do not conform to the textbook categories of TOX, TWiST and REL for two reasons. First, the patients in the MIC study have extensive disease and are receiving palliative treatments. They are therefore unlikely to experience time totally free of the symptoms of disease and, thus, do not experience TWiST. This is likely to be a common problem in many quality-of-life studies. Second, the quality-of-life data are only collected during the treatment stage of their survival; hence, during the period of analysis (i.e. 18 weeks from entry to trial) they are constantly in the TOX state.

Since the data did not conform to the standard Q-TWiST health states, alternative progressive health states had to be considered which could be derived from the data and were also clinically meaningful. MAL, expressed as a two-level variable, no malaise (MAL = 0) and malaise (MAL = 1, 2 or 3), was used to define 'good' and 'poor' states of quality of life, respectively. Various sequences of good and poor quality-of-life periods were experienced by patients within the 18 weeks from study entry and these were explored to help define the progressive sequence of health states (see Table 23).

**TABLE 23** Sequences of good (G) and poor (P) quality of life in the MIC data

Sequence	Frequency
G	19
GP	19
GPG	2
GP GP	3
P	36
PG	18
PGP	10
PGPG	2

To incorporate the sequences experienced by **all** patients (i.e. to include both GPGP and PGPG, where G = good and P = poor), the definition of progressive health states needed to include five progressive health states (i.e. GPGPG or PGPGP). This was considered to be beyond both the limit of intelligibility and the limited amount of data. Hence, although a few patients would need to be excluded from the analysis, a definition consisting of four progressive health states was considered preferable.

Having decided on a definition consisting of four progressive health states, there were two possible sequences to consider; PGPG or GPGP. The first option would exclude three patients with a GPGP sequence and the second option would exclude two patients with a PGPG pattern. In general, given several options for a definition, the final choice should be based on what is most clinically meaningful. In this case, neither option seemed clinically preferable to the other and there was no clear clinical explanation for the potential continued fluctuation between the two health states. The decision, therefore, to use the second option, GPGP, was arbitrary, although it did minimise the number of patient exclusions ( $n = 2$ ). To determine if using a different definition affects the conclusions of the analysis a sensitivity analysis should be performed.

The progressive health states were defined as GOOD1, POOR1, GOOD2 and POOR2, where the numbers indicate the first and second visits to the 'same' health state. With this sort of definition of progressive health states, where a patient returns to a health state previously visited, the degree of similarity between the first and second visit to a state needs to be considered. For example, a return to a good quality-of-life state after a patient has been in a poor state (i.e. GOOD2) might describe a very different experience to the first good quality-of-life state (i.e. GOOD1).

Having defined the four progressive health states, the date of exit from each formed successive endpoints for analysis. The time from entry to study to each endpoint was calculated for each patient, with entry and exit times set equal if a state was skipped. For example, if a patient was in a poor state on entry to the study, then they were assumed to have skipped the first state, and their exit time from GOOD1 was set at zero. Also, for example, if a patient exited POOR1 at 18 weeks and did not therefore experience the other two states within the 18-week analysis period, then the exit times from GOOD2 and POOR2 were both set at 18 weeks.

Kaplan–Meier survival curves for successive endpoints were calculated and overlaid to give a partitioned survival plot for each treatment group (see Figures 21 and 22). The area under each curve, with an upper time limit of 18 weeks, was obtained from SAS/STAT output<sup>27</sup> (corresponding to equation 28) and gave the restricted mean survival times from date of entry to trial to each endpoint (see Table 24). The differences between successive means gave the mean time spent in each health state (see Table 25).

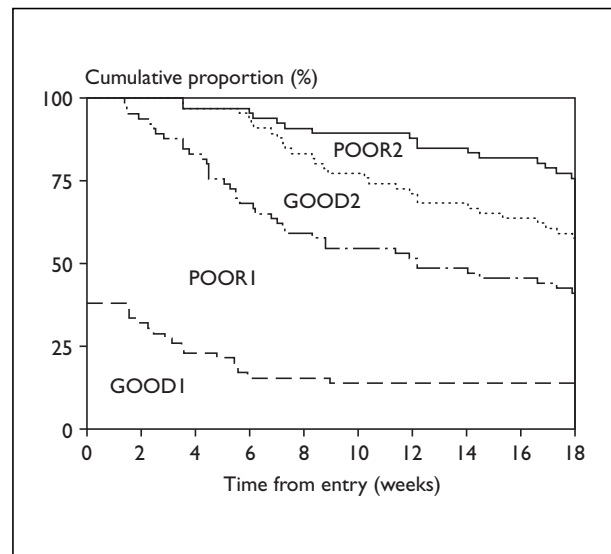


FIGURE 21 Partitioned survival analysis of the MIC data: CT arm

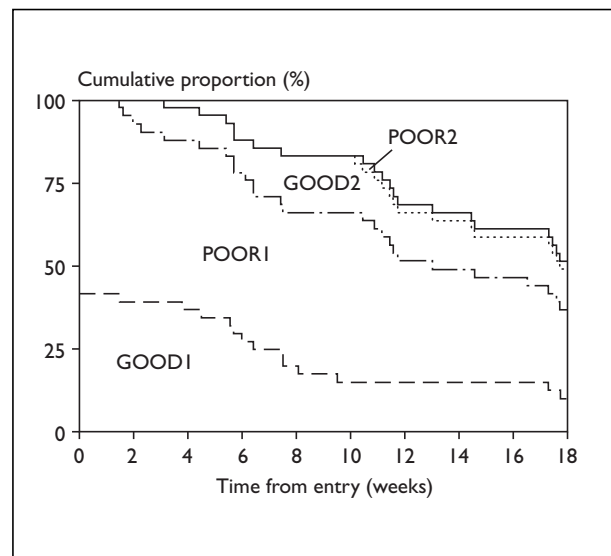


FIGURE 22 Partitioned survival analysis of the MIC data: PAL arm

TABLE 24 Restricted mean survival times from study entry to exiting each health state

Health state	CT arm: restricted mean survival time (weeks)	PAL arm: restricted mean survival time (weeks)
GOOD1	3.37	4.22
POOR1	11.47	12.18
GOOD2	14.63	14.33
POOR2	16.29	14.52

**TABLE 25** Restricted mean survival times spent in each health state

Health state	CT arm: restricted mean survival time (weeks)	PAL arm: restricted mean survival time (weeks)
GOOD1	3.37	4.22
POOR1	8.10	7.96
GOOD2	3.16	2.15
POOR2	1.66	0.19

The Q-TWiST model for the MIC data was defined as follows:

$$Q-TWiST = t_{GOOD1} + u_1 t_{POOR1} + t_{GOOD2} + u_2 t_{POOR2} \quad (29)$$

where  $t_i$  represents time spent in state  $i$  and the utility values,  $u_1$  and  $u_2$  ( $0 \leq u_1 \leq 1$ ,  $0 \leq u_2 \leq 1$ ) are unknown and reflect the reduction in quality of life during the first and second visits, respectively, to the poor quality-of-life state. This model assumes that the quality of life in the good state, whether at first or second visit, was equivalent to perfect health. It also assumes that the quality of life experienced in the poor state differed depending on whether it was the first or second visit.

Utility values are not known and so a threshold utility analysis was undertaken (see Figure 23). Substituting the mean times spent in each state (see Table 25) into equation 29 gives the following equations for the two treatment groups:

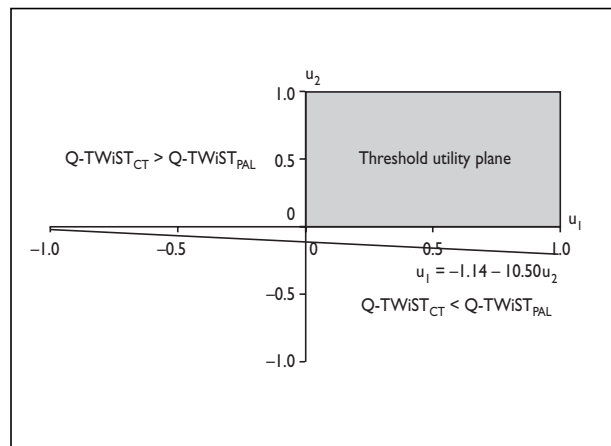
$$Q-TWiST_{CT} = 3.37 + 8.10u_1 + 3.16 + 1.66u_2 \quad (30)$$

$$Q-TWiST_{PAL} = 4.22 + 7.96u_1 + 2.15 + 0.19u_2 \quad (31)$$

Equating equations 30 and 31 gives the threshold line (see Figure 23), indicating pairs of utility coefficients for which the treatments have equal Q-TWiST. The threshold line is given by

$$u_1 = -1.14 - 10.50u_2 \quad (32)$$

The threshold utility plane, i.e. the plane containing all possible pairs of values for  $u_1$  and  $u_2$  ( $0 \leq u_1 \leq 1$  and  $0 \leq u_2 \leq 1$ ), lies completely above this line in the region where  $Q-TWiST_{CT} > Q-TWiST_{PAL}$  (see Figure 23). Thus, whatever values are chosen for  $u_1$  and  $u_2$ , Q-TWiST within 18 weeks of entry to the study is always greater on the chemotherapy arm than on the palliative arm of the trial. A confidence region for the threshold line could not be calculated because



**FIGURE 23** Threshold utility analysis for the MIC data

bootstrap estimates of the standard error for Q-TWiST were not readily available.

### Investigating the effect of imposing an upper time limit

Using an upper time limit for the analysis allows for the possibility that further follow-up could alter the conclusions of the analysis. There is a need to examine the potential effects of having further follow-up time, in terms of changing the conclusions based on the restricted means. Glasziou and colleagues<sup>1</sup> suggest two possible analyses for this purpose.

One method is to examine the difference in Q-TWiST for a range of truncation times. Using a reasonable set of values for the utility weights, Q-TWiST for each treatment can be calculated for a range of upper time limits from 0 to  $L$ , the chosen upper time limit. The difference between treatments can then be plotted against the restriction time. This analysis can be enhanced further by plotting not only the treatment difference for a particular set of utility coefficients as a curve but also the whole range of possible treatment effects for all possible utility coefficients between 0 and 1 as a shaded region. This is called the Q-TWiST gain function.<sup>25</sup> Based on clinical knowledge, a subjective judgement can be made as to whether the graph is sensible and has reached a level of stability.

The second method suggested<sup>1</sup> is to fit a parametric survival model either to the entire distribution or just to the tail of the distribution. In this way, quality-adjusted survival can be extrapolated beyond the time limit  $L$  and the conclusions regarding which treatment is superior can be examined. Gelber and colleagues<sup>30</sup> discuss projecting survival estimates beyond

the follow-up period of a trial by fitting parametric survival models to the tails of Kaplan–Meier survival curves and using the estimated models to make projections. They suggest that Weibull and log-normal models are likely to be the most useful. For the method to be successful in practice, there must be sufficient data in the tail of the survival curve for adequate estimation of a parametric model.

### Extending Q-TWiST to incorporate covariates

In the standard Q-TWiST method, covariates are included by stratifying the sample of patients by the covariate values and performing a separate analysis on each stratum.<sup>1</sup> The problem with this approach is that the sample sizes in some strata may not be sufficient for estimating the mean quality-adjusted survival with reasonable accuracy. In order to avoid stratification and thus enable the entire sample of patients to be used, Cole and colleagues<sup>31</sup> extended the Q-TWiST methodology to allow for covariates using a Cox proportional hazards model.<sup>32</sup> Accelerated failure time regression could be used in a similar manner.<sup>25</sup>

The survivor functions, for time from study entry to exiting each successive health state, are estimated using the results from Cox proportional hazards models fitted to each endpoint, instead of the Kaplan–Meier method. When comparing two treatments, the treatment can be included either as a covariate in a Cox regression or, if the proportional hazards assumption is not valid, a stratified regression can be used. The proportional hazards assumption can be checked for each health state endpoint by plotting  $\log(-\log S(t))$  against  $\log t$  for the different levels of a covariate. If the proportional hazard assumption is valid, the curves will be parallel.

The use of Cox regression models allows quality-adjusted survival, for specific sets of covariate values, to be estimated for each treatment, allowing treatment comparisons in specific subgroups of patients. In a study of zidovudine therapy in asymptomatic HIV-infected patients, Lenderking and colleagues<sup>33</sup> used proportional hazards models to investigate differences between treatments in terms of quality-adjusted survival, for different baseline values of CD4+ cell counts. In a breast cancer application, Gelber and colleagues<sup>34</sup> used covariate values to define two patient profiles corresponding to a good and a poor prognosis, and used Cox models to compare treatments in terms of quality-adjusted survival in both types of patient. A further application in melanoma enabled the effect of the initial

stage of disease on quality-adjusted survival to be investigated.<sup>35</sup>

### Extending Q-TWiST to perform meta-analysis

Meta-analysis is an analytical technique used to summarise quantitatively the results from a number of different studies all addressing the same basic research question.<sup>36</sup> Cole and colleagues<sup>37</sup> developed a meta-analysis methodology for combining the Q-TWiST results from individual trials in which patient-level data is not required. The method is a modified version of the standard Q-TWiST analysis and is summarised by Gelber and colleagues.<sup>25</sup> Regression models are used to combine trials, enabling an overall comparison of treatments in terms of quality-adjusted survival that accounts for the differing follow-up times.

The method was applied to data from eight clinical trials comparing adjuvant chemotherapy with no adjuvant systemic therapy in premenopausal women with breast cancer.<sup>37,38</sup> Gelber and colleagues<sup>39</sup> also used the technique to combine the results from nine trials in order to compare the effect of adjuvant chemotherapy plus tamoxifen with tamoxifen alone in postmenopausal node-positive breast cancer.

### Critical appraisal of quality-adjusted survival analysis

Quality-adjusted survival analysis is based on the concept of combining quality-of-life and survival data in a QALY model, with TWiST and Q-TWiST being special forms of QALYs. The assumptions and criticisms of the QALY model (see pages 58 and 61) also therefore apply to quality-adjusted survival analysis and these assumptions have been discussed in relation to a Q-TWiST model.<sup>40,41</sup> A parametric approach to quality-adjusted survival analysis, as discussed later in chapter 18, allows a more general form for the function combining quality-of-life and survival data and may overcome some of the assumptions of the QALY model.

The concept of creating a quality-adjusted survival time for each patient in a study to be used as an endpoint in treatment comparison instead of survival time seems sensible. However, subject-based approaches in general suffer from the problem of informative censoring. Thus, population-based approaches to quality-adjusted survival analysis, which overcome this problem, are preferable in that respect.

Partitioned survival analysis is a population-based approach to quality-adjusted survival analysis, usually used with a Q-TWiST model but applicable to any QALY model. It has the advantage of being able to include covariates and requires few assumptions,<sup>1</sup> being based on either the non-parametric methods of Kaplan and Meier or the semi-parametric method of Cox regression. It has two main limitations:<sup>1</sup>

- the need to use progressive health states
- the need to restrict the period of analysis to an upper time limit.

Further, it has the disadvantage of not being readily accessible to researchers, since calculation of CIs for quality-adjusted survival uses the bootstrap method.

Partitioned survival analysis is sometimes difficult to apply because of the need for progressive health states. It may be possible to overcome this by specifying different phases of the same state, as in the MIC study, but this can become clumsy and may lose clinical meaning. An alternative method of analysis, multistate survival analysis, discussed in the next chapter, does not require progressive states and is thus a more flexible way of modelling the data.

There is no unique way of dividing the survival time of patients into periods of differing quality of life and the accuracy depends on the frequency of quality-of-life assessments. Different divisions should be considered as part of a sensitivity analysis. Missing quality-of-life data will also cause difficulties since, although quality-adjusted survival analysis deals with informative drop-out caused by death, it does not deal with other reasons for drop-out. Values for the missing data can be imputed or it may be possible to incorporate into the model the time spent as a drop-out, with an appropriate weighting factor to reflect quality of life. Alternatively, methods that explicitly model the drop-out process, such as multistate survival analysis (see chapter 17) or simultaneous modelling (see chapter 19) should be considered.

A further limitation of partitioned survival analysis is the need to restrict the period of analysis to an upper time limit. In some situations, as with the MIC study, the quality-of-life data may only be collected for a limited time, and the period of analysis will automatically be restricted to an upper time limit. Otherwise, methods to investigate the effect of imposing an upper time limit should be

applied. The parametric approach to quality-adjusted survival analysis, as discussed in chapter 18, overcomes this limitation.

Quality-adjusted survival is calculated using a weighted sum of time spent in each health state. The choice of values for the weights is based on the perceived quality of life experienced in each health state and can be problematic. Consideration of a range of possible values should form part of a sensitivity analysis and, in particular, threshold utility analysis allows all possible combinations of values to be investigated.

## References

1. Glasziou PP, Simes RJ, Gelber RD. Quality adjusted survival analysis. *Stat Med* 1990;**9**:1259–76.
2. Beacon HJ. The statistical analysis of self assessed quality of life data in cancer clinical trials [PhD thesis]. London School of Hygiene and Tropical Medicine, University of London; 1996.
3. Hwang JS, Tsauo JY, Wang JD. Estimation of expected quality adjusted survival by cross-sectional survey. *Stat Med* 1996;**15**:93–102.
4. Hinkley DV. Bootstrap methods. *J R Stat Soc B* 1988;**50**:321–37.
5. Efron B. Censored data and the bootstrap. *J Am Stat Assoc* 1981;**76**:312–19.
6. Ganiats TG, Miller CJ, Kaplan RM. Comparing the quality-adjusted life-year output of 2 treatment arms in a randomized trial. *Med Care* 1995;**33**:AS245–54.
7. Gelber RD, Gelman RS, Goldhirsch A. A quality-of-life-oriented endpoint for comparing therapies. *Biometrics* 1989;**45**:781–95.
8. Gelber RD, Goldhirsch A. A new endpoint for the assessment of adjuvant therapy in postmenopausal women with operable breast cancer. *J Clin Oncol* 1986;**4**:1772–9.
9. Gelber RD, Goldhirsch A, Castiglione M, Price K, Isley M, Coates A. Time without symptoms and toxicity (TWiST): a quality-of-life-oriented endpoint to evaluate adjuvant therapy. In: Salmon SE, editor. Adjuvant therapy of cancer V. Orlando: Grune and Stratton; 1987; 455–65.
10. Willemse PHB, van Lith J, Mulder NH, Aalders JG, Bouma J, de Vries EGE, *et al.* Risks and benefits of cisplatin in ovarian cancer. A quality-adjusted survival analysis. *Eur J Cancer* 1990;**26**:345–52.

11. Willemse PHB, de Vries EGE, Kloppenburg M, Fontein DL, Aalders JG, Boonstra H, *et al.* Carboplatin with cyclophosphamide in patients with advanced ovarian cancer: an efficacy and quality-adjusted survival analysis. *Int J Gynecol Cancer* 1992;**2**:236–43.
12. Feldstein ML. Quality-of-life-adjusted survival for comparing cancer treatments: a commentary on TWiST and Q-TWiST. *Cancer* 1991;**67**:851–4.
13. Glimelius B, Hoffman K, Graf W, Haglund U, Nyren O, Pahlman L, *et al.* Cost-effectiveness of palliative chemotherapy in advanced gastrointestinal cancer. *Ann Oncol* 1995;**6**:267–74.
14. Allen-Mersh TG, Earlam S, Fordy C, Abrams K, Houghton J. Quality of life and survival with continuous hepatic-artery floxuridine infusion for colorectal liver metastases. *Lancet* 1994;**344**:1255–60.
15. Brunner KW. Adjuvant therapies in breast cancer and quality of life: a critical review of the TWiST concept. In: Senn HJ, Goldhirsch A, Gelber RD, Osterwalder B, editors. *Adjuvant therapy of primary breast cancer*. Berlin: Springer-Verlag; 1989; 239–43.
16. Goldhirsch A, Gelber RD, Simes RJ, Glasziou P, Coates AS. Costs and benefits of adjuvant therapy in breast cancer: a quality-adjusted survival analysis. *J Clin Oncol* 1989;**7**:36–44.
17. Gelber RD, Goldhirsch A, Cavalli F. Quality-of-life-adjusted evaluation of adjuvant therapies for operable breast cancer. *Ann Intern Med* 1991;**114**:621–8.
18. Gelber RD, Goldhirsch A. Methods for assessing treatment efficacy in trials for adjuvant therapy for breast cancer. In: Senn HJ, Goldhirsch A, Gelber RD, Osterwalder B, editors. *Adjuvant therapy of primary breast cancer*. Berlin: Springer-Verlag; 1989; 211–19.
19. Rosenthal MA, Webster PJ, Gebiski VJ, Stuartharris RC, Langlands AO, Boyages J. The cost of treating small cell lung cancer. *Med J Aust* 1992;**156**:605–10.
20. Cole BF, Solal-Celigny P, Le Page E. Interferon alpha for the treatment of advanced follicular lymphoma: an analysis of quality-of-life-adjusted survival. *Blood* 1995;**86**:1745.
21. Gelber RD, Goldhirsch A, Cole BF, Weiland HS, Schroeder G, Krook JE. A quality-adjusted time without symptoms or toxicity (Q-TWiST) analysis of adjuvant radiation therapy and chemotherapy for resectable rectal cancer. *J Natl Cancer Inst* 1996;**88**:1039–45.
22. Gelber RD, Lenderking WR, Cotton DJ, Cole BF, Fischl MA, Goldhirsch A, *et al.* Quality-of-life evaluation in a clinical trial of zidovudine therapy in patients with mildly symptomatic HIV infection. *Ann Intern Med* 1992;**116**:961–6.
23. Schwartz CE, Cole BF, Gelber RD. Measuring patient-centered outcomes in neurologic disease: extending the Q-TWiST method. *Arch Neurol* 1995;**52**:754–62.
24. Schwartz CE, Cole BF, Vickrey BG, Gelber RD. The Q-TWiST approach to assessing health-related quality of life in epilepsy. *Qual Life Res* 1995;**4**:135–41.
25. Gelber RD, Cole BF, Gelber S, Goldhirsch A. Comparing treatments using quality-adjusted survival: the Q-TWiST method. *Am Statistician* 1995;**49**:161–9.
26. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;**53**:457–81.
27. SAS Institute Inc. SAS/STAT users guide, v.6, vol. 2. 4th ed. Cary, NC: SAS Institute Inc; 1989.
28. Breslow N, Crowley J. A large sample study of the life table and product limit estimates under random censorship. *Ann Stat* 1974;**2**:437–53.
29. Goodwin PJ, Field R, Evans WK, Pater J. Cost-effectiveness of cancer chemotherapy: an economic evaluation of a randomised trial in small-cell lung cancer. *J Clin Oncol* 1988;**6**:1537–47.
30. Gelber RD, Goldhirsch A, Cole BF. Parametric extrapolation of survival estimates with applications to quality of life evaluation of treatments. *Control Clin Trials* 1993;**14**:485–99.
31. Cole BF, Gelber RD, Goldhirsch A. Cox regression models for quality adjusted survival analysis. *Stat Med* 1993;**12**:975–87.
32. Cox DR. Regression models and life tables. *J R Stat Soc B* 1972;**34**:187–220.
33. Lenderking WR, Gelber RD, Cotton DJ, Cole BF, Goldhirsch A, Volberding PA, *et al.* Evaluation of the quality of life associated with zidovudine treatment in asymptomatic human immunodeficiency virus infection. *N Engl J Med* 1994;**330**:738–43.
34. Gelber RD, Goldhirsch A, Cole BF. Evaluation of effectiveness: Q-TWiST. *Cancer Treat Rev* 1993;**19**:73–84.
35. Cole BF, Gelber RD, Kirkwood JM, Goldhirsch A, Barylak E, Borden E. Quality-of-life-adjusted survival analysis of interferon alfa-2b adjuvant treatment of high-risk resected cutaneous melanoma: an Eastern Cooperative Oncology Group study. *J Clin Oncol* 1996;**14**:2666–73.
36. Sutton AJ, Abrams KR, Jones DJ, Sheldon TA, Song F. Systematic reviews of trials and other studies. *Health Technol Assess* 1998;**2**(19).
37. Cole BF, Gelber RD, Goldhirsch A. A quality-adjusted survival meta-analysis of adjuvant chemotherapy for premenopausal breast cancer. *Stat Med* 1995;**14**:1771–84.

38. Gelber RD, Cole BF, Goldhirsch A. Adjuvant chemotherapy for premenopausal breast cancer: a meta-analysis using quality-adjusted survival. *Cancer J Sci Am* 1995;1:114–21.
39. Gelber RD, Cole BF, Goldhirsch A, Rose C, Fisher B, Osborne CK, *et al.* Adjuvant chemotherapy plus tamoxifen compared with tamoxifen alone for postmenopausal breast cancer: meta-analysis of quality-adjusted survival. *Lancet* 1996;347:1066–71.
40. Till JE, de Haes JCJM. Quality-adjusted survival analysis. *J Clin Oncol* 1991;9:525–6.
41. Gelber RD, Goldhirsch A. Quality-adjusted survival analysis [reply]. *J Clin Oncol* 1991;9:526.



# Chapter 17

## Multistate survival analysis

### Introduction

Multistate models have been advocated in a number of reviews and discussions regarding the analysis of quality of life data as a possible means of analysing quality of life and survival data simultaneously.<sup>1-6</sup> They were first proposed in a medical context by Fix and Neyman in 1951.<sup>7</sup> The approach is discussed in some detail by Olschewski and Schumacher<sup>8</sup> in their paper on the analysis of quality-of-life data in cancer clinical trials.

The statistical background to multistate survival analysis is derived from the analysis of event-history data<sup>9</sup> and stochastic processes.<sup>10,11</sup> Many of the theoretical aspects of multistate models fall within a counting process framework.<sup>12-14</sup> The study of events occurring in individuals over time generates event-history data. In studies such as these, individuals can be thought of as occupying one of a finite number of states at any point in time and the movement between states can be described by conditional probabilities or transition rates. This dynamic process is known as a stochastic process. Quality-of-life assessment in clinical trials generates event-history type data, with events being defined as entry and exit from pre-defined health states. The movement of individuals between quality-of-life states can then be considered as a stochastic process and modelled accordingly.

As with quality-adjusted survival analysis, multistate survival analysis starts by defining a finite number of health states, including death, that patients experience during the study. Defining these health states and the possible transitions between them describes the multistate model. The transition rates, which describe the movement between health states, can then be modelled, possibly using covariates. In this way, the time-dependent structure and dynamic nature of quality-of-life data can be incorporated into an analysis comparing treatments and the effect of explanatory variables on transition rates from one state to another can be investigated. Multistate models overcome some of the limitations of the Q-TWiST approach, such as the requirement for progressive health states.

Multistate models in survival analysis have been applied in a variety of clinical settings such as diabetes,<sup>15,16</sup> liver transplantation,<sup>17</sup> bone marrow transplantation,<sup>18</sup> heart transplantation,<sup>19</sup> breast cancer,<sup>20</sup> prostate cancer<sup>21</sup> and HIV.<sup>22</sup> These applications, however, used clinical criteria rather than quality-of-life data to define health states. There are limited examples of applications of multistate models to quality-of-life data.<sup>23,24</sup> In this study, the multistate survival analysis methodology is illustrated by application to quality-of-life data collected in the MIC study.

### Defining the model

#### Health states

The set of health states chosen for the model should be clinically meaningful and fully describe the experiences of the patients. The states should be mutually exclusive and exhaustive. The number of states should be restricted so that the model does not become overcomplicated and also to ensure that the number of patients passing from one state to another is sufficient for adequate modelling of the data.

There are two main types of health states. A transient state is one that a patient can pass through during the course of their follow-up and an absorbing state is one that a patient cannot leave once it has been entered. The standard model for survival analysis corresponds to the simplest multistate model, in which the patient can be in one of two possible states, a transient 'alive' state or an absorbing 'death' state. The competing risks model<sup>25</sup> is an extension of this two-state survival model and forms a multistate model, with one transient alive state and several absorbing death states corresponding to different causes of death. In terms of modelling quality-of-life data, it is more relevant to extend the simple two-state survival model so that there are several transient alive states and one absorbing death state.

The simplest version of this sort of multistate model is the **three-state disability model** (also called the **illness-death model**), where there are two transient alive states: 'alive without disability'

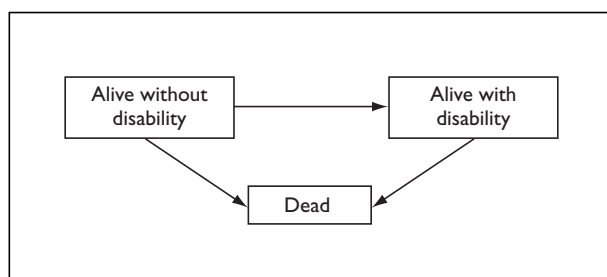


FIGURE 24 A three-state disability model

and ‘alive with disability’, and one absorbing death state. Andersen<sup>15</sup> modelled diabetes data using a model with this structure (see Figure 24).

### Quality-of-life states for the MIC study

The MIC quality of life data were modelled using a three-state illness–death model (see Figure 25). The patients’ health states were determined from the malaise question, which asked if they had been feeling generally ill. Patients were categorised as being in an ‘alive and well’ state if the patient had no malaise (MAL = 0) and were categorised as ‘alive and ill’ if the patient had malaise at any level (MAL = 1, 2 or 3).

The health states used in the model could have been defined using different cut-offs for responses to the malaise question or could have been based on other questions asked in the study or on a cut-off value for the MQS. A more complex model with more than three states was considered but such a formulation would have meant that there were too few transitions for adequate modelling of the data.

### Health state transitions

The movement between states can be described using either a transition probability or a transition rate. A transition probability is the likelihood of an individual moving from one state to another within a specified period; a transition rate is the instantaneous potential of transition at any point in time.<sup>26</sup> While the transition probability can take values between 0 and 1, the transition rate, sometimes called a transition intensity, has no upper bound. In the simple two-state survival model, the transition rate from a transient alive state to an absorbing death state is the standard hazard rate function for the survival time distribution (see page 41).

The two measures are related<sup>26</sup> and, when the instantaneous transition rate  $r$  remains constant during the period,  $t_1$  to  $t_2$ , a transition probability  $p$  can be estimated using

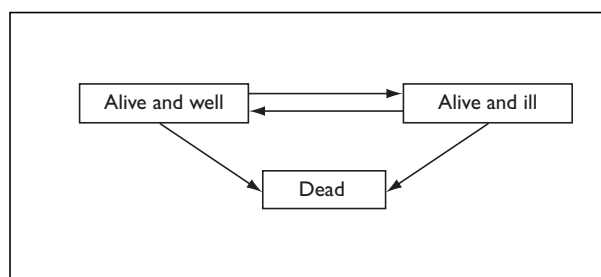


FIGURE 25 Multistate model for the MIC study

$$p = 1 - \exp(-r(t_2 - t_1)) \quad (33)$$

In some situations, patients in the trial may experience ‘recovery’ during the follow-up period and in these circumstances the model may need to include reverse transitions, allowing the ability to return to a state previously occupied. Andersen and colleagues<sup>27</sup> model data from a liver cirrhosis trial as a three-state illness–death model but they allow patients with low prothrombin index to recover and return to the ‘alive with normal prothrombin index’ state, thus incorporating reverse transitions in their model. If reverse transitions are possible then repeat transitions may occur, that is, a patient may experience a particular transition more than once during follow-up. In some circumstances, it may be preferable to model repeat and initial transitions as separate events.<sup>28</sup> This would only be possible if the number of patients experiencing repeat transitions was large enough for adequate modelling.

In the MIC study, patients moved between health states, often several times, until finally moving to the absorbing state of death. The model, therefore, included reverse and, hence, repeat transitions (shown in Figure 25 by arrows in both directions between the alive states).

### Assumptions of the underlying stochastic process

In the most general model, the transition rates would depend on the whole history of the patient. However, it is often reasonable to assume that at any time point the state currently occupied by a patient contains all the information relevant to that patient’s future course. Under this assumption, the model represents a Markov process. If the transition rates from each state are conditional on the duration of time spent in the state, often called the **sojourn time**, then the model represents a semi-Markov process.<sup>9</sup>

It may be necessary to define the health states of the model so that the Markov assumption is

valid. For example, in the MIC study, if the transition of patients between the alive-well and death states depended on whether the patient had originally been in the alive-ill state, then it would be necessary to create two alive-well states: an 'always been alive and well' state and a 'was ill but now alive and well' state. In this way the validity of the Markov assumption could be retained.

If the transition rates remain constant over time (i.e. are independent of time) then the process is time-homogeneous; otherwise, if they vary over time (i.e. are functions of time), the process is time-inhomogeneous. In some situations the transition rates may be 'piecewise' constant, that is, constant over defined subdivisions of the follow-up time.

## Modelling the transition rates

In a standard survival model, the transition rate from the alive state to the dead state is commonly represented by a Cox regression model.<sup>29</sup> The application of Cox regression models to the more general multistate framework, which allows several transient disease states between entry to study and death, has been discussed.<sup>19,30-33</sup> The exact dates of transition from one state to another, or estimates of them, are needed for this type of analysis. Cox regression models have been used to model the transition rates in various multistate survival analysis applications.<sup>15,17,18,20,28</sup>

The transition rate from state  $i$  to state  $j$ ,  $\lambda_{ij}(t)$  can be modelled using a Cox regression model as follows:

$$\lambda_{ij}(t) = \lambda_{0ij}(t) \exp(\beta_{ij}^T \mathbf{x}_{ij}) \quad (34)$$

where  $\lambda_{0ij}(t)$  is a baseline transition rate for the transition from  $i$  to  $j$ ,  $\mathbf{x}_{ij}$  is a vector of covariates specific to that transition and  $\beta_{ij}$  is a vector of unknown regression coefficients specific to that transition. The model could be generalised to include time-dependent covariates by replacing  $\mathbf{x}_{ij}$  with  $\mathbf{x}_{ij}(t)$ .

### Markov or semi-Markov

When modelling the transition rates, consideration needs to be given to whether the process is Markov or semi-Markov. Semi-Markov processes have been discussed in relation to multistate models.<sup>30-32,34</sup> In a Markov process, as discussed above, the transition rate to another state depends only on the present state occupied, while in a semi-Markov process the transition rate is also dependent upon the duration of time spent in the present state.<sup>13</sup>

The type of process is related to the timescale on which time in the model for a transition rate is measured, that is, it relates to the time at which the clock starts. If, in modelling the transition rate from state  $i$  to state  $j$ , time ( $t$ ) is measured from the time of entry to the study (i.e.  $t = 0$ ), then the model represents a Markov process, since duration of time in state  $i$  is not included. Otherwise, if time is measured in the model from time of entry to state  $i$ , say  $w$  in relation to study entry time, then time is included in the model as  $t - w$  rather than  $t$  and, with sojourn time now included, the model represents a semi-Markov process.<sup>20</sup> In a semi-Markov process, the clock is effectively 'reset to zero' every time a state is entered.<sup>9</sup>

Information regarding the history of the process prior to entering state  $i$  may be included as covariates. In particular, when individuals can experience the same transition more than once, the transition rate from state  $i$  to state  $j$  may depend on such aspects as whether state  $i$  has been occupied before, the number of times state  $i$  has been visited before and the total time previously spent in state  $i$ . If the covariates in the model do not include information regarding states prior to the current one, then the model implicitly assumes that the changes of state form a Markov process.<sup>19</sup> The time from study entry to entry to state  $i$  could also be included as a covariate.<sup>20</sup> When the time origin is taken as date of entry to study, then the covariates containing information on history before entering state  $i$  will be time-dependent, while if the origin is taken as date of entry to state  $i$ , then the history is already determined at the origin and the information will be included as fixed covariates.

### Semi-parametric or fully-parametric

The transition rates can either be modelled semi-parametrically using a Cox regression model where the underlying baseline transition rate is left unspecified<sup>19</sup> or parametrically by assuming the transition times follow a specific distribution, thus giving a parametric form to the baseline transition rate.<sup>35,36</sup> The most commonly used distributions are the exponential and Weibull distribution, the exponential being just a special form of Weibull distribution (see page 41). If an exponential distribution is assumed for the transition times from state  $i$  to state  $j$ , then the underlying baseline transition rate is constant:

$$\lambda_{0ij}(t) = \lambda_{ij} \quad (35)$$

and the model is, thus, time-homogeneous, while for a Weibull distribution the underlying baseline transition rate takes the form

$$\lambda_{0ij}(t) = \lambda_{ij} \gamma_{ij} t^{\gamma_{ij}-1} \quad (36)$$

and the model is time-inhomogeneous. If an exponential distribution is assumed for the transition times then, since the underlying transition rate is not dependent on time, the models under the assumption of either a Markov or a semi-Markov process will be equivalent. This, however, is not the case for other distributions, such as Weibull, since the baseline transition rate changes over time.

### Estimating parameters

Kay<sup>19</sup> showed that the partial likelihood for a transition from state  $i$  to state  $j$  is identical to the partial likelihood for the standard Cox regression model,<sup>37</sup> with transitions from  $i$  to states other than  $j$  treated as censored data. In situations where individuals may experience the same transition more than once, the partial likelihood is still valid.<sup>19</sup> Thus, for each transition in the model, estimates of the  $\beta$  regression coefficients are obtained by maximising the relevant partial likelihood.

Significance testing and CIs for the  $\beta$  parameters can be based either on the asymptotic normality of the distribution of the estimators or on the large sample likelihood ratio test for nested models,<sup>19</sup> where changes in  $-2 \log(\text{likelihood})$  are compared to a chi-square distribution. As with a standard regression analysis, the selection of covariates can be made using backward elimination, forward elimination or a stepwise approach.

### Model checking

The assumption of proportional hazards for covariates in the model should be checked using standard graphical techniques and the fit of model can be assessed using residual analysis.<sup>38</sup> These standard techniques require independent observations and so, in the case of repeat transitions, assessment of the model is problematic.

### Computing issues

When modelling the transition rate from state  $i$  to state  $j$  as a semi-Markov process, that is, with the time origin taken as date of entry to state  $i$ , the PHREG procedure (SAS Institute)<sup>39</sup> can be used to fit semi-parametric models and the LIFEREG procedure (SAS Institute)<sup>40</sup> can be used to fit parametric models.

In the parametric modelling of transition times  $T_i$ , the SAS software fits an accelerated failure time model, which has the general form

$$\log(T_i) = \mu + \alpha^T \mathbf{x}_i + \sigma \varepsilon_i \quad i = 1, \dots, n \quad (37)$$

where  $\mathbf{x}_i$  is a vector of covariates,  $\alpha$  is a vector of unknown regression coefficients,  $\sigma$  is an unknown scale parameter,  $\mu$  is an unknown intercept parameter, and  $\varepsilon_i$  is an error term where errors are assumed to come from a known distribution. If transition times are assumed to follow a Weibull distribution then there is a direct correspondence between the parameters under an accelerated failure time model ( $\mu$ ,  $\sigma$  and  $\alpha$  in equation 37) and those under a proportional hazards model ( $\lambda$ ,  $\gamma$  and  $\beta$  in equations 34 and 36). The relationships are given by the following formulae:<sup>38</sup>

$$\begin{aligned} \lambda &= \exp(-\mu/\sigma) \\ \gamma &= 1/\sigma \\ \beta &= (-\alpha/\sigma) \end{aligned} \quad (38)$$

When modelling the transition rate from state  $i$  (the base state) to state  $j$  as a Markov process, that is the time origin is taken as date of entry to study, SAS software cannot be used since it will not allow for times of entry to the base state other than zero. Other statistical software, in particular the *agreg* function in S-PLUS (Statistical Sciences),<sup>41</sup> allows entry times to the base state other than zero in semi-parametric models and thus can be used for Markov modelling. Specially written functions in S-PLUS<sup>23</sup> needed to be written to fit Weibull–Markov models.

## Modelling transition rates in the MIC data

### Introduction

A three-state illness–death model was used to describe the MIC quality-of-life and survival data (see *Figure 25*). For reasons of data availability, the analysis was restricted to the 18-week period from study entry (see page 14). Before any modelling of the data could be undertaken, various assumptions relating to the quality-of-life data needed to be made.

Quality of life was assessed at distinct time points and assumptions were necessary to infer values over continuous time. These assumptions were the same as those made to calculate QALYs (see page 58). In particular, changes in quality of life were assumed to occur at the midpoint between assessments. This enabled the exact dates of

transition between the two alive states to be estimated. Dates of death for patients in the study were known and so exact dates for transitions to death were available. Sensitivity analysis should be used to assess the impact of the assumptions on the conclusions of the analysis.

Drop-outs in the data were dealt with by imputing values using a 'last value carried forward' approach. As well as other methods of imputation (see Table 7), other strategies for dealing with drop-outs in the context of multistate survival analysis were considered. Patients who dropped out of the quality-of-life study because of death were not a problem since their transition to death was included in the model. Transition times for patients who dropped out for reasons other than death could be censored at date of drop-out, with date of drop-out possibly defined as the next planned assessment date, although this is likely to result in informative censoring. Another option considered was to include 'drop-out' as a state in the model but this made the model too complex for the limited amount of data available.

### Models fitted to the data

During the 18-week period from study entry, 113 transitions in total were experienced and these data are used to model the four transition rates (see Table 26). The transition rates were modelled using a Cox regression model with only a covariate for treatment ( $trt$ ) included. In all cases the comparison is chemotherapy ( $trt = 1$ ) against standard palliative treatment ( $trt = 0$ ). Although treatment was the only covariate in the model, Kaplan–Meier survival estimates and log-rank tests could not be used for treatment comparison since the MIC data includes repeat transitions and therefore observations would not be independent.

For illustrative purposes a variety of models were fitted, all adaptations of the basic Cox regression model according to whether the process was assumed to be Markov or semi-Markov, and whether a parametric form for the underlying transition rate was specified (see Table 27).

The transition rates were analysed one at a time. When modelling the transition rate from state  $i$  to state  $j$ , only individuals who occupied state  $i$  at some point during the analysis period contributed to the model, although individuals who passed out of state  $i$  at some point and then returned back to it later contributed twice to the model. If an individual was in state  $i$ , then they were 'at risk' for the state  $i$  to state  $j$  transition. If they moved to state  $k$  rather than state  $j$ , then the time for the  $i$ – $j$  transition was censored at the time of passing to state  $k$ . If they did not move from state  $i$  before the 18-week time limit, the  $i$ – $j$  transition time was censored at 18 weeks.

The set-up of the data for analysis was different depending on whether the model being fitted was Markov (Models 1 and 4) or semi-Markov (Models 2 and 5). The data set-up for the time-homogeneous model (Model 3) was the same as that for a semi-Markov model.

- **Markov model:** The data for the transition from state  $i$  to state  $j$  consisted of one line per patient per visit to state  $i$ . Each line consisted of the time of entry to and exit from state  $i$  in relation to trial time and an indicator variable which was 1 for an exit to state  $j$ , representing an actual event, and 0 for either an exit to state  $k$  or an exit time censored at 18 weeks.
- **Semi-Markov model:** The data for the transition from state  $i$  to state  $j$  consisted of one line per patient per visit to state  $i$ . Each line consisted of the duration of time spent in state  $i$  on that visit and an indicator variable which was 1 for an exit to state  $j$ , representing an actual event, and 0 for either an exit to state  $k$  or an exit time censored at 18 weeks.

S-PLUS was used to fit Markov models, with the `agreg` function<sup>41</sup> used for semi-parametric models (Model 1) and specially written functions<sup>23</sup> used for Weibull models (Model 4). SAS/STAT software was used to fit semi-Markov models, with the PHREG procedure<sup>39</sup> used for semi-parametric

TABLE 26 Frequency of each transition in the MIC data

Transition	Number at risk	Total	CT arm	PAL arm
Alive and well → alive and ill	79	39	28	11
Alive and ill → alive and well	106	37	27	10
Alive and well → dead	79	10	3	7
Alive and ill → dead	106	27	13	14

**TABLE 27** Models fitted to transition rates in the MIC study

Model	Distribution for transition times	Process	Model for transition rate from state $i$ to state $j$
1	Unspecified	Markov	$\lambda_{ij}(t trt) = \lambda_{0ij}(t) \exp(\beta_{ij}trt)Y_{ij}(t)$ where: $\lambda_{0ij}(t)$ is baseline transition rate $\beta_{ij}$ is the regression parameter for treatment $Y_{ij}(t)$ is the 'at risk' process – if an individual is in state $i$ at time $t$ and therefore at risk for transition to state $j$ then $Y_{ij}(t) = 1$ , otherwise $Y_{ij}(t) = 0$ $t$ represents time from study entry
2	Unspecified	Semi-Markov	$\lambda_{ij}(t trt, t_i) = \lambda_{0ij}(t - t_i) \exp(\beta_{ij}trt)$ where: $\lambda_{0ij}(t - t_i)$ is the baseline transition rate $\beta_{ij}$ is the regression parameter for treatment $t_i$ represents time from study entry to entry to state $i$
3	Exponential, i.e. transition rates assumed to be constant over time	Not applicable (Markov equivalent to semi-Markov)	$\lambda_{ij}(t trt) = \lambda_{ij} \exp(\beta_{ij}trt)$ where: $\lambda_{ij}$ is the constant baseline transition rate $\beta_{ij}$ is the regression parameter for treatment $t$ can be either time from study entry or time from entry to state $i$
4	Weibull, i.e. transition rates assumed to vary over time	Markov	$\lambda_{ij}(t trt) = \lambda_{ij}Y_{ij}t^{Y_{ij}-1} \exp(\beta_{ij}trt)Y_{ij}(t)$ where: $\lambda_{ij}$ and $Y_{ij}$ are the scale and shape parameters for the Weibull distribution $\beta_{ij}$ is the regression parameter for treatment $Y_{ij}(t)$ is the 'at-risk' process described in Model 1 $t$ represents time from study entry
5	Weibull, i.e. transition rates assumed to vary over time	Semi-Markov	$\lambda_{ij}(t trt, t_i) = \lambda_{ij}Y_{ij}(t - t_i)^{Y_{ij}-1} \exp(\beta_{ij}trt)$ where: $\lambda_{ij}$ and $Y_{ij}$ are the scale and shape parameters for the Weibull distribution $\beta_{ij}$ is the regression parameter for treatment $t_i$ represents time from study entry to state $i$

**TABLE 28** Comparison of treatment regression parameters in all five multistate models

Transition	Model 1: semi-parametric Markov	Model 2: semi-parametric semi-Markov	Model 3: exponential Markov/ semi-Markov	Model 4: Weibull Markov	Model 5: Weibull semi- Markov
Well → ill	0.48 (−0.22, 1.18)	0.47 (−0.23, 1.17)	0.46 (−0.24, 1.15)	0.47 (0.10, 0.84)	0.45 (−0.25, 1.15)
Ill → well	0.40 (−0.32, 1.13)	0.31 (−0.41, 1.04)	0.34 (−0.39, 1.07)	0.36 (−0.02, 0.74)	0.34 (−0.39, 1.06)
Well → dead	−1.39 (−2.74, −0.04)	−1.26 (−2.61, 0.09)	−1.33 (−2.68, 0.03)	−1.40 (−2.53, −0.27)	−1.27 (−2.63, 0.08)
Ill → dead	−0.77 (−1.53, −0.01)	−0.71 (−1.46, 0.05)	−0.73 (−1.48, 0.03)	−0.78 (−1.32, −0.23)	−0.70 (−1.45, 0.06)

models (Model 2) and the LIFEREG procedure<sup>40</sup> used for parametric models (Models 3 and 5).

### Results for treatment comparison

The semi-parametric models (Models 1 and 2) give estimates only for the treatment regression parameter, while the parametric models (Models 3, 4 and 5) also give estimates for the parameters

of the underlying distributions for the transition times. The five different models were only fitted for illustrative purposes. The effect of treatment on each transition rate is the main interest and thus the estimates of the treatment regression parameter given by each model are given here (see Table 28). Determination of which model best fits the data was not relevant here, hence,

parameter estimates for the exponential and Weibull distributions have not been given.

The treatment parameter estimates for the five models are reasonably comparable, although there is some discrepancy regarding the width of the CIs, which warrants further investigation. The results for the semi-parametric Markov model (Model 1) suggest that, although the effects were not statistically significant, chemotherapy increased the risk of transition between the two different quality-of-life states. There was also statistically significant evidence to suggest that chemotherapy reduced the risk of transition to death from both the well and ill quality-of-life states. Although the significance of the treatment effects in the equivalent semi-Markov model (Model 2) differed from the Markov model, the point estimates and also, to a great extent, the CIs did not greatly differ. Thus, in this study, taking account of duration of time in a state has relatively little effect on the transition rates out of the state. The significance of the treatment effects for the parametric models (Models 3–5) also differ but, again, there is little to discriminate between the point estimates and CIs, indicating that assuming such underlying distributions for the transition times between states does not greatly influence the transition rates.

## Alternative modelling approaches when exact transition times are not known

### Introduction

In many prospective studies, although the exact time of death is known, it is not possible to observe the actual transitions of patients from one health state to another, all that is known for a patient is the health state occupied at certain follow-up times. The transition times in this situation are interval-censored data and the transition rates cannot be modelled directly using the methods discussed above. One approach, used in the analysis of the MIC data above, is to estimate the exact transition dates from the data. The most widely used approximations to exact transition dates are the actual follow-up dates or the mid-point between follow-up dates. Estimates of transition rates obtained from approximated data may be incorrect.<sup>27</sup> Alternative approaches to modelling transition rates, that account for the interval censored data, are discussed here. The application of these methods is difficult in practice because standard software cannot be used, although special FORTRAN computer programs have been developed.<sup>42–44</sup>

### Models in which time is continuous

In situations when exact transition times are not known, Kay<sup>42</sup> proposed modelling transition rates using a general continuous-time Markov chain. Transition rates are assumed to be constant over time, although piecewise constant transition rates could be accommodated, by splitting time at a number of pre-defined points and estimating separate transition rates for each period. This method of multistate modelling has been applied to the data in a number of studies.<sup>16,22,27,45–47</sup>

Estimates of transition rates are obtained using maximum likelihood methods, where the likelihood function is formulated from the probability of each individual's passage through the health states. The standard relationship between the matrix of transition rates and the matrix of transition probabilities, as described by Kolmogorov equations,<sup>10,11</sup> enables the maximum likelihood estimation of transition rates. The maximisation process is an iterative procedure and computer routines in FORTRAN are available to compute parameter estimates together with their standard errors.<sup>42</sup>

Covariates such as treatment can be included in the modelling process as a proportional factor over the constant baseline transition rates, as in a Cox model,<sup>42</sup> so that the transition rate from state  $i$  to state  $j$  can be represented by

$$\lambda_{ij}(\mathbf{x}) = \lambda_{0ij} \exp(\beta_{ij}^T \mathbf{x}) \quad (39)$$

where  $\lambda_{0ij}$  is the constant baseline transition rate,  $\mathbf{x}$  is the vector of covariates and  $\beta_{ij}$  is the vector of associated regression parameters for that particular transition. A computer program written in FORTRAN called MARKOV<sup>43</sup> has been designed to allow the inclusion of covariates in the model. The number of model parameters may be reduced either by assuming the effect of the covariates is either the same for all progressive and for all regressive transitions or is the same for all transitions.<sup>16</sup>

Hypothesis testing can be carried out using either likelihood ratio tests or Wald tests but opinions regarding which is optimal for various circumstances lack consistency. In the comparison of transition rates (i.e.  $H_0: \lambda_i = \lambda_j$ ), the Wald test, which calculates a  $\chi^2$  test statistic based on estimated transition rates and associated covariances, has been advocated.<sup>42</sup> Gentleman and colleagues,<sup>22</sup> however, have sometimes found the normal approximation to be unreliable and therefore recommend the use of likelihood ratio tests for

such comparisons. Other hypothesis tests of interest might be regarding the comparison of regression parameters for different transition rates (i.e.  $H_0: \beta_{ij} = \beta_{mn}$ ). In this case, Marshall and Jones<sup>16</sup> recommend the Wald test, whereas Kay<sup>42</sup> recommends likelihood ratio tests. Marshall and Jones<sup>16</sup> recommend likelihood ratio tests for covariate selection within the modelling of each transition rate.

Modelling transition times using a general continuous-time Markov chain is only valid when the pattern of follow-up times, called the **examination scheme**, is noninformative for the disease process.<sup>48</sup> This condition is equivalent to noninformative censoring in survival analysis. In situations where examinations take place because of patient self-selection, the examination scheme may be informative and estimates of transition rates may be biased.

Related methods have been described for studies in which individuals move from state to state, according to a continuous-time Markov process, but examination times are fixed. Methods have been proposed for situations in which the data is in aggregate form, that is, it consists of the number of individuals in each state at the fixed time points.<sup>49</sup> Alternatively, if individual patient data is known, a more flexible method that assumes transition rates are proportional overcomes the limitations of Kay's method of constant transition rates.<sup>50</sup>

### Models in which time is a discrete variable

In some studies it may be preferable to view follow-up time as a series of discrete time points, at a given interval apart, rather than a continuous scale. In this way follow-up time is divided into segments called Markov cycles. For example, Myers and colleagues<sup>21</sup> measured time in units of 90 days where time point 0 was at study entry and time point 1 was 90 days later, and so on. They then determined a patient's state at each time point from the last available measurement in the preceding 90-day interval.

Various considerations are relevant when deciding on the length of the time interval. The model assumes that each patient may make only one transition during each cycle; hence the time interval needs to be kept short enough for this assumption to be tenable but needs to be long enough to allow a reasonable number of patients to experience transitions. The tenability of the Markov property will depend on the length of the time

interval employed. For a time-homogeneous Markov process, the length of the cycles must be uniform, while for time-inhomogeneous processes the cycles may vary in length.

### Time-inhomogeneous process

If the Markov process is time-inhomogeneous then each cycle, which may vary in length, will have its own transition probability matrix. Each matrix can be estimated from the data and can be used to estimate, for each fixed time point  $t$  defining the cycles, the probabilities of being in each state.<sup>51</sup>

Suppose the process consists of  $s$  states and  $\mathbf{P}_{t-1,t}$  is the  $(s \times s)$  transition matrix for the time interval  $(t-1, t)$ , then the maximum likelihood estimate for the  $(i, j)$ th element of the matrix is given by

$$\mathbf{P}_{t-1,t}(i, j) = n_{t-1,t}(i, j) / n_{t-1}(i) \quad (40)$$

where  $n_{t-1,t}(i, j)$  is the number of transitions from state  $i$  at time  $t-1$  to state  $j$  at time  $t$ , and  $n_{t-1}(i)$  is the number of individuals in state  $i$  at time  $t-1$  that are not censored in the interval  $(t-1, t)$ .

Given these estimated transition matrices, the vector  $\mathbf{p}_t$  containing the unconditional probabilities of being in each state  $i$  ( $i = 1$  to  $s$ ) at time  $t$  can be estimated using

$$\mathbf{p}_t = \mathbf{p}_0 \prod_{k=1}^t \mathbf{P}_{k-1,k} = \mathbf{p}_{t-1} \mathbf{P}_{t-1,t} \quad (41)$$

where  $\mathbf{p}_0$  is the initial probability vector estimated from the data using the proportion of patients in each state at time 0.

This methodology is analogous to a standard life-table analysis, in that the probabilities are estimated using observed frequencies and the cumulative survival probability for any time  $t$  is the product of all the preceding interval survivals.<sup>51</sup> The method is based on the assumption that incomplete follow-up is not related to patient outcome.

### Time-homogeneous process

If the Markov process is assumed to be time-homogeneous, then the probability of changing from state  $i$  to state  $j$  between times  $t-1$  and  $t$  is equal to that between times  $t$  and  $t+1$ . Based on this assumption, estimates of transition probabilities can be calculated from the data. The probability of transition from state  $i$  to state  $j$  can be estimated by calculating

$$p_{ij} = \sum_t n_{t-1,t}(i, j) / \sum_t n_{t-1}(i) \quad (42)$$



where  $n_{i-1,i}(i, j)$  and  $n_{i-1}(i)$  are as specified above in equation 40.

These can be presented in the form of a transition probability matrix,  $\mathbf{P}$ , which provides a useful summary description of the movement between states.<sup>21</sup> The transition probabilities in the matrix can be modelled using log-linear models<sup>52</sup> or the matrix can be calculated to the  $n$ th power to indicate, for each starting state, the probabilities of being in each state after  $n$  cycles. These can be plotted on a graph over time.<sup>53</sup>

The expected duration in each state can be calculated from the transition probability matrix.<sup>53</sup> A submatrix,  $\mathbf{P}^*$ , of the transition probability matrix can be formed from the probabilities of transition between the non-absorbing states. The matrix  $\mathbf{N}$ , known as the fundamental matrix of an absorbing chain, can then be calculated as

$$\mathbf{N} = (\mathbf{I} - \mathbf{P}^*)^{-1} \quad (43)$$

where  $\mathbf{I}$  is the identity matrix. This matrix gives the expected number of cycles patients would spend in each health state given their initial state. The matrix  $\mathbf{V}$  containing the variances of the expected number of cycles can be calculated as

$$\mathbf{V} = \mathbf{N} (2\mathbf{N}^* - \mathbf{I}) - \mathbf{N}^2 \quad (44)$$

where  $\mathbf{N}^*$  is a copy of the  $\mathbf{N}$  matrix with only the diagonal elements preserved and zeros elsewhere.<sup>54</sup>

## Critical appraisal of multistate survival analysis

Multistate survival analysis can be problematic since the method requires extensive data and is based on quite strong simplifying assumptions.<sup>3</sup> However, as quality of life increases in importance in clinical trials and becomes more routinely collected, the problem of lack of data should be less of an issue.<sup>4</sup> Furthermore, if multistate models are proposed at the design stage of a clinical trial, then collection of data can be planned so that it yields adequate and appropriate data.

The data requirements for multistate modelling are strict, with not only dates of entry, death and censoring needed but also 'exact' dates of transition between health states. Transition dates can be estimated and the accuracy is determined by

the frequency of the quality-of-life assessments. Alternative methods that do not require exact transition dates are available but are not readily accessible to researchers because specialised computer software is required.

Defining health states for a multistate model may be problematic, despite the fact that progressive health states are not required. The definitions based on quality-of-life data are subjective and different definitions need to be considered as part of a sensitivity analysis. The investigator has to make decisions on which quality-of-life variable to use, the number of health states to be included and the cut-off values used to discriminate between health states. At one extreme, the model needs to be complex enough to be clinically meaningful and to ensure that information from the data is utilised to a maximum. At the other extreme, a simple model is needed to allow an adequate number of transitions between health states, both to enable transition rates to be estimated with sufficient precision and to ease interpretation of the analysis.

Definition of the health states may be such that clinically important information is lost. It may not be possible to include the most clinically important transition in the model because of the small number of participants in the study experiencing it. For example, in the MIC study the moderately ill to severely ill transition could not be modelled because of too few patients but this could be a very important transition from a clinical viewpoint. Also the transition between health states may result from very small changes in quality of life while large changes in quality of life are not reflected in the model. For example, if the MQS had been used in the MIC study to define health states and a value of 1 had been used to discriminate between the alive-well and alive-ill states, then a small change in MQS from 0.9 to 1.1 would result in a transition while a large change from, say, 1.1 to 3 would not be reflected in the model.

A multistate model, despite its complexity, may be preferable to an overall survival model since a greater biological insight may be gained by analysing the steps in a disease process.<sup>15</sup> Different covariates may affect different transitions and the effects of important covariates may be lost when considering just overall survival. Examples of applications of multistate models to quality-of-life data are limited. Such applications, including the possibility of the inclusion of drop-out states in the model to overcome problems of informative

drop-out, require further investigation. Methods discussed in this chapter have been based on a classical approach and, although there has been some work on Bayesian approaches to multistate modelling,<sup>23,55</sup> further work in this field is needed.

## References

- Fayers PM, Jones DR. Measuring and analysing quality of life in cancer clinical trials: a review. *Stat Med* 1983;**2**:429–46.
- Schumacher M, Olschewski M, Schulgen G. Assessment of quality-of-life in clinical trials. *Stat Med* 1991;**10**:1915–30.
- Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Spiegelhalter DJ, Jones DR. Quality-of-life assessment: can we keep it simple? *J R Stat Soc A* 1992;**155**:353–93.
- Abrams KR. In discussion of Cox, *et al.* Quality-of-life assessment: can we keep it simple? *J R Stat Soc A* 1992;**155**:353–93.
- Olschewski M, Schulgen G, Schumacher M. In discussion of Cox, *et al.* Quality-of-life assessment: can we keep it simple? *J R Stat S A* 1992;**155**:353–93.
- Abrams KR. In: Papers from the conference on methodological and ethical issues in clinical trials: Discussion. *Stat Med* 1993;**12**:1523–31.
- Fix E, Neyman J. A simple stochastic model of recovery, relapse and loss of patients. *Hum Biol* 1951;**23**:205–41.
- Olschewski M, Schumacher M. Statistical analysis of quality of life data in cancer clinical trials. *Stat Med* 1990;**9**:749–63.
- Clayton D. The analysis of event history data: a review of progress and outstanding problems. *Stat Med* 1988;**7**:819–41.
- Cox DR; Miller HD. The theory of stochastic process. London: Chapman and Hall; 1965.
- Chiang CL. An introduction to stochastic processes and their applications. Huntington, NY: Robert E. Krieger; 1980.
- Andersen PK, Borgan O. Counting process models for life history data: a review. *Scand J Stat* 1985;**12**:97–158.
- Andersen PK, Borgan O, Gill RD, Keiding N. Statistical models based on counting processes. New York: Springer-Verlag; 1993.
- Fleming TR, Harrington DP. Counting processes and survival analysis. New York: John Wiley; 1991.
- Andersen PK. Multistate models in survival analysis: a study of nephropathy and mortality in diabetes. *Stat Med* 1988;**7**:661–70.
- Marshall G, Jones RH. Multistate models and diabetic retinopathy. *Stat Med* 1995;**14**:1975–83.
- Hansen BE, Thorogood J, Hermans J, Ploeg RJ, van Bockel JH, van Houwelingen JC. Multistate modelling of liver transplantation data. *Stat Med* 1994;**13**:2517–29.
- Klein JP, Keiding N, Copelan EA. Plotting summary predictions in multistate survival models: probabilities of relapse and death in remission for bone marrow transplantation patients. *Stat Med* 1993;**12**:2315–32.
- Kay R. The analysis of transition times in multistate stochastic processes using proportional hazard regression models. *Comm Stat – Theory Methodol* 1982;**11**:1743–56.
- Kay R. Multistate survival analysis: an application in breast cancer. *Methods Inf Med* 1984;**23**:157–62.
- Myers LE, Paulson DF, Berry WR, Cox EB, Laszlo J, Stanley W. A time-dependent statistical model which relates current clinical status to prognosis: application to advanced prostatic cancer. *J Chronic Dis* 1980;**33**:491–9.
- Gentleman RC, Lawless JF, Lindsey JC, Yan P. Multistate Markov models for analyzing incomplete disease history data with illustrations for HIV disease. *Stat Med* 1994;**13**:805–21.
- Abrams KR. Bayesian survival analysis [PhD thesis]. Liverpool: University of Liverpool; 1992.
- Olschewski M. Methods of evaluating and analyzing quality of life in clinical trials. *Biometrics* 1984;**40**:264.
- David HA, Moeschberger ML. The theory of competing risks. London: Charles Griffin; 1978.
- Miller DK, Homan SM. Determining transition probabilities: confusion and suggestions. *Med Decis Making* 1994;**14**:52–8.
- Andersen PK, Hansen LS, Keiding N. Assessing the influence of reversible disease indicators on survival. *Stat Med* 1991;**10**:1061–7.
- Islam MA. Multistate survival models for transitions and reverse transitions: an application to contraceptive use data. *J R Stat Soc A* 1994;**157**:441–55.
- Cox DR. Regression models and life tables. *J R Stat Soc B* 1972;**34**:187–220.
- Hsieh FY, Crowley J, Tormey DC. Some test statistics for use in multistate survival analysis. *Biometrika* 1983;**70**:111–9.
- Cox DR. Some remarks on semi-Markov processes in medical statistics. In: Proceedings international symposium on semi-Markov processes and their applications; Brussels, June 1984.
- Prentice RL, Williams BJ, Peterson AV. On the regression analysis of multivariate failure time data. *Biometrika* 1981;**68**:373–9.

33. Islam MA, Singh KP. Multistate survival models for partially censored data. *Environmetrics* 1992;**3**:223–34.
34. Lagakos SW, Sommer CJ, Zelen M. Semi-Markov models for partially censored data. *Biometrika* 1978;**65**:311–17.
35. Beck GJ. Stochastic survival models with competing risks and covariates. *Biometrics* 1979;**35**:427–38.
36. Lagakos SW. A stochastic model for censored survival data in the presence of an auxiliary variable. *Biometrics* 1976;**32**:551–9.
37. Cox DR. Partial likelihood. *Biometrika* 1975;**62**:269–76.
38. Collett D. Modelling survival data in medical research. London: Chapman and Hall; 1994.
39. SAS Institute Inc. SAS/STAT software: changes and enhancements; release 6.07. SAS technical report P-229. Cary, NC: SAS Institute Inc; 1992.
40. SAS Institute Inc. SAS/STAT users guide, v. 6, vol. 2. 4th ed. Cary, NC: SAS Institute Inc; 1989.
41. Statistical Science Inc. S-PLUS user's manual. Seattle: Statistical Sciences; 1991.
42. Kay R. A Markov model for analysing cancer markers and disease states in survival studies. *Biometrics* 1986;**42**:855–65.
43. Marshall G, Guo WS, Jones RH. MARKOV: a computer program for multistate Markov models with covariables. *Comput Methods Programs Biomed* 1995;**47**:147–56.
44. Guo WS, Marshall G. ORDMKV: a computer program fitting proportional odds model for multistate Markov process. *Comput Methods Programs Biomed* 1995;**46**:257–63.
45. Longini IM, Clark WS, Byers RH, Ward JW, Darrow WW, Lemp GF, *et al.* Statistical analysis of the stages of HIV infection using a Markov model. *Stat Med* 1989;**8**:831–43.
46. Gottschau A, Høgh B. Interval censored survival data and multistate compartmental models in the analysis of first appearance of plasmodium falciparum parasites in infants. *Stat Med* 1995;**14**:2727–36.
47. Lu Y, Stitt FW. Using Markov processes to describe the prognosis of HIV-1 infection. *Med Decis Making* 1994;**14**:266–72.
48. Gruger J, Kay R, Schumacher M. The validity of inferences based on incomplete observations in disease state models. *Biometrics* 1991;**47**:595–605.
49. Kalbfleisch JD, Lawless JF, Vollmer WM. Estimation in Markov models from aggregate data. *Biometrics* 1983;**39**:907–19.
50. Thornquist MD. Proportional hazards model for repeated measures with monotonic ordinal response. *Biometrics* 1993;**49**:721–30.
51. Hillis A, Maguire M, Hawkins BS, Newhouse MM. The Markov process as a general method for nonparametric analysis of right-censored medical data. *J Chronic Dis* 1986;**39**:595–604.
52. Leviton A, Schulman J, Kammerman L, Porter D, Slack W, Graham JR. A probability model of headache recurrence. *J Chronic Dis* 1980;**33**:407–12.
53. Silverstein MD, Albert DA, Hadler NM, Ropes MW. Prognosis in SLE: comparisons of Markov model to life table analysis. *J Clin Epidemiol* 1988;**41**:623–33.
54. Beck JR, Pauker SG. The Markov process in medical prognosis. *Med Decis Making* 1983;**3**:419–58.
55. Charlesworth A, Skene AM. A multi-state Markov process model for quality of life and survival in large heart failure clinical trials. *Control Clin Trials* 1997;**18**(3S):165S.



## Chapter 18

# Incorporating results from a multistate model into a Q-TWiST analysis

### Introduction

The quality-adjusted survival analysis discussed previously (see chapter 16) is, in general, a non-parametric technique for comparing health technologies in terms of both quality of life and survival simultaneously. Cole and colleagues<sup>1</sup> devised a parametric approach to quality-adjusted survival analysis that enabled many of the limitations of the Q-TWiST model to be overcome. Essentially the method uses the results from a multistate survival analysis to carry out a more refined Q-TWiST analysis. This chapter is based purely on this paper,<sup>1</sup> which is a seminal one on this topic.

The method was illustrated by Cole and colleagues<sup>1</sup> by data from a breast cancer clinical trial. The health states used in the analysis were the usual Q-TWiST states (i.e. TOX, TWiST and REL), which were determined using clinical criteria rather than quality-of-life data. However, quality-of-life data rather than clinical criteria could be used to define the health states in the model.

### Methodology

As with both Q-TWiST and multistate survival analysis, the analysis starts by defining the health states that patients occupy during the study. Possible transitions between health states are also defined. A multistate survival analysis approach is then used to model the transition intensities. Each successive transition is represented by a cause-specific hazard function which is conditional upon previous transitions. The model is equivalent to a semi-Markov model with hazard functions being conditional on the current state, the transitions to the current state and the duration of the current state. Covariates can be included in the model. Cole and colleagues<sup>1</sup> used conditional log-normal distributions and conditional Weibull distributions to model their hazard functions and included covariates using accelerated failure-time models. Parameter estimates are obtained by maximum likelihood.

Using the models fitted to the health state transitions, the expected amount of time spent in each health state can be estimated using simulation for specific sets of covariate values. Standard errors for these estimates can be obtained using the bootstrap method. Cole and colleagues<sup>1</sup> estimated the average time spent in each health state for two sets of covariate values, one relating to a good prognosis and the other to a poor prognosis. The estimates were restricted to a 10-year follow-up period but, in general, the estimates can be unrestricted.

The next stage is to combine the average time spent in each health state with utility measures representing the quality of life experienced in each health state. Cole and colleagues<sup>1</sup> referred to the function which combines quality and quantity of life into a composite measure as a quality function. The non-parametric Q-TWiST model is a quality function but more general functions can be defined to allow the utility coefficients for a health state to depend on the entry and exit times and the duration of time spent in the state. Quality functions can also be defined to incorporate discounting. Cole and colleagues<sup>1</sup> used the standard non-parametric Q-TWiST quality function in their analysis and compared treatments in terms of quality-adjusted survival using various arbitrary utility values. The significance of the difference between treatments can be established from the CIs for the difference in quality-adjusted survival.

### Conclusions

The parametric approach to quality-adjusted survival analysis overcomes many of the limitations associated with a non-parametric Q-TWiST analysis. It does not restrict estimates to the follow-up period of the data and allows covariate values to be easily included. In theory it does not require progressive health states, although there may be strong assumptions in relation to repeat transitions. The method also allows quality of life and survival to be combined in a more general way than the Q-TWiST model with, for example, utility coefficients being functions

of time rather than just constants and discount rates being included where appropriate.

The method requires the number of observations and the number of states to be such that there are sufficient data for estimating each of the conditional cause-specific hazard functions. If estimation is based on few data, the variability of the estimates will be large giving perhaps excessive variability in the quality-adjusted survival.

In conclusion, when a particular parametric model is appropriate and a reasonably large amount of data is available, this method will give more efficient estimation than a non-parametric approach.

## **Reference**

1. Cole BF, Gelber RD, Anderson KM. Parametric approaches to quality-adjusted survival analysis. *Biometrics* 1994;**50**:621–31.

## Chapter 19

# Modelling quality of life and survival as two simultaneous processes

In studies where quality of life and survival are both important endpoints, the change in quality of life over time and the time to death can be considered as two simultaneous processes occurring in patients, and can be modelled as such. Models for analysing longitudinal quality-of-life data and survival as two separate processes were discussed earlier in this review (see chapters 5–9 and chapters 10–13, respectively). Simultaneous modelling allows survival data to be incorporated into the model for quality-of-life data, thus adjusting for informative missing data caused by drop-out due to death. Conversely, the analysis of survival data with quality of life as a time-dependent covariate is enhanced in simultaneous modelling, since covariate values are estimated from the model for quality-of-life data over time fitted to all subjects.

The simultaneous modelling of repeatedly measured covariates and survival data has been described for covariate data in the form of normally distributed clinical markers<sup>1,2</sup> but not for quality-of-life data, which may be non-normally distributed continuous measures or even ordinal measures. Models are set up for each process and Gibbs sampling is used to simultaneously fit the models in a single analysis.<sup>1,2</sup> Faucett and Thomas<sup>1</sup> refer to the two parts of the model as the ‘covariate tracking model’ for the repeated measures process and the ‘disease risk model’ for the survival process. In their simultaneous modelling they assume a random effects model for the covariate tracking part (see equations 45 and 46 below) and a proportional hazards model for the disease risk part (see equation 47).

Thus,  $z_{ij}$ , the single continuous time-dependent covariate for the  $j$ th measurement of subject  $i$  at time  $t_{ij}$  is modelled using:

$$z_{ij} = x_i(t_{ij}) + \varepsilon_{ij} \quad (45)$$

where  $x_i(t_{ij})$  is the value of the true unobserved covariate at time  $t_{ij}$  and  $\varepsilon_{ij}$  are independent, normally distributed errors, and

$$x_i(t) = \alpha_i + \beta_i t \quad (46)$$

where the random effects  $\alpha_i$  and  $\beta_i$  have a bivariate normal distribution.

At the same time, the hazard of death for subject  $i$  is modelled using

$$\lambda_i(t) = \lambda_0(t) \exp(\gamma x_i(t)) \quad (47)$$

where  $\lambda_0(t)$  represents the underlying baseline hazard function and  $\gamma$  is the regression coefficient for estimation.

The method is flexible in that it allows for unequally spaced and missing repeated measures data, with varying numbers of observations per subject, and it allows for censored survival times. Simulation was used to compare results from modelling each process separately with results from the combined model and showed that the separate models underestimated parameters while the combined model virtually eliminated the bias.<sup>1</sup>

Instead of modelling the survival process simultaneously with a longitudinal measurements process, a more general approach would be to model the drop-out process,<sup>3</sup> which includes other reasons for drop-out apart from death. Lindsey<sup>3</sup> mentions that any appropriate standard longitudinal repeated measures model can be used to analyse, conditionally, all observed responses up to drop-out, and his paper is focussed instead on modelling the drop-out process. A survival model in a log-linear form is used to model the drop-out process, with subjects not dropping out being uninformatively censored. The longitudinal measurements can be included as covariates in the drop-out model. If the study is such that drop-outs occur for different reasons, modelling them as distinct risk processes should be considered. Lindsey<sup>3</sup> considers that the drop-out process should be modelled together with the longitudinal measurements process, since it is an integral part of the phenomenon under study.

## References

1. Faucett CL, Thomas DC. Simultaneously modelling censored survival data and repeatedly measured covariates: a Gibbs sampling approach. *Stat Med* 1996;15:1663–85.

2. Berzuini C. Medical monitoring. In: Gilks WR, Richardson S, Spiegelhalter DJ, editors. Markov chain Monte Carlo in practice. London: Chapman and Hall; 1995; 321–37.
3. Lindsey JK. Modelling longitudinal measurement dropouts as a survival process with time-varying covariates. Limburgs Universitair Centrum, Belgium: Technical Report; 1997.



## Chapter 20

# Simultaneous analysis of quality-of-life and survival data: summary and discussion

In studies in which quality of life and survival are both important endpoints, it may be advantageous to assess health technologies in terms of these endpoints simultaneously. Three different approaches can be used to achieve this. One approach is to combine quality and quantity of life into a single endpoint and use quality-adjusted survival analysis to compare treatments. Another approach is to use multistate models to describe the movement of patients between various health states, defined by levels of quality of life and death, and to explore how treatments differ in terms of these movements. Finally, a simultaneous modelling approach considers quality of life and survival as two simultaneous processes and describes the data in terms of two inter-linked models.

There is some controversy regarding the simultaneous analysis of quality-of-life and survival data. One view held is that the time-dependent structure

of the individual quality-of-life process can best be accounted for when quality and quantity of survival are analysed simultaneously,<sup>1</sup> while another is that methods which attempt to combine quality of life and survival into a single measure are generally inappropriate.<sup>2,3</sup> It may be preferable to analyse and report quality-of-life and survival outcomes separately so that any conflict between them in terms of treatment differences is apparent.<sup>4</sup>

If simultaneous analysis is deemed to be appropriate, the choice of method depends on the aims of the study, the nature of the disease and treatments, and the quality-of-life data collected. The strengths and limitations of each method should be considered carefully (see *Table 29*). Quality-adjusted survival analysis is the most straightforward approach but may be difficult to apply because of the need for progressive health states. Multistate modelling is more flexible but more complex and has stringent data requirements. In

**TABLE 29** Comparison of quality-adjusted survival analysis and multistate survival analysis

Quality-adjusted survival analysis	Multistate survival analysis
<b>Aim</b>	
<ul style="list-style-type: none"> <li>Compares treatments in terms of a composite measure of quality and quantity of life</li> </ul>	<ul style="list-style-type: none"> <li>Compares treatments in terms of movements between different health states</li> </ul>
<b>Limitations</b>	
<ul style="list-style-type: none"> <li>Need to be able to define a finite, mutually exclusive and exhaustive set of health states that are clinically meaningful and fully describe the experiences of patients</li> <li>For partitioned survival analysis, the health states must be progressive</li> <li>QoL assessments need to be frequent enough to be able to estimate transition times adequately</li> <li>The assumptions underlying the QALY model may not be valid</li> <li>For a subject-based approach, a cut-off time for the analysis needs to be chosen to minimise the number of censored survival times, otherwise a population-based approach should be used</li> <li>Partitioned survival analysis requires the period of analysis to be restricted to an upper time limit</li> <li>Calculation of CIs in a partitioned survival analysis requires specialised methods</li> </ul>	<ul style="list-style-type: none"> <li>Need to be able to define a finite, mutually exclusive and exhaustive set of health states that are clinically meaningful and fully describe the experiences of patients</li> <li>QoL assessments need to be frequent enough to be able to estimate exact transition times adequately, otherwise specialised software needs to be available for the use of alternative methods</li> <li>Sample size needs to be large enough to ensure the numbers of patients passing from one state to another is sufficient for adequate modelling of the data</li> </ul>
<b>Strengths</b>	
<ul style="list-style-type: none"> <li>Enables covariates to be included</li> </ul>	<ul style="list-style-type: none"> <li>Enables different covariates to be considered for different transitions</li> </ul>

both cases, defining health states can be problematic and different definitions should be considered as part of a sensitivity analysis. The accuracy of the results from both methods depends on the frequency of the quality-of-life assessments. Simultaneous modelling could potentially provide the most flexible and powerful approach but further research of the method is required before it can be fully evaluated.

Methods that simultaneously analyse quality-of-life and survival data are based on a single measure of quality of life. This may cause problems in situations with several distinct quality-of-life measures, when the analysis would need to be repeated for each, and large numbers of quality-of-life endpoints could lead to problems of interpretation and multiple testing. All methods, however, not only have the advantage of allowing for informative drop-out due to death but also have the potential to be extended to deal with informative drop-out for reasons other than death. Further research

into the application of these methods to quality-of-life data is required.

## References

1. Olschewski M, Schumacher M. Statistical analysis of quality of life data in cancer clinical trials. *Stat Med* 1990;**9**:749–63.
2. Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Spiegelhalter DJ, Jones DR. Quality-of-life assessment: can we keep it simple? *J R Stat Soc A* 1992;**155**:353–93.
3. Hopwood P, Stephens RJ, Machin D. Approaches to the analysis of quality of life data: experiences gained from a Medical Research Council Lung Cancer Working Party palliative chemotherapy trial. *Qual Life Res* 1994;**3**:339–52.
4. Fletcher A, Gore SM, Jones DR, Fitzpatrick R, Spiegelhalter DJ, Cox DR. Quality of life measures in health care. II: design, analysis and interpretation. *BMJ* 1992;**305**:1145–8.

## Chapter 2 I

# Conclusions and recommendations

### Overall summary and discussion

In health technology assessment with both quality of life and survival as endpoints, the appropriate method of analysis depends on which endpoint is most important. In some situations, the primary focus of the analysis will be quality of life, with consideration also given to survival but only in terms of the problematical impact the latter may have on the quality-of-life data. In other situations, the focus will be primarily on survival with the analysis adjusting for the effects of quality of life and, in further scenarios, both endpoints will be important.

Quality-of-life data are usually longitudinal in nature and, as with most longitudinal studies, the data suffer from the problem of missing values, especially as a result of patient drop-out. The missing values do not only cause the data to be unbalanced but also cause problems for the analysis because the missing data mechanism is likely to be non-ignorable. This is because, in quality-of-life studies in which survival is also an issue, patients are generally severely ill and drop-out caused by illness and death is a common occurrence. This drop-out process may be informative and needs to be accounted for in any analysis of the data.

Another problem with quality-of-life data is in the multivariate nature of the endpoint. Quality-of-life data may comprise a single global measure but often consist of a set of item or dimension scores. The analysis may require each quality-of-life endpoint to be considered separately but, if the number of endpoints is large, multiple testing and estimation may be a problem. Methods that handle the multivariate nature of the quality-of-life endpoint have been proposed but they either suffer from the disadvantage of providing results purely in terms of hypothesis testing and do not give an estimate of treatment effect, or they do not deal with the problem of informative drop-out. The analysis should either focus on a few key quality-of-life measures or should attempt to combine measures into an aggregate score on which the analysis can then be based.

Any analysis of longitudinal quality-of-life data should begin descriptively, using plots of individual

patient profiles and group profiles, to give insight into the data before any formal testing or modelling is carried out. Group profiles should be interpreted with caution since, if informative drop-out is present, group summary measures may well be biased. A descriptive analysis can be used to explore the drop-out problem in the data, although establishing if drop-out is informative is not possible.

The application of standard methods of longitudinal data analysis to quality-of-life data is generally problematic because of the likely presence of informative drop-out. The simple approach of using summary measures does not fully capture the dynamic nature of the data and may be problematic in the presence of informative drop-out; although the more complex modelling techniques model the change in quality of life over time, they assume the missing data mechanism is ignorable. These methods could therefore give biased results and invalid conclusions when informative drop-out is present in the data. Modelling techniques that deal with informative drop-out have been developed and their application to quality-of-life data needs investigation.

If the treatment comparison is in terms of survival, then the analysis can adjust for the effects of quality of life by including it as a covariate in a standard survival model. This also enables the prognostic value of quality of life to be explored. Quality of life can be included either in terms of baseline values as a fixed covariate or in terms of changing values over time as a time-dependent covariate. If assessments of quality of life are infrequent or data are missing for reasons other than death, then it may be difficult to adjust for changing quality of life with any degree of accuracy. The analysis may be improved by modelling quality of life and survival as two simultaneous processes. In this approach, values of quality of life incorporated into a survival analysis as a covariate are estimated from the model for quality-of-life data over time fitted to all subjects.

In studies where quality of life and survival are both important endpoints, it may be

advantageous to assess health technologies in terms of these endpoints simultaneously. Three different approaches can be used to achieve this. One approach is to combine quality and quantity of life into a single endpoint and use quality-adjusted survival analysis to compare treatments; another is to use multistate models to describe the movement of patients between various health states, defined by levels of quality of life and death, and to explore how treatments differ in terms of these movements (see *Table 29* for a comparison of these two approaches). Finally, in a simultaneous modelling approach, quality of life and survival are considered as two simultaneous processes and the data are described in terms of two inter-linked models.

The method of quality-adjusted survival analysis is technically straightforward, except for the calculation of CIs, but may be difficult to apply because it requires survival time to be partitioned into progressive health states. These may not always be easy to define in a clinically relevant way. A further limitation is the need to restrict the period of analysis to an upper time limit, although this can be overcome using a parametric approach. Quality-adjusted survival analysis deals with the problem of informative drop-out due to death and has the potential to be extended to deal with drop-out for reasons other than death.

Multistate survival analysis does not necessarily require progressive health states and provides a more flexible means for modelling quality-of-life and survival data. The inclusion of death as a health state in the model enables the analysis to deal with informative drop-out due to death, and it may be possible to deal with drop-out for other reasons by including a 'drop-out' health state. Depending on the multistate model, methods are reasonably accessible to researchers, except in those situations where exact transition times are not known. In general, multistate survival analysis can be problematic since the method requires extensive data for adequate modelling of the data.

Modelling quality of life and survival as two simultaneous processes has already been suggested as a method that may enhance the analysis of survival data with quality of life as a time-dependent covariate. Further, this approach allows both the interrelationship between the two processes to be assessed and, in particular, for the changes in quality of life over time to be explored while adjusting for informative drop-out due to death. Other reasons for drop-out

can be included by modelling the drop-out rather than the survival process explicitly. The benefits of simultaneous modelling in a quality-of-life context need further investigation.

## Recommendations

### Implications for the design and conduct of HTA research

This review of the methods proposed in the scientific and medical literature for the analysis of quality-of-life and survival data, together with their application to data from a previously conducted study, has given rise to a series of recommendations for practitioners and researchers. These recommendations should complement those given by other groups reviewing quality-of-life assessment in clinical trials.<sup>1,2</sup>

### Obtaining appropriate data

Consideration of the proposed analysis of quality-of-life and survival data at the design stage of a study could result in better quality and more appropriate data for analysis. The method of analysis may influence decisions regarding the frequency and timing of quality-of-life assessments and also the instrument used to measure quality of life. Some methods, such as those using exact dates of health state transitions, may require frequent collection of data to ensure accurate assessment of changes in quality of life. The choice of fixed or varying assessment times may be partly influenced by the proposed analytical method.

In choosing the instrument with which to measure quality of life, consideration should be given to the number of quality-of-life endpoints that will be measured. Some instruments result in a single global measure of quality of life, while others produce a number of dimension-specific scores. In some cases, large numbers of quality-of-life measures will be produced. Most methods of analysis are based on a single measure of quality of life and the analysis will need to be repeated for major quality-of-life endpoints. If a large number of quality-of-life measures are produced then both the analysis and the interpretation may become problematic. Use of a single global measure of quality of life simplifies statistical analysis and should be aimed at when sensible and justifiable.<sup>3</sup> Otherwise, to overcome problems of multiple testing, the most important quality-of-life endpoints on which hypothesis testing will be based should be specified at the design stage, with all other outcomes assumed to be of secondary importance.

Missing data cause the most serious problems for the analysis of quality-of-life data and thus the study design, in terms of quality-of-life assessment, should be chosen to minimise the chance of non-compliance. Patient drop-out is still likely to occur and recording reasons for drop-out during data collection will maximise insight into the drop-out process, thus enabling modelling of separate drop-out processes when necessary. The number of patients required for a study will also depend on the method of analysis. Multistate survival analysis, for example, needs a relatively large overall sample size, so that there are sufficient patients in each transition for adequate modelling of the data.

- The method of analysis needs to be decided upon at the design stage of a study so that appropriate quality-of-life data can be collected. Issues to consider are:
  - the quality-of-life instrument to be used
  - the frequency and timing of quality-of-life assessments
  - the need to minimise non-compliance
  - the collection of additional information, such as reason for drop-out
  - the sample size required.

### Choosing the appropriate method

There are several factors affecting the choice of method for analysing quality-of-life and survival data. The research question that the study aims to answer is a major factor and, if the choice is being made after data collection rather than at the design stage, the nature of the quality-of-life data that has been collected will also have an effect. The analysis needs to produce results that are relevant and accessible to health service professionals and patients. Some methods of analysis are complex and interpretation of results may be difficult.

In broad terms, it may be argued that, in view of the problems of data definition and collection, too much sophistication in analysis would be misguided and the methods used for analysis should broadly match the quality of the data.<sup>4</sup>

In general, when choosing the appropriate method of analysis the advantages and disadvantages of each method should be weighed against each other, with consideration given to the quality of the data.

- The choice of method should be based on the research question that the study aims to answer. The advantages and disadvantages of each method should be considered carefully, together with the relevance and interpretability of the results to clinicians and patients.

Analysis of longitudinal quality-of-life data must consider the problem of informative drop-out. Standard methods for longitudinal data analysis assume missing data mechanisms are ignorable and should be avoided because they may produce biased results and invalid conclusions. Methods that deal with informative drop-out should be used. The appropriate method will depend on whether drop-out is primarily caused by death or occurs for multiple reasons.

- Methods used to analyse longitudinal quality-of-life data must allow for informative drop-out.

### Reporting the analysis

The method used for analysis of quality-of-life and survival data should be described clearly when reporting the results. Some justification for the choice could also be included. To analyse the data, various assumptions usually need to be made, whatever method is used. Assumptions such as these should be reported clearly. In the MIC study, for example, most methods of analysis were based on the major assumption that changes in quality of life occurred midway between assessments rather than on the actual assessment date. Sensitivity analysis should be carried out to assess the effect of the assumptions on the conclusions of the analysis.

- Methods used should be reported clearly, with details of definitions and assumptions used in the analysis.
- Sensitivity analysis should be carried out to assess the robustness of conclusions to any critical assumptions made in the analysis.

### Recommendations for further research

Some of the methods described in this report have, to date, been applied to only a few examples of quality-of-life and survival data. Although used fairly extensively, the application of quality-adjusted survival analysis has been based mainly on clinical data rather than on patient-assessed quality-of-life data; hence, further work evaluating the use of such a technique in a quality-of-life context needs to be addressed. There has been limited use of other techniques, such as hierarchical modelling, multistate modelling and simultaneous modelling, in the quality-of-life field and further research is needed. Most methods are based on a classical approach and the application of Bayesian methods requires further investigation. Although the available computer software has not been reviewed in this project, it has become apparent through worked examples that development of software may be required; a review of this would form a useful supplement to this review.

The number of applications of the methodology discussed in this review to quality-of-life data should continue to expand; the methodological areas most likely to develop rapidly are multistate survival analysis and simultaneous modelling. The problems of analysing longitudinal data with informative drop-outs are also being addressed in other parallel areas, such as AIDS research; developments within this context should also be monitored.

- Further experience in the application of quality-adjusted survival analysis techniques to quality-of-life data is needed to enable a proper evaluation of such methods.
- Further research is needed to develop hierarchical models, multistate models and simultaneous modelling methods in their practical application to quality-of-life and survival data using both classical and Bayesian approaches. Consideration should be given as to how such methods could deal with the multivariate nature of the quality-of-life endpoint.
- A full review of available computer software for methods that simultaneously analyse quality-of-

life and survival data is needed to highlight areas requiring further development.

- Progress in the most rapidly developing areas, multistate survival analysis and simultaneous modelling, should be monitored, together with parallel areas of methodological development, such as in AIDS research.

## References

1. Nayfield SG, Ganz PA, Moinpour CM, Cella DF, Hailey BJ. Report from a National Cancer Institute (USA) workshop on quality of life assessment in cancer clinical trials. *Qual Life Res* 1992;1:203–10.
2. Hopwood P, Stephens RJ, Machin D. Approaches to the analysis of quality of life data: experiences gained from a Medical Research Council Lung Cancer Working Party palliative chemotherapy trial. *Qual Life Res* 1994;3:339–52.
3. Olschewski M, Schumacher M. Statistical analysis of quality of life data in cancer clinical trials. *Stat Med* 1990;9:749–63.
4. Fayers PM, Jones DR. Measuring and analysing quality of life in cancer clinical trials: a review. *Stat Med* 1983;2:429–46.



## Acknowledgements

We would like to thank the HTA programme for their continuing financial support for this project.

We would like to thank the Cancer Research Campaign Trials Unit at Birmingham University for allowing Cindy Billingham to be seconded to Leicester University for the duration of the project. Particular thanks go to Janet Dunn who kindly instigated and negotiated the arrangement.

We are indebted to Dr Michael Cullen, Consultant and Honorary Reader in Medical Oncology at the Queen Elizabeth Hospital in Birmingham, for

allowing us to use the MIC data as the illustrative example in this report. We also gratefully acknowledge the assistance of Sue Spriggs from the Clinical Sciences Library at Leicester University and Andrew Booth from the Sheffield Centre for Health and Related Research (SCHARR) for their advice on searching the literature.

We would also like to thank all our colleagues at Leicester and Birmingham for their invaluable support and encouragement.

Our thanks are also due to the referees for their perseverance in reading the report and the quality of their comments.







## Bibliography

All 1127 references found by the search strategy discussed in chapter 2 are listed here, including those which are not referenced in the report.

The following publications were found too late for inclusion in the study but are likely to be of particular interest.

Staquet MJ, Hays RD, Fayers PM, editors, 1998. Quality of life assessment in clinical trials: methods and practice. New York: Oxford University Press.

Bernhard J, Gelber RD, editors, 1998. Workshop on missing data in quality of life research in cancer clinical trials: practical and methodological issues. *Stat Med*;17:5–7.

### Report bibliography

Anon., 1985. Validity and reliability studies of the diabetes control and complications trials quality of life assessment. *Control Clin Trials*;6:223.

Anon., 1993. The Oregon health care proposal and the Americans with disabilities act. *Harvard Law Rev*;106:1296–313.

Aalen OO, Husebye E, 1991. Statistical analysis of repeated events forming renewal processes. *Stat Med*;10:1227–40.

Aaronson NK, 1988. The assessment of subjective response in prostatic cancer clinical research. *Am J Clin Oncol: Cancer Clin Trials*;11:S43–7.

Aaronson NK, 1989. Quality of life assessment in clinical trials: methodologic issues. *Control Clin Trials*;10:S195–208.

Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, 1993. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst*;85:365–76.

Abrams KR, 1992. In discussion of Cox *et al.*: Quality-of-life assessment: can we keep it simple? *J R Stat Soc A*; 155:353–93.

Abrams KR, 1992. Bayesian survival analysis [PhD thesis]. Liverpool: University of Liverpool.

Abrams KR, 1993. In: Papers from the conference on methodological and ethical issues in clinical trials: Discussion. *Stat Med*;12:1523–31.

Adams PC, Gregor JC, Kertesz AE, Valberg LS, 1995. Screening blood donors for hereditary hemochromatosis: decision analysis model based on a 30-year database. *Gastroenterology*;109:177–88.

Adams PC, Kertesz AE, Valberg LS, 1995. Screening for hemochromatosis in children of homozygotes: prevalence and cost-effectiveness. *Hepatology*;22:1720–7.

Agresti A, 1989. A survey of models for repeated ordered categorical response data. *Stat Med*;8:1209–24.

Agresti A, 1990. Analysing repeated categorical response data. In: Agresti A. Categorical data analysis. New York: Wiley and Sons.

Agresti A, 1993. Distribution-free fitting of logit models with random effects for repeated categorical responses. *Stat Med*;12:1969–87.

Allen D, Chant ADB, 1992. QALYfied arterial reconstruction. *Lancet*;339:1117.

Allen D, Lee RH, Lowson K, 1989. The use of QALYs in health service planning. *Int J Health Planning Manag*;4:261–73.

Allen DR, Fergus MEB, Chant ADB, 1990. Cost-effectiveness of treating critical ischemia using quality-adjusted life years. *Br J Surg*;77:A348.

Allen-Mersh TG, Earham S, Fordy C, Abrams K, Houghton J, 1994. Quality of life and survival with continuous hepatic-artery floxuridine infusion for colorectal liver metastases. *Lancet*;344:1255–60.

Allison PD, 1984. Event history analysis – regression for longitudinal event data. London: Sage Publications.

Anda R, Williamson D, Jones D, Macera C, Eaker E, Glassman A, Marks J, 1993. Depressed affect, hopelessness, and the risk of ischemic heart disease in a cohort of US adults. *Epidemiology*;4:285–94.

Andersen PK, 1986. Time-dependent covariates and Markov processes. In: Moolgavkar S, Prentice RL, editors. Modern statistical methods in chronic disease epidemiology. New York: Wiley; 82–103.

Andersen PK, 1988. Multistate models in survival analysis: a study of nephropathy and mortality in diabetes. *Stat Med*;7:661–70.

Andersen PK, Borgan O, 1985. Counting process models for life history data: a review. *Scand J Stat*;12:97–158.

Andersen PK, Borgan O, Gill RD, Keiding N, 1993. Statistical models based on counting processes. New York: Springer-Verlag.

- Andersen PK, Gill RD, 1982. Cox's regression model for counting processes: a large sample study. *Ann Stat*;10:1100–20.
- Andersen PK, Hansen LS, Keiding N, 1991. Assessing the influence of reversible disease indicators on survival. *Stat Med*;10:1061–7.
- Anderson JP, Bush JW, Berry CC, 1988. Internal consistency analysis: a method for studying the accuracy of function assessment for health outcome and quality of life evaluation. *J Clin Epidemiol*;41:127–37.
- Anderson JP, Kaplan RM, Schneiderman LJ, 1994. Effects of offering advance directives on quality-adjusted life expectancy and psychological well-being among ill adults. *J Clin Epidemiol*;47:761–72.
- Andrykowski MA, Curran SL, Studts JL, Cunningham L, Carpenter JS, McGrath PC, *et al.*, 1996. Psychosocial adjustment and quality-of-life in women with breast cancer and benign breast problems: a controlled comparison. *J Clin Epidemiol*;49:827–34.
- Anie KA, Jones PW, Hilton SR, Anderson HR, 1996. A computer-assisted telephone interview technique for assessment of asthma morbidity and drug use in adult asthma. *J Clin Epidemiol*;49:653–6.
- Antczak A, Weinstein MC, 1987. An outcome measure for dentistry: quality-adjusted tooth-years. *J Dent Res*;66:299.
- Ashby J, O'Hanlon M, Buxton MJ, 1994. The time trade-off technique: how do the valuations of breast cancer patients compare to those of other groups? *Qual Life Res*;3:257–65.
- Ashby M, Neuhaus JM, Hauck WW, Bacchetti P, Heilbron DC, Jewell NP, *et al.*, 1992. An annotated bibliography of methods for analysing correlated categorical data. *Stat Med*;11:67–99.
- Aubuchon JP, Birkmeyer JD, 1994. Safety and cost-effectiveness of solvent-detergent-treated plasma. In search of a zero-risk blood supply. *JAMA*;272:1210–14.
- Badia X, Alonso J, 1995. Re-scaling the Spanish version of the Sickness Impact Profile: an opportunity for the assessment of cross-cultural equivalence. *J Clin Epidemiol*;48:949–57.
- Badia X, Alonso J, 1996. Validity and reproducibility of the Spanish version of the Sickness Impact Profile. *J Clin Epidemiol*;49:359–65.
- Baillie SR, 1995. Uses of ringing data for the conservation and management of bird populations: a ringing scheme perspective. *J Appl Stat*;22:967–87.
- Bakker C, van der Linden S, 1995. Health-related utility measurement: an introduction. *J Rheumatol*;22:1197–9.
- Bakker CH, Rutten-van Molken M, van Doorslaer E, Bennett K, van der Linden S, 1993. Health-related utility measurement in rheumatology: an introduction. *Patient Educ Counsel*;20:145–52.
- Baladi JF, Menon D, Otten N, 1996. An economic evaluation of finasteride for treatment of benign prostatic hyperplasia. *Pharmacoeconomics*;9:443–54.
- Barendregt JJ, Bonneux L, van der Maas PJ, 1994. Health expectancy: an indicator for change. *J Epidemiol Community Health*;48:482–7.
- Barnett DB, 1991. Assessment of quality-of-life. *Am J Cardiol*;67:C41–4.
- Barosi G, 1993. Recombinant erythropoietin in the treatment of chronic renal failure: an economic analysis. *G Ital Farm Clin*;7:157–61.
- Barr RD, Feeny D, Furlong W, Weitzman S, Torrance GW, 1995. A preference-based approach to health-related quality-of-life for children with cancer. *Int J Pediatr Hematol/Oncol*;2:305–15.
- Barrett BJ, Parfrey PS, Foley RN, Detsky AS, 1994. An economic analysis of strategies for the use of contrast media for diagnostic cardiac catheterization. *Med Decis Making*;14:325–35.
- Barrett-Connor E, 1994. Postmenopausal estrogen and the risk of breast cancer. *Ann Epidemiol*;4:177–80.
- Barry MJ, Mulley AG Jr, Fowler FJ, Wennberg JW, 1988. Watchful waiting vs immediate transurethral resection for symptomatic prostatism. The importance of patient's preferences. *JAMA*;259:3010–17.
- Bass EB, Pitt HA, Lillemoe KD, 1993. Cost-effectiveness of laparoscopic cholecystectomy versus open cholecystectomy. *Am J Surg*;165:466–71.
- Bass EB, Steinberg EP, Pitt HA, Saba GP, Lillemoe KD, Kafonek DR, *et al.*, 1991. Cost-effectiveness of extracorporeal shock-wave lithotripsy versus cholecystectomy for symptomatic gallstones. *Gastroenterology*;101:189–99.
- Bayliss-Brown P, Glicher SR, Sayers G, Nord E, 1992. An alternative to QALYs: saved young life equivalent (SAVE). *BMJ*;305:1365–6.
- Beacon HJ, 1996. The statistical analysis of self assessed quality of life data in cancer clinical trials [PhD thesis]. London School of Hygiene and Tropical Medicine, University of London.
- Beacon HJ, Thompson S, 1996. Multi-level models for repeated measurement data: application to quality of life data in clinical trials. *Stat Med*;15:2717–32.
- Beard MEJ, Inder AB, Allen JR, Hart DNJ, Heaton DC, Spearing RL, 1991. The costs and benefits of bone-marrow transplantation. *N Z Med J*;104:303–5.
- Beck GJ, 1979. Stochastic survival models with competing risks and covariates. *Biometrics*;35:427–38.
- Beck JR, Kattan MW, Miles BJ, 1994. A critique of the decision analysis for clinically localized prostate cancer. *J Urol*;152:1894–9.
- Beck JR, Pauker SG, 1983. The Markov process in medical prognosis. *Med Decis Making*;3:419–58.
- Beck JR, Pauker SG, Gottlieb JE, Klein K, Kassires JP, 1982. A convenient approximation of life expectancy (the 'DEALE'). II. Use in medical decision-making. *Am J Med*;73:889–97.

- Beck JR, Salem DN, Estes NAM, Pauker SG, 1987. A computer-based Markov decision analysis of the management of symptomatic bifascicular block: the threshold probability for pacing. *J Am Coll Cardiol*;9:920–35.
- Belyea M, Barefoot J, Siegler I, Hooker K, 1988. Religion, well-being, social contacts and survival: a conceptual replication. *Am J Epidemiol*;128:927.
- Benesmalone M, Young TB, Weber S, 1991. The relation of sleep disordered breathing and treatment to quality-of-life. *Am J Epidemiol*;134:788–9.
- Bennett CL, Matchar D, McCrory D, McLeod DG, Crawford ED, Hillner BE, 1996. Cost-effective models for flutamide for prostate carcinoma patients: are they helpful to policy makers? *Cancer*;77:1854–61.
- Bennett K, Torrance G, Tugwell P, 1991. Methodologic challenges in the development of utility measures of health-related quality-of-life in rheumatoid arthritis. *Control Clin Trials*;12:S118–28.
- Berndt ER, Griliches Z, Rappaport NJ, 1995. Econometric estimates of price indexes for personal computers in the 1990s. *J Econometrics*;68:243–68.
- Bernhard J, Gelber RD, editors, 1998. Workshop on missing data in quality of life research in cancer clinical trials: practical and methodological issues. *Stat Med*;17:5–7.
- Berzuini C, 1995. Medical monitoring. In: Gilks WR, Richardson S, Spiegelhalter DJ, editors. Markov chain Monte Carlo in practice. London: Chapman and Hall.
- Best NG, Spiegelhalter DJ, Thomas A, Brayne CEG, 1996. Bayesian analysis of realistically complex models. *J R Stat Soc A*;159:323–42.
- Biem HJ, Detsky AS, Armstrong PW, 1990. Management of asymptomatic chronic aortic regurgitation with left ventricular dysfunction: a decision analysis. *J Gen Intern Med*;5:394–401.
- Billingham LJ, Abrams KR, Jones DR, 1999. Methods for the analysis of quality of life and survival data in health technology assessment. *Health Technol Assess*;3(10).
- Birch S, 1986. QALYs or short straws. *BMJ*;293:338.
- Birch S, Gafni A, 1992. Cost-effectiveness utility analyses: do current decision rules lead us to where we want to be. *J Health Econ*;11:279–96.
- Birch S, Gafni A, 1995. Ethical dimensions of health technology assessment: being economic with economic principles. *Int J Technol Manag*;511:60–72.
- Birch S, Maynard A, 1985. Performance indicators and performance assessment in the UK National Health Service: implications for management and planning. *Int J Health Planning Manag*;1:143–56.
- Birkmeyer JD, Aubuchon JP, Littenberg B, O'Connor GT, Nease RF, Nugent WC, et al., 1994. Cost-effectiveness of preoperative autologous donation in coronary artery bypass grafting. *Ann Thorac Surg*;57:161–9.
- Birkmeyer JD, Goodnough LT, Aubuchon JP, Noordsij, PG, Littenberg B, 1993. The cost-effectiveness of preoperative autologous blood donation for total hip and knee replacement. *Transfusion*;33:544–51.
- Bleichrodt H, 1995. QALYs and HYE: under what conditions are they equivalent? *J Health Econ*;14:17–37.
- Blondel B, Saurel-cubizolles MJ, Chwalow J, 1991. Social and psychological evaluation. *Rev Epidemiol Sante Publique*;39:S31–42.
- Blossfeld H, Hamerle A, Mayer KU, 1989. Event history analysis – statistical theory and application in the social sciences. Lawrence Erlbaum Associates.
- Blumenschein K, Johannesson M, 1996. Economic evaluation in health care: a brief history and future directions. *Pharmacoeconomics*;10:114–22.
- Blumenschein K, Johannesson M, 1996. Incorporating quality-of-life changes into economic evaluations of health care: an overview. *Health Policy*;36:155–66.
- BMDP Statistical Software Inc, 1988. BMDP statistical software manual, vol. 2. CA: University of California Press.
- Bocquet H, Pous J, Charlet JP, Grand A, 1996. Measuring the burden for carers of dependent elderly: the Zarit inventory. *Rev Epidemiol Sante Publique*;44:57–65.
- Bodger K, Daly MJ, Heatley RV, Williams DRR, 1996. Clinical economics review: gastroenterology. *Aliment Pharmacol Ther*;10:55–60.
- Bombardier C, Raboud J, 1991. A comparison of health-related quality-of-life measures for rheumatoid arthritis research. *Control Clin Trials*;12:S243–56.
- Bone MR, 1992. International efforts to measure health expectancy. *J Epidemiol Community Health*;46:555–8.
- Bonnetterre J, Schraub S, Lecomte S, Mercier M, 1996. Quality-of-life as an outcome in breast cancer: clinical application. *Pharmacoeconomics*;9:23–9.
- Bordley RF, 1994. Making social trade-offs among lives, disabilities, and cost. *J Risk Uncertainty*;9:135–49.
- Bosch JL, Hunink MGM, 1996. The relationship between descriptive and valuational quality of life measures in patients with intermittent claudication. *Med Decis Making*;16:217–25.
- Bostrom B, Woods WG, Nesbit ME, et al., 1987. Successful reinduction of patients with acute lymphoblastic leukemia who relapse following bone marrow transplantation. *J Clin Oncol*;5:376–81.
- Bouchet C, Guillemin F, Briancon S, 1995. Comparison of 3 quality-of-life measures for longitudinal studies in rheumatoid arthritis. *Rev Epidemiol Sante Publique*;43:250–8.
- Bouchet C, Guillemin F, Briancon S, 1996. Nonspecific effects in longitudinal studies: impact on quality-of-life measures. *J Clin Epidemiol*;49:15–20.

- Bowen DJ, Kestin M, McTiernan A, Carrell D, Green P, 1995. Effects of dietary fat intervention on mental health in women. *Cancer Epidemiol Biomarkers Prev*;4:555–9.
- Bowling A, 1991. Measuring health: a review of quality of life measurement scales. Buckingham; Open University Press.
- Bowling A, 1996. Measuring disease: a review of disease-specific quality of life measurement scales. Buckingham; Open University Press.
- Bowling A, 1996. The effects of illness on quality-of-life: findings from a survey of households in Great Britain. *J Epidemiol Community Health*;50:149–55.
- Boyle MH, Torrance GW, Sinclair JC, Horwood SP, 1983. Economic evaluation of neonatal intensive care of very low birth-weight infants. *N Engl J Med*;308:1330–7.
- Bozzette SA, Duan N, Berry S, Kanouse DE, 1994. Analytic difficulties in applying quality-of-life outcomes to clinical trials of therapy for HIV disease. *Psychology Health*;9:143–56.
- Bradley R, 1991. Bovine spongiform encephalopathy (BSE): the current situation and research. *Eur J Epidemiol*;7:532–44.
- Brasel KJ, Weigelt JA, 1996. Blunt thoracic aortic trauma: a cost–utility approach for injury detection. *Arch Surg*;131:619–25.
- Brazier J, Deverill M, Green C, Harper R, Booth A, 1999. A review of the use of health status measures in economic evaluation. *Health Technol Assess*;3(9).
- Brazier J, Dixon S, 1995. The use of condition specific outcome measures in economic appraisal. *Health Econ*;4:255–64.
- Breslow N, Crowley J, 1974. A large sample study of the life table and product limit estimates under random censorship. *Ann Stat*;2:437–53.
- Brooks MM, Gorkin L, Schron EB, Wiklund I, Campion J, Ledingham RB, 1994. Moricizine and quality-of-life in the Cardiac Arrhythmia Suppression Trial II (CAST II). *Control Clin Trials*;15:437–49.
- Broome J, 1993. QALYs. *J Publ Econ*;50:149–67.
- Brown ML, Nayfield SG, Shibley LM, 1994. Adjuvant therapy for stage III colon cancer: economics returns to research and cost-effectiveness of treatment. *J Natl Cancer Inst*;86:424–30.
- Brown PB, 1992. An alternative to QALYs: saved young life equivalent (SAVE). *BMJ*;305:1365.
- Brownie C, Hines JE, Nichols JD, Pollock KH, Hestbeck JB, 1993. Capture–recapture studies for multiple strata including non-Markovian transitions. *Biometrics*;49:1173–87.
- Bruce ML, Hoff RA, 1994. Social and physical health risk factors for first onset major depressive disorder in a community sample. *Soc Psychiatry Psychiat Epidemiol*;29:165–71.
- Bruinvels DJ, Kievit J, Habbema JDF, van de Velde CJH, 1993. Carcinoembryonic antigen monitoring in the follow-up of colorectal cancer patients: a decision analysis. *Theor Surg*;8:173–9.
- Brunner KW, 1989. Adjuvant therapies in breast cancer and quality of life: a critical review of the TWiST concept. In: Senn HJ, Goldhirsch A, Gelber RD, Osterwalder B, editors. Adjuvant therapy of primary breast cancer. Berlin: Springer-Verlag; 239–43.
- Bryan S, Parkin D, Donaldson C, 1991. Chiropractic and the QALY: a case-study in assigning categories of disability and distress to patients. *Health Policy*;18:169–85.
- Buchwald H, Fitch LL, Matts JP, Johnson JW, Hansen BJ, Stuenkel MR, *et al.*, 1993. Perception of quality-of-life before and after disclosure of trial results: a report from the program on the surgical control of the hyperlipidemias (POSCH). *Control Clin Trials*;14:500–10.
- Buckingham JK, 1987. Choosing between projects. The use of graphs to evaluate the use of benefit/cost ratios in cost effectiveness studies. *Hosp Health Services Rev*;83:268–71.
- Buckingham K, 1993. A note on HYE (healthy years equivalent). *J Health Econ*;12:301–9.
- Buckingham K, 1995. Economics, health and health economics: HYE versus QALYs – a response. *J Health Econ*;14:397–8.
- Bullinger M, Hasford J, 1991. Evaluating quality-of-life measures for clinical trials in Germany. *Control Clin Trials*;12:S91–105.
- Bulpitt CJ, Fletcher AE, 1990. Economic assessments in randomized controlled trials. *Med J Aust*;153:S16–19.
- Bulpitt CJ, Fletcher AE, 1993. Cost-effectiveness of the treatment of hypertension. *Clin Exp Hypertens*;15:1131–46.
- Buron C, le Gales C, Fericelli AM, 1994. Health-related quality of life approach and economic measurement of utility [abstract]. *Qual Life Res*;3:50–1.
- Burrows C, Brown K, 1993. QALYs for resource allocation: probably not and certainly not now. *Aust J Publ Health*;17:278–86.
- Busch MP, Korelitz JJ, Kleinman SH, Lee SR, Aubuchon JP, Schreiber GB, *et al.*, 1995. Declining value of alanine aminotransferase in screening of blood donors to prevent posttransfusion hepatitis-B and hepatitis-C virus infection. *Transfusion*;35:903–10.
- Busschbach JJV, Hensing DJ, Decharro FT, 1993. The utility of health at different stages in life: a quantitative approach. *Soc Sci Med*;37:153–8.
- Campbell M, Gibbard N. Frequency of use of quality-of-life instruments. To be published.
- Cantor SB, Clover RD, Thompson RF, 1994. A decision-analytic approach to postexposure rabies prophylaxis. *Am J Public Health*;84:1144–8.

- Cantor SB, Spann SJ, Volk RJ, Cardenas MP, Warren MM, 1995. Prostate cancer screening: a decision analysis. *J Fam Pract*;41:33–41.
- Carlens E, Dahlstrom G, Nou E, 1970. Comparative measurements of quality of survival of lung cancer patients after diagnosis. *Scand J Respir Dis*;51:268–75.
- Carr-Hill R, Morris J, 1991. The Q in QALYs. *BMJ*;303:1480.
- Carr-Hill RA, 1989. Assumptions of the QALY procedure. *Soc Sci Med*;29:469–77.
- Carr-Hill RA, 1991. Is the QALY a technical solution to a political problem? Of course not – a response. *Int J Health Serv*;21:371–2.
- Carr-Hill RA, 1991. Allocating resources to health care: is the QALY (quality adjusted life year) a technical solution to a political problem? *Int J Health Serv*;21:351–63.
- Carr-Hill RA, Morris J, 1991. Current practice in obtaining the Q in QALYs: a cautionary note. *BMJ*;303:699–701.
- Cartwright A, 1992. Social class differences in health and care in the year before death. *J Epidemiol Community Health*;46:54–7.
- Caudry M, Bonnel C, Floquet A, Marsault C, Quetin P, Pujol J, *et al.*, 1995. A randomized study of bolus fluorouracil plus folinic acid versus 21-day fluorouracil infusion alone or in association with cyclophosphamide and mitomycin C in advanced colorectal carcinoma. *Am J Clin Oncol: Cancer Clin Trials*;18:118–25.
- Chan CLH, Villar RN, 1996. Obesity and quality-of-life after primary hip arthroplasty. *J Bone Joint Surg Br*;78:78–81.
- Chang RW, Pellissier JM, Hazen GB, 1996. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. *JAMA*;275:858–65.
- Chang RW, Pellissier JM, Hazen GB, 1996. A cost-effectiveness analysis of total hip arthroplasty for osteoarthritis of the hip. *JAMA*;275:858–65.
- Chapman S, 1993. QALYs: a personal experience. *Aust J Publ Health*;17:397.
- Charlesworth A, Skene AM, 1997. A multi-state Markov process model for quality of life and survival in large heart failure clinical trials. *Control Clin Trials*;18:165S.
- Charny MC, Lewis PA, Farrow SC, 1989. Choosing who shall not be treated in the NHS. *Soc Sci Med*;28:1331–8.
- Chen MK, Yang GL, 1979. A quantitative index of evaluating patient care with longitudinal data. *Int J Epidemiol*;8:265–71.
- Cheung AP, Wren BG, 1992. A cost-effectiveness analysis of hormone replacement therapy in the menopause. *Med J Aust*;156:312–16.
- Chiang CL, 1980. An introduction to stochastic processes and their applications. Huntington, NY: Robert E Krieger.
- Chopra SK, Herdman CP, Laws D, Thomas AM, Wood P, 1989. The treatment of osteo-arthritis in general practice. Mefenamic acid (Ponstan Forte® compared to naproxen (Naprosyn®). *Clin Trials J*;26:401–10.
- Chow SC, Ki FYC, 1994. On statistical characteristics of quality of life assessment. *J Biopharm Stat*;4:1–17.
- Chow SC, Ki FYC, 1996. Statistical issues in quality-of-life assessment. *J Biopharm Stat*;6:37–48.
- Churchill DN, Wallace JE, Ludwin D, Beecroft ML, Taylor DW, 1991. A comparison of evaluative indexes of quality-of-life and cognitive function in hemodialysis patients. *Control Clin Trials*;12:S159–67.
- Clayton D, 1988. The analysis of event history data: a review of progress and outstanding problems. *Stat Med*;7:819–41.
- Cleary PD, Greenfield S, McNeil BJ, 1991. Assessing quality-of-life after surgery. *Control Clin Trials*;12:S189–203.
- Cleeton DL, Goepfrich VT, Weisbrod BA, 1992. What does the consumer price index for prescription drugs really measure? *Health Care Financ Rev*;13:45–51.
- Coast J, 1992. Reprocessing data to form QALYs. *BMJ*;305:87–90.
- Coast J, 1993. Developing the QALY concept: exploring the problems of data acquisition. *Pharmacoeconomics*;4:240–6.
- Coast J, 1993. The role of economic evaluation in setting priorities for elective surgery. *Health Policy*;24:243–57.
- Cogo R, Zavanella A, Gini M, 1984. Treatment of intrinsic or extrinsic asthma and chronic obstructive pulmonary disease. An extempore combination of bronchodilator (Salbutamol) with antihistamine (Oxatomide). *Clin Trials J*;21:384–92.
- Cohen A, Stessman J, Ginsberg GM, Issler C, Hammerman-Rozenberg R, Friedman R, *et al.*, 1995. The Jerusalem seventy year olds longitudinal study. II: Background results from the initial home interview. *Eur J Epidemiol*;11:685–92.
- Cohen DJ, Breall JA, Ho KKL, Kuntz RE, Goldman L, Baim DS, *et al.*, 1994. Evaluating the potential cost-effectiveness of stenting as a treatment for symptomatic single vessel coronary disease: use of a decision-analytic model. *Circulation*;89:1859–74.
- Cohen J, 1996. Preferences, needs and QALYs. *J Med Ethics*;22:267–72.
- Cole BF, Gelber RD, Anderson KM, 1994. Parametric approaches to quality-adjusted survival analysis. *Biometrics*;50:621–31.
- Cole BF, Gelber RD, Goldhirsch A, 1993. Cox regression models for quality-adjusted survival analysis. *Stat Med*;12:975–87.

- Cole BF, Gelber RD, Goldhirsch A, 1995. A quality-adjusted survival meta analysis of adjuvant chemotherapy for premenopausal breast cancer. *Stat Med*;14:1771–84.
- Cole BF, Gelber RD, Kirkwood JM, Goldhirsch A, Barylak E, Borden E, 1996. Quality-of-life-adjusted survival analysis of interferon alpha-2b adjuvant treatment of high risk resected cutaneous melanoma: an Eastern Cooperative Oncology Group Study. *J Clin Oncol*;14:2666.
- Cole BF, Solalceligny P, le Page E, 1995. Interferon-alpha for the treatment of advanced follicular lymphoma: an analysis of quality-of-life-adjusted survival. *Blood*;86:1745.
- Cole RP, Shakespeare V, Shakespeare P, Hobby JAE, 1994. Measuring outcome in low priority plastic surgery patients using quality-of-life indexes. *Br J Plast Surg*;47:117–21.
- Coles J, 1986. The myths and realities of DRGs in the NHS. *Hosp Health Services Rev*;82:28–31.
- Colice GL, Birkmeyer JD, Black WC, Littenberg B, Silvestri G, 1995. Cost-effectiveness of head CT in patients with lung cancer without clinical evidence of metastases. *Chest*;108:1264–71.
- Collett D, 1994. Modelling survival data in medical research. London: Chapman and Hall.
- Connors AF Jr, Dawson NV, Arkes HR, Roach MJ, 1990. Decision making in support: physician perceptions and preferences. *J Clin Epidemiol*;43:59–62S.
- Connors AF Jr, Dawson NV, Lynn J, 1990. Decision making in support: introduction. *J Clin Epidemiol*;43:47–9S.
- Cook DJ, Guyatt GH, Juniper E, Griffith L, McIlroy W, Willan A, *et al.*, 1993. Interviewer versus self-administered questionnaires in developing a disease-specific, health-related quality-of-life instrument for asthma. *J Clin Epidemiol*;46:529–34.
- Cook J, Richardson J, Street A, 1994. A cost–utility analysis of treatment options for gallstone disease: methodological issues and results. *Health Econ*;3:157–68.
- Cooper DN, 1994. Net investment in agricultural tractors. *J Agric Econ*;45:339–50.
- Cooper SD, Raymer JH, Pellizzari ED, Thomas KW, 1995. The identification of polar organic compounds found in consumer products and their toxicological properties. *J Exposure Anal Environ Epidemiol*;5:57–75.
- Corral FP, Abaira V, 1995. Autoperception and satisfaction with health. Two Med Care markers in elderly hospitalized patients: quality-of-life as an outcome estimate of clinical practice. *J Clin Epidemiol*;48:1031–40.
- Corry JF, Lonning PE, 1994. Systemic therapy in breast cancer: efficacy and cost–utility. *PharmacoEconomics*; 5:198–212.
- Corten P, Mercier C, Pelc I, 1994. ‘Subjective quality of life’: clinical model for assessment of rehabilitation treatment in psychiatry. *Soc Psychiatry Psychiatr Epidemiol*;29:178–83.
- Cory DC, 1985. Congestion costs and quality-adjusted user fees: a methodological note. *Land Econ*;61:452–5.
- Coste J, Fermanian J, Venot A, 1995. Methodological and statistical problems in the construction of composite measurement scales: a survey of 6 medical and epidemiologic journals. *Stat Med*;14:331–45.
- Coste J, Walter E, Venot A, 1995. A new approach to selection and weighting of items in evaluative composite measurement scales. *Stat Med*;14:2565–80.
- Cowen ME, Chartrand M, Weitzel WF, 1994. A Markov model of the natural history of prostate cancer. *J Clin Epidemiol*;47:3–21.
- Cox DR, 1972. Regression models and life tables. *J R Stat Soc B*;34:187–220.
- Cox DR, 1975. Partial likelihood. *Biometrika*;62:269–76.
- Cox DR, 1984. Some remarks on semi-Markov processes in medical statistics. In: Proceedings international symposium on semi-Markov processes and their applications. Brussels: June 1984.
- Cox DR, Fitzpatrick R, Fletcher AE, Gore SM, Spiegelhalter DJ, Jones DR, 1992. Quality-of-life assessment: can we keep it simple? *J R Stat Soc A*; 155:353–93.
- Cox DR, Miller HD, 1965. The theory of stochastic process. London: Chapman and Hall.
- Craig N, Walker D, 1996. Choice and accountability in health promotion: the role of health economics. *Health Educ Res*;11:355–60.
- Cribb A, Haycox A, 1989. Economic analysis in the evaluation of health promotion. *Community Med*;11:299–305.
- Cubbon J, 1991. The principle of QALY maximization as the basis for allocating health care resources. *J Med Ethics*;17:181–4.
- Cubbon J, 1992. Unprincipled QALYs – a response. *J Med Ethics*;18:100.
- Cullen MH, Billingham LJ, Woodroffe CM, Chetiyawardana AD, Gower NH, Joshi R, *et al.*, 1999. Mitomycin, ifosfamide and cisplatin (in unresectable non-small cell lung cancer: effects on survival and quality of life. *J Clin Oncol*. In press.
- Culyer AJ, Wagstaff A, 1992. QALYs versus HYE: a theoretical exposition. Discussion paper 99. York: Centre for Health Economics, University of York.
- Culyer AJ, Wagstaff A, 1993. QALYs versus HYE. *J Health Econ*;12:311–23.
- Culyer AJ, Wagstaff A, 1995. QALYs versus HYE – a reply to Gafni, Birch and Mehrez. *J Health Econ*;14:39–45.

- Dales RE, Belanger R, Shamji FM, Leech J, Crepeau A, Sachs HJ, 1994. Quality-of-life following thoracotomy for lung cancer. *J Clin Epidemiol*;47:1443–49.
- Daly E, Gray A, Barlow D, McPherson K, Roche M, Vessey M, 1993. Measuring the impact of menopausal symptoms on quality-of-life. *BMJ*;307:836–40.
- Daly E, Vessey MP, Barlow D, Gray A, McPherson K, Roche M, 1996. Hormone replacement therapy in a risk–benefit perspective. *Maturitas*;23:247–59.
- Danese MD, Powe NR, Sawin CT, Ladenson PW, 1996. Screening for mild thyroid failure at the periodic health examination: a decision and cost-effectiveness analysis. *JAMA*;276:285–92.
- Danford DA, 1990. QALYs: their ethical implications. *JAMA*;264:2503.
- Danoff B, Kramer S, Irwin P, Gottlieb A, 1983. Assessment of the quality of life in long-term survivors after definitive radiotherapy. *Am J Clin Oncol: Cancer Clin Trials*;6:339–45.
- David HA, Moeschberger ML, 1978. The theory of competing risks. London: Charles Griffin.
- Davies AM, 1987. Epidemiological approaches to disability in the elderly: from measurement to the projection of need. *Rev Epidemiol Sante Publique*;35:241–7.
- Davis CS, 1991. Semi-parametric and non-parametric methods for the analysis of repeated measurements with applications to clinical trials. *Stat Med*;10:1959–80.
- Davis M, Ettinger W, Neuhaus J, Barclay J, Segal M, 1991. Obesity and psychological well-being in relation to knee osteoarthritis and knee pain. *Am J Epidemiol*;134:788.
- Dazord A, Gerin P, 1994. Evaluation of treatment effects on patients quality-of-life: what approach – from the concepts to the instruments. *Clin Trials Meta-Anal*;29:1–8.
- de Cosse JJ, Cennerazzo WJ, 1996. A quality-adjusted time without symptoms or toxicity (Q-TWiST) analysis of adjuvant radiation therapy and chemotherapy for resectable rectal cancer. *J Natl Cancer Inst*;88:1686.
- de Gruttola V, Tu XM, 1994. Modelling progression of CD4 lymphocyte count and its relationship to survival time. *Biometrics*;50:1003–14.
- de Haes JCJM, de Koning HJ, van Oortmarssen GJ, van Agt HME, de Bruyn AE, van der Maas PJ, 1991. The impact of a breast cancer screening program on quality-adjusted life-years. *Int J Cancer*;49:538–44.
- de Haes JCJM, van Knippenberg FCE, 1989. Quality of life instruments for cancer patients: Babel's Tower revisited. *J Clin Epidemiol*;42:1239–41.
- de Koning HJ, van Ineveld BM, van Oortmarssen GJ, de Haes JCJM, Collette HJA, Hendriks JHCL, et al., 1991. Breast cancer screening and cost-effectiveness: policy alternatives, quality-of-life considerations and the possible impact of uncertain factors. *Int J Cancer*;49:531–7.
- de Stavola BL, Christensen E, 1996. Multilevel models for longitudinal variables prognostic for survival. *Lifetime Data Anal*;2:329–47.
- Defo BK, 1995. Epidemiology and control of infant and early childhood malaria: a competing risks analysis. *Int J Epidemiol*;24:204–17.
- Deliege D, 1983. Indicators of physical, mental and social wellbeing. *World Health Stat Q*;36:349–93.
- de Pauw S, Levine M, Guyatt G, 1987. Comparison of quality of life instruments in a setting of a clinical trial. *Control Clin Trials*;8:293–4.
- Depelsmacker P, 1988. Marketing, expenditure and quality-adjusted price effects on market share evolution in a segmented Belgian car market (1972–81). *Appl Econ*;20:15–30.
- Derdeyn CP, Powers WJ, 1996. Cost-effectiveness of screening for asymptomatic carotid atherosclerotic disease. *Stroke*;27:1944–50.
- Desbiens NA, 1992. Should all patients with atrial fibrillation be screened with echocardiography? *J Gen Intern Med*;7:131–9.
- Desch CE, Hillner BE, Smith TJ, Retchin SM, 1993. Should the elderly receive chemotherapy for node-negative breast cancer? A cost-effectiveness analysis examining total and active life-expectancy outcomes. *J Clin Oncol*;11:777–82.
- Detsky AS, McLaughlin JR, Abrams HB, et al., 1986. A cost–utility analysis of the home parenteral nutrition program at Toronto General Hospital: 1970–1982. *J Parenteral Enteral Nutr*;10:49–57.
- Dexter F, 1996. Application of cost–utility and quality-adjusted life years analyses to monitored anesthesia care for sedation only. *J Clin Anesth*;8:286–8.
- Deyo RA, Diehr P, Patrick DL, 1991. Reproducibility and responsiveness of health status measures: statistics and strategies for evaluation. *Control Clin Trials*;12:S142–58.
- Diamond GA, Denton TA, Matloff JM, 1993. Fee-for-benefit: a strategy to improve the quality of health care and control costs through reimbursement incentives. *J Am Coll Cardiol*;22:343–52.
- Dicostanzo F, Tagliaventi M, Carlini P, Zironi S, Mazzocchi B, Bella M, et al., 1996. High-dose folinic acid and fluorouracil with or without ifosfamide is an inactive combination in advanced pancreatic cancer: a randomized phase II study of the Italian oncology group for Clin Res (GOIRC). *Am J Clin Oncol: Cancer Clin Trials*;19:307–10.
- Dieras V, Marty M, Tublana N, Corette L, Morvan F, Serin D, et al., 1995. Phase II randomized study of paclitaxel versus mitomycin in advanced breast cancer. *Semin Oncol*;22:33–9.
- Dietrich GV, 1996. Cost-efficiency of preoperative autologous blood deposit compared with homologous blood transfusion. *Zentralbl Chir*;121:841–6.

- Diggle P, Kenward MG, 1994. Informative drop-out in longitudinal data analysis. *Appl Stat*;43:49–93.
- Diggle PJ, Liang K, Zeger SL, 1994. Analysis of longitudinal data. Oxford: Clarendon Press.
- Dippel DWJ, Avezaat CJJ, van der Ende ME, Habbema JDF, 1995. Clinical decision analysis in the case of a 60 year old HIV positive woman with a cerebral arteriovenous malformation. *Ned Tijdschr Geneesk*;139:2362–9.
- Dippel DWJ, Touwotten F, Habbema JDF, 1992. Management of children with acute pharyngitis: a decision analysis. *J Fam Pract*;34:149–59.
- Dippel DWJ, van Crevel H, Lindsay KW, Hijdra A, Habbema JDF, 1995. Management of subarachnoid hemorrhage: a decision analysis. *Cerebrovasc Dis*;5:350–7.
- Dippel DWJ, Vermeulen M, Braakman R, Habbema JDF, 1994. Transient ischemic attacks, carotid stenosis, and an incidental intracranial aneurysm: a decision analysis. *Neurosurgery*;34:449–57.
- Disch DL, Greenberg ML, Holzberger PT, Malenka DJ, Birkmeyer JD, 1994. Managing chronic atrial fibrillation: a Markov decision analysis comparing warfarin, quinidine, and low-dose amiodarone. *Ann Intern Med*;120:449–57.
- Dobson AJ, 1990. An introduction to generalized linear models. London: Chapman and Hall.
- Doel SR, Millar LJ, McEwan S, 1992. Quinapril and nifedipine in mild-to-moderate hypertension: a randomized, open, multicenter, parallel group comparison. *Clin Trials Meta-Anal*;28:29–38.
- Dolan P, 1994. Search for a critical appraisal of EuroQol: a response by the EuroQol group to Gafni and Birch. *Health Policy*;28:67–9.
- Domen RE, 1996. Preoperative autologous blood donation: clinical, economic, and ethical issues. *Cleve Clin J Med*;63:295–300.
- Donaldson C, Atkinson A, Bond J, Wright K, 1988. Should QALYs be program specific? *J Health Econ*;7:239–57.
- Donaldson C, Atkinson A, Bond J, Wright K, 1988. QALYs and long-term care for elderly people in the UK: scales for assessment of quality of life. *Age Ageing*;17:379–87.
- Donaldson C, Mooney G, 1991. Needs assessment, priority setting, and contracts for health care: an economic view. *BMJ*;303:1529–30.
- Donaldson C, Wright K, 1990. Program-specific QALYs – a reply. *J Health Econ*;8:489–91.
- Doubilet P, McNeil BJ, Weinstein MC, 1983. Optimal strategies for the diagnosis and treatment of coronary artery disease: analysis using microcomputer. *Med Decis Making*;3:23–8.
- Dougherty CJ, 1994. Quality-adjusted life years and the ethical values of health care. *Am J Phys Med Rehabil*;73:61–5.
- Dougherty CJ, 1995. Quality-adjusted life years and the ethical values of health care. *Am J Phys Med Rehabil*;74:S29–33.
- Downs SM, McNutt RA, Margolis PA, 1991. Management of infants at risk for occult bacteremia: a decision analysis. *J Pediatr*;118:11–20.
- Drewett RF, Minns RJ, Sibly TF, 1992. Measuring outcome of total knee replacement using quality-of-life indexes. *Ann R Coll Surg Engl*;74:286–90.
- Drummond M, Torrance G, Mason J, 1993. Cost-effectiveness league tables: more harm than good. *Soc Sci Med*;37:33–40.
- Drummond MF, 1987. Resource allocation decisions in health care: a role for quality of life assessments? *J Chronic Dis*;40:605–16.
- Drummond MF, 1987. Quality-adjusted life-years. *Lancet*;i:1372–3.
- Drummond MF, Heyse J, Cook J, McGuire A, 1993. Selection of end-points in economic evaluations of coronary heart disease interventions. *Med Decis Making*;13:184–90.
- Drummond MF, Mohide EA, Tew M, Streiner DL, Pringle DM, Gilbert JR, 1991. Economic evaluation of a support program for caregivers of demented elderly. *Int J Technol Assess Health Care*;7:209–19.
- Duncan C, Jones K, Moon G, 1995. Psychiatric morbidity: a multilevel approach to regional variations in the UK. *J Epidemiol Community Health*;49:290–5.
- Dusheiko GM, Roberts JA, 1995. Treatment of chronic type B and type C hepatitis with interferon alfa: an economic appraisal. *Hepatology*;22:1863–73.
- Ebbesen LS, Guyatt GH, McCartney N, Oldridge NB, 1990. Measuring quality of life in cardiac spouses. *J Clin Epidemiol*;43:481–7.
- Ebrahim S, Britts S, Wu A, 1991. The valuation of states of ill health: the impact of age and disability. *Age Ageing*;20:37–40.
- Eckman MH, Beshansky JR, Durand-Zaleski I, Levine HJ, Pauker SG, 1990. Anticoagulation for noncardiac procedures in patients with prosthetic heart valves: does low risk mean high cost? *JAMA*;263:1513–21.
- Eckman MH, Greenfield S, Mackey WC, Wong JB, Kaplan S, Sullivan L, et al., 1995. Foot infections in diabetic patients: decision and cost-effectiveness analyses. *JAMA*;273:712–20.
- Eckman MH, Levine HJ, Pauker SG, 1993. Effect of laboratory variation in the prothrombin time ratio on the results of oral anticoagulant therapy. *N Engl J Med*;329:696–702.
- Edgar A, 1995. Weighting health states and strong evaluation. *Bioethics*;9:240–51.
- Efron B, 1981. Censored data and the bootstrap. *J Am Stat Assoc*;76:312–19.



- Ekelund R, 1991. National survey of oral health care in Finnish municipal old peoples homes. *Community Dent Oral Epidemiol*;19:169–72.
- Elvik R, 1995. The validity of using health state indexes in measuring the consequences of traffic injury for public health. *Soc Sci Med*;40:1385–8.
- Elwes RDC, 1996. Health economics and the selection of antiepileptic drug treatment. *Eur J Neurol*;3:39–43.
- Enghofer E, 1992. Quality of life as 'eye-catcher' and promoter in the pharmaceutical industry. *Wien Med Wochenschr*;142:536–8.
- Epstein RS, Deverka PA, Chute CG, Panser L, Oesterling JE, Lieber MM, *et al.*, 1992. Validation of a new quality-of-life questionnaire for benign prostatic hyperplasia. *J Clin Epidemiol*;45:1431–45.
- Eraker SA, Baker S, Miyamoto JM, 1984. Parameter estimates for a quality adjusted life year utility model. *Clin Res*;32:A294.
- Erickson P, Wilson RW, 1994. The term years of healthy life: misunderstood, defended, and challenged. A short term for quality-adjusted life years. *Am J Public Health*;84:866.
- Esrey SA, 1996. Water, waste, and well-being: a multicountry study. *Am J Epidemiol*;143:608–23.
- Etchason J, Petz L, Keeler E, Calhoun L, Kleinman S, Snider C, *et al.*, 1995. The cost-effectiveness of pre-operative autologous blood donations. *N Engl J Med*;332:719–24.
- Evans RW, 1987. Quality-adjusted life-years. *Lancet*;i:1372.
- Fairclough DL, 1995. Summary measures for comparison of QOL in clinical trials [abstract]. *Qual Life Res*;4:423.
- Fallowfield L, 1990. The quality of life: the missing measurement in health care. London: Souvenir Press.
- Falvey RE, Fried HO, Richards B, 1992. An hedonic guide to New Orleans restaurants. *Q Rev Econ Finance*;32:123–33.
- Fanshel S, Bush JW, 1970. A health-status index and its applications to health services outcomes. *Operations Res*;18:1021–66.
- Faucett CL, Thomas DC, 1996. Simultaneously modelling censored survival data and repeatedly measured covariates: a Gibbs sampling approach. *Stat Med*;15:1663–85.
- Fayers P, 1995. MRC quality of life studies using a daily diary card: practical lessons learned from cancer trials. *Qual Life Res*;4:343–52.
- Fayers PM, Jones DR, 1983. Measuring and analysing quality of life in cancer clinical trials: a review. *Stat Med*;2:429–46.
- Feeny D, Guyatt G, Patrick D, 1991. Proceedings of the international conference on the measurement of quality-of-life as an outcome in clinical trials: June 14–17, 1989: introduction. *Control Clin Trials*;12:S79–80.
- Feldstein ML, 1991. Quality-of-life-adjusted survival for comparing cancer treatments: a commentary on TWiST and Q-TWiST. *Cancer*;67:851–4.
- Fentonlee D, Imrie CW, 1993. Pancreatic necrosis: assessment of outcome related to quality-of-life and cost of management. *Br J Surg*;80:1579–82.
- Fergany N, 1994. Quality-of-life indexes for Arab countries in an international context. *Int Stat Rev*;62:187–202.
- Ferguson BM, Keown PA, 1995. An introduction to utility measurement in health care. *Infect Control Hosp Epidemiol*;16:240–47.
- Ferland P, Pampalon R, Sauve J, 1996. Indicator of needs for regional resource allocation in Quebec. *Can J Public Health*;87:280–5.
- Fernandezcornejo J, Jans S, 1994. Quality-adjusted price and quantity indexes for pesticides. *Am J Agric Econ*;76:1259.
- Fernandezcornejo J, Jans S. 1995. Quality-adjusted price and quantity indexes for pesticides. *Am J Agric Econ*;77:645–59.
- Fernandezlopez JA, Siegrist J, Hernandezmejia R, Broer M, Cuetoespinar A, 1994. Study of quality-of-life on rural hypertensive patients: comparison with the general population of the same environment. *J Clin Epidemiol*;47:1373–80.
- Feuer EJ, Hankey BF, Gaynor JJ, Wesley MN, Baker SG, Meyer JS, 1992. Graphical representation of survival curves associated with a binary non-reversible time dependent covariate. *Stat Med*;11:455–74.
- Finkelstein DM, Cassileth BR, Bonomi PD, Ruckdeschel JC, Ezdinli EZ, Wolter JM, 1988. A pilot study of the Functional Living Index Cancer (FLIC) scale for the assessment of quality of life for metastatic lung cancer patients. An Eastern Cooperative Oncology Group study. *Am J Clin Oncol: Cancer Clin Trials*;11:630–3.
- Fiscella K, Franks P, 1996. Cost-effectiveness of the transdermal nicotine patch as an adjunct to physicians smoking cessation counseling. *JAMA*;275:1247–51.
- Fitzgerald JM, 1989. The taste for bequests and well-being of widows: a model of life insurance demand by married couples. *Rev Econ Stat*;71:206–14.
- Fitzpatrick R, Davey C, Buxton MJ, Jones DR, 1998. Evaluating patient-based outcome measures for use in clinical trials. *Health Technol Assess*;2(14).
- Fitzpatrick R, Fletcher A, Gore SM, Jones, DR, Spiegelhalter DJ, Cox DR, 1992. Quality of life measures in health care. I: applications and issues in assessment. *BMJ*;305:1074–7.
- Fix E, Neyman J, 1951. A simple stochastic model of recovery, relapse and loss of patients. *Hum Biol*;23:205–41.

- Fleming C, Wasson JH, Albertsen PC, Barry MJ, Wennberg JE, 1993. A decision analysis of alternative treatment strategies for clinically localized prostate cancer. *JAMA*;269:2650–8.
- Fleming ST, 1991. The relationship between quality and cost: pure and simple? *Inquiry*;28:29–38.
- Fleming TR, Harrington DP, 1991. Counting processes and survival analysis. New York: John Wiley.
- Fletcher A, 1991. Pressure to treat and pressure to cost: a review of cost-effectiveness analysis. *J Hypertension*;9:193–8.
- Fletcher A, Gore SM, Jones DR, Fitzpatrick R, Spiegelhalter DJ, Cox DR, 1992. Quality of life measures in health care. II: design, analysis and interpretation. *BMJ*;305:1145–8.
- Fletcher SW, 1994. Summing up: epidemiology's effect on women's health and vice versa. *Ann Epidemiol*;4:174–5.
- Follmann D, Wittes J, Cutler JA, 1992. The use of subjective rankings in clinical trials with an application to cardiovascular disease. *Stat Med*;11:427–37.
- Foppa I, Noack H, Minder CE, 1995. The relation of reported symptoms to social, individual, and behavioral indicators of ill health. Is the number of reported symptoms a unique general dimension of ill health? *J Clin Epidemiol*;48:941–8.
- Ford ES, Eaker E, 1991. Ischemic heart disease, well-being, and depression. *Am J Epidemiol*;134:762.
- Fordham R, 1993. Hip fractures and QALYs. *J Bone Joint Surg Br*;75:163–4.
- Forjuoh SN, Guyer B, Strobino DM, Keyl PM, Dienerwest M, Smith GS, 1995. Risk factors for childhood burns: a case-control study of Ghanaian children. *J Epidemiol Community Health*;49:189–93.
- Fox-Rushby JA, 1995. The Gambia: cost and effectiveness of a mobile maternal health care service, West Kiang. *World Health Stat Q*;48:23–7.
- Francom SF, Chuang-Stein C, Landis JR, 1989. A log-linear model for ordinal data to characterize differential change among treatments. *Stat Med*;8:571–82.
- Freedberg KA, Tosteson ANA, Cotton DJ, Goldman L, 1992. Optimal management strategies for HIV-infected patients who present with cough or dyspnea: a cost-effectiveness analysis. *J Gen Intern Med*;7:261–72.
- French MT, Mauskopf JA, Teague JL, Roland EJ, 1996. Estimating the dollar value of health outcomes from drug abuse interventions. *Med Care*;34:890–910.
- Friedman HL, 1994. Reproductive health in adolescence. *World Health Stat Q*;47:31–5.
- Fries JF, 1991. The hierarchy of quality-of-life assessment, the health assessment questionnaire (HAQ) and issues mandating development of a toxicity index. *Control Clin Trials*;12:S106–17.
- Fryback DG, 1993. QALYs, HYE, and the loss of innocence. *Med Decis Making*;13:271–2.
- Fujita A, Takabatake H, Tagaki S, Sohda T, Sekine K, 1996. Evaluation of chemotherapy for stage IV non small cell lung cancer employing a regression tree type method for quality-adjusted survival analysis to determine prognostic factors. *Jpn J Cancer Chemother*;23:433–8.
- Fulgente T, Onofri M, Del Re ML, Ferracci F, Bazzano S, Ghilardi MF, et al., 1990. Laevo-acetylcarnitine (Nicetile®) treatment of senile depression. *Clin Trials J*;27:155–63.
- Gabriel SA, Rosenthal SS, 1996. Commutes, neighborhood effects, and earnings: an analysis of racial discrimination and compensating differentials. *J Urban Econ*;40:61–83.
- Gabriel SE, 1995. Economic evaluation using mathematical models: the case of misoprostol prophylaxis. *J Rheumatol*;22:1412–14.
- Gabriel SE, Champion ME, Ofallon WM, 1994. A cost-utility analysis of misoprostol prophylaxis for rheumatoid arthritis patients receiving nonsteroidal antiinflammatory drugs. *Arthritis Rheum*;37:333–41.
- Gafni A, 1989. The quality of QALYs (quality-adjusted-life-years): do QALYs measure what they at least intend to measure? *Health Policy*;13:81–3.
- Gafni A, 1991. Measuring the adverse effects of unnecessary hypertension drug therapy: QALYs vs HYE. *Clin Invest Med*;14:266–70.
- Gafni A, 1994. The standard gamble method: what is being measured and how it is interpreted. *Health Serv Res*;29:207–24.
- Gafni A, 1996. HYE: do we need them and can they fulfil the promise? *Med Decis Making*;16:215–16.
- Gafni A, Birch S, 1991. Equity considerations in utility-based measures of health outcomes in economic appraisals: an adjustment algorithm. *J Health Econ*;10:329–42.
- Gafni A, Birch S, 1993. Searching for a common currency: critical appraisal of the scientific basis underlying European harmonization of the measurement of health related quality-of-life (EuroQol(C)). *Health Policy*;23:219–28.
- Gafni A, Birch S, 1993. Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem. *Can Med Assoc J*;148:913–17.
- Gafni A, Birch S, 1995. Preferences for outcomes in economic evaluation: an economic approach to addressing economic problems. *Soc Sci Med*;40:767–76.
- Gafni A, Birch S, Mehrez A, 1993. Economics, health and health economics: HYE versus QALYs. *J Health Econ*;12:325–39.

- Gage BF, Cardinalli AB, Albers GW, Owens DK, 1995. Cost-effectiveness of warfarin and aspirin for prophylaxis of stroke in patients with nonvalvular atrial fibrillation. *JAMA*;274:1839–45.
- Gage BF, Cardinalli AB, Owens DK, 1996. The effect of stroke and stroke prophylaxis with aspirin or warfarin on quality-of-life. *Arch Intern Med*;156:1829–36.
- Gail MH, 1981. Evaluating serial cancer marker studies in patients at risk of recurrent disease. *Biometrics*;37:67–78.
- Gandal N, 1994. Hedonic price indexes for spreadsheets and an empirical test for network externalities. *Rand J Econ*;25:160–70.
- Ganiats TG, Browner DK, Kaplan RM, 1996. Comparison of two methods of calculating quality-adjusted life years. *Qual Life Res*;5:162–4.
- Ganiats TG, Miller CJ, Kaplan RM, 1995. Comparing the quality-adjusted life-year output of 2 treatment arms in a randomized trial. *Med Care*;33:AS245–54.
- Ganz PA, Lee JJ, Sim MS, Polinsky ML, Schag CAC, 1992. Exploring the influence of multiple variables on the relationship of age to quality-of-life in women with breast cancer. *J Clin Epidemiol*;45:473–85.
- Ganz PA, Schag CAC, Cheng HL, 1990. Assessing the quality of life: a study in newly diagnosed breast cancer patients. *J Clin Epidemiol*;43:75–86.
- Garratt AM, Ruta DA, Russell I, MacLeod K, Brunt P, McKinlay A, *et al.*, 1996. Developing a condition-specific measure of health for patients with dyspepsia and ulcer-related symptoms. *J Clin Epidemiol*;49:565–71.
- Gaynor M, Trapani JM, 1994. Quantity, quality and the welfare effects of United States airline deregulation. *Appl Econ*;26:543–50.
- Geelhoed E, Harris A, Prince R, 1994. Cost-effectiveness analysis of hormone replacement therapy and life-style intervention for hip fracture. *Aust J Publ Health*;18:153–60.
- Gelber RD, Cole BF, Gelber S, Goldhirsch A, 1995. Comparing treatments using quality-adjusted survival: the Q-TWiST method. *Am Stat*;49:161–9.
- Gelber RD, Cole BF, Goldhirsch A, 1995. Adjuvant chemotherapy for premenopausal breast cancer: a meta-analysis using quality-adjusted survival. *Cancer J Sc Am*;1:114–21.
- Gelber RD, Cole BF, Goldhirsch A, Rose C, Fisher B, Osborne CK, *et al.*, 1996. Adjuvant chemotherapy plus tamoxifen compared with tamoxifen alone for postmenopausal breast cancer: metaanalysis of quality-adjusted survival. *Lancet*;347:1066–71.
- Gelber RD, Gelman RS, Goldhirsch A, 1989. A quality-of-life-oriented endpoint for comparing therapies. *Biometrics*;45:781–95.
- Gelber RD, Goldhirsch A, 1986. A new endpoint for the assessment of adjuvant therapy in postmenopausal women with operable breast cancer. *J Clin Oncol*;4:1772–9.
- Gelber RD, Goldhirsch A, 1989. Methods for assessing treatment efficacy in trials for adjuvant therapy for breast cancer. In: Senn HJ, Goldhirsch A, Gelber RD, Osterwalder B, editors. *Adjuvant therapy of primary breast cancer*. Berlin: Springer-Verlag; 211–19.
- Gelber RD, Goldhirsch A, 1991. Quality-adjusted survival analysis – reply. *J Clin Oncol*;9:526.
- Gelber RD, Goldhirsch A, Castiglione M, Price K, Isley M, Coates A, 1987. Time without symptoms and toxicity (TWiST): a quality-of-life-oriented endpoint to evaluate adjuvant therapy. In: Salmon SE, editor. *Adjuvant therapy of cancer V*. Orlando: Grune and Stratton; 455–65.
- Gelber RD, Goldhirsch A, Cavalli F, 1991. Quality-of-life-adjusted evaluation of adjuvant therapies for operable breast cancer. *Ann Intern Med*;114:621–8.
- Gelber RD, Goldhirsch A, Cole BF, 1993. Evaluation of effectiveness: Q-TWiST. *Cancer Treat Rev*;19:73–84.
- Gelber RD, Goldhirsch A, Cole BF, 1993. Parametric extrapolation of survival estimates with applications to quality-of-life evaluation of treatments. *Control Clin Trials*;14:485–99.
- Gelber RD, Goldhirsch A, Cole BF, Weiland HS, Schroeder G, Krook JE, 1996. A quality-adjusted time without symptoms or toxicity (Q-TWiST) analysis of adjuvant radiation therapy and chemotherapy for resectable rectal cancer. *J Natl Cancer Inst*;88:1039–45.
- Gelber RD, Lenderking WR, Cotton DJ, Cole BF, Fischl MA, Goldhirsch A, *et al.*, 1992. Quality-of-life evaluation in a clinical trial of zidovudine therapy in patients with mildly symptomatic HIV infection. *Ann Intern Med*;116:961–6.
- Gemke RJB, Bonsel GJ, 1996. Reliability and validity of a comprehensive health-status measure in a heterogeneous population of children admitted to intensive care. *J Clin Epidemiol*;49:327–33.
- Gentleman RC, Lawless JF, Lindsey JC, Yan P, 1994. Multistate Markov models for analyzing incomplete disease history data with illustrations for HIV disease. *Stat Med*;13:805–21.
- Gerard K, 1992. Cost-utility in practice: a policy makers guide to the state-of-the-art. *Health Policy*;21:249–79.
- Gerard K, Dobson M, Hall J, 1993. Framing and labeling effects in health descriptions: quality adjusted life years for treatment of breast cancer. *J Clin Epidemiol*;46:77–84.
- Gertler PJ, Waldman DM, 1992. Quality-adjusted cost functions and policy evaluation in the nursing home industry. *J Political Econ*;100:1232–56.
- Gibson D, Harvey AJ, Everett V, Parmar MKB, 1994. Is double data entry necessary? The CHART trials. *Control Clin Trials*;15:482–8.

- Gilbert H, Kagan AR, Cassidy F, *et al.*, 1981. Glioblastoma multiforme is not a uniform disease! *Cancer Clin Trials*;4:87–9.
- Giorgi F, Cellerino R, Gramazio A, Tummarello D, Menichetti ET, Giordani P, *et al.*, 1996. Assessing quality-of-life in patients with cancer: a comparison of a visual-analog and a categorical model. *Am J Clin Oncol: Cancer Clin Trials*;19:394–9.
- Glasziou PP, Simes RJ, Gelber RD, 1990. Quality adjusted survival analysis. *Stat Med*;9:1259–76.
- Glicher SR, 1992. Alternative to QALYs: saved young life equivalent (SAVE). *BMJ*;305:1365.
- Glimelius B, Hoffman K, Graf W, Haglund U, Nyren O, Pahlman L, *et al.*, 1995. Cost-effectiveness of palliative chemotherapy in advanced gastrointestinal cancer. *Ann Oncol*;6:267–74.
- Glimelius B, Hoffman K, Sjoden PO, Jacobsson G, Sellstrom H, Enander LK, *et al.*, 1996. Chemotherapy improves survival and quality-of-life in advanced pancreatic and biliary cancer. *Ann Oncol*;7:593–600.
- Glutzer DE, Freedberg KA, Bauchner H, 1995. Management of childhood lead poisoning: clinical impact and cost-effectiveness. *Med Decis Making*;15:13–24.
- Goel V, Deber RB, Detsky AS, 1989. Nonionic contrast media: economic analysis and health policy development. *Can Med Assoc J*;140:389–95.
- Goldhirsch A, Gelber RD, Simes RJ, Glasziou P, Coates AS, 1989. Costs and benefits of adjuvant therapy in breast cancer: a quality-adjusted survival analysis. *J Clin Oncol*;7:36–44.
- Goldstein G, 1990. Urbanization, health and well-being: a global perspective. *Statistician*;39:121–33.
- Goldstein H, 1995. Multilevel statistical models. 2nd ed. London: Edward Arnold.
- Goodnough LT, Grishaber JE, Birkmeyer JD, Monk TG, Catalona WJ, 1994. Efficacy and cost-effectiveness of autologous blood predeposit in patients undergoing radical prostatectomy procedures. *Urology*;44:226–31.
- Goodwin PJ, Field R, Evans WK, Pater J, 1988. Cost-effectiveness of cancer chemotherapy: an economic evaluation of a randomised trial in small-cell lung cancer. *J Clin Oncol*;6:1537–47.
- Goossens MEJB, Rutten-van Molken MPMH, Leidl RM, Bos SGPM, Vlaeyen JWS, Teekengruben NJG, 1996. Cognitive educational treatment of fibromyalgia: a randomized clinical trial. 2. Economic evaluation. *J Rheumatol*;23:1246–54.
- Gore SM, 1988. Integrated reporting of quality and length of life: a statistician's perspective. *Eur Heart J*;9:228–34.
- Gorsky RD, Koplan JP, Peterson HB, Thacker SB, 1994. Relative risks and benefits of long-term estrogen replacement therapy: a decision analysis. *Obstet Gynecol*;83:161–6.
- Gottlieb LK, Schwartz B, Pauker SG, 1983. Glaucoma screening. A cost-effectiveness analysis. *Surv Ophthalmol*;28:206–26.
- Gottlieb RH, Mooney C, Mushlin AI, Rubens DJ, Fultz PJ, 1996. The prostate: decreasing cost-effectiveness of biopsy with advancing age. *Invest Radiol*;31:84–90.
- Gottschau A, Hogh B, 1995. Interval censored survival data and multistate compartmental models in the analysis of first appearance of plasmodium falciparum parasites in infants. *Stat Med*;14:2727–36.
- Gould AL, 1980. A new approach to the analysis of clinical drug trials with withdrawals. *Biometrics*;36:721–7.
- Gournay K, Brooking J, 1995. The community psychiatric nurse in primary care: an economic analysis. *J Adv Nurs*;22:769–78.
- Grand A, Grosclaude P, Bocquet H, Pous J, Albaredo JL, 1990. Disability, psychosocial factors and mortality among the elderly in a rural French population. *J Clin Epidemiol*;43:773–82.
- Green SB, 1997. Does assessment of quality of life in comparative cancer trials make a difference? A discussion. *Control Clin Trials*;18:306–10.
- Gregor JC, Ponich TP, Detsky AS, 1996. Should ERCP be routine after an episode of idiopathic pancreatitis: a cost–utility analysis. *Gastrointest Endosc*;44:118–23.
- Gregory JR, McMurtrey MJ, Mountain CF, 1985. A surgical approach to the treatment of pericardial effusion in cancer patients. *Am J Clin Oncol: Cancer Clin Trials*;8:319–23.
- Groenvold M, Bjorner JB, Klee MC, Kreiner S, 1995. Test for item bias in a quality-of-life questionnaire. *J Clin Epidemiol*;48:805–16.
- Groot CUD, Hagenbeek A, Verdonck LF, Lowenberg B, Rutten FFH, 1995. Cost-effectiveness of ABMT in comparison with CHOP chemotherapy in patients with intermediate and high grade malignant non-Hodgkins lymphoma (NHL). *Bone Marrow Transplant*;16:463–70.
- Gruger J, Kay R, Schumacher M, 1991. The validity of inferences based on incomplete observations in disease state models. *Biometrics*;47:595–605.
- Grunberg SM, 1995. Economic impact of antiemesis. *Oncology*;9:155–60.
- Grunberg SM, Boutin N, Ireland A, Miner S, Silveira J, Ashikaga T, 1996. Impact of nausea/vomiting on quality-of-life as a visual analog scale derived utility score. *Support Care Cancer*;4:435–9.
- Grundmann RT, 1996. Quality assurance in surgery: more than the documentation of postoperative complications. *Zentralbl Chir*;121:157–66.
- Gudex C, 1986. QALYs and their use by the Health Service. Discussion paper 20. York: Centre for Health Economics, University of York.
- Gudex C, 1996. Measuring patient benefit in mental illness. *Eur Psychiatry*;11:155–8.

- Gudex C, Kind P, 1987. The QALY toolkit. Discussion Paper. York: Centre for Health Economics, University of York.
- Gudex C, Kind P, 1991. The Q in QALYs. *BMJ*;303:1137.
- Gudex C, Kind P, 1991. Chiropody and the QALY: a case study in assigning categories and distress to patients. *Health Policy*;19:79–80.
- Guillemin F, Bombardier C, Beaton D, 1993. Cross-cultural adaptation of health-related quality-of-life measures: literature review and proposed guidelines. *J Clin Epidemiol*;46:1417–32.
- Guillemin F, Bombardier C, Beaton D, 1994. Cross-cultural validation for quality-of-life – response. *J Clin Epidemiol*;47:1466.
- Gunnell D, Harvey I, Smith L, 1995. The invasive management of angina: issues for consumers and commissioners. *J Epidemiol Community Health*;49:335–43.
- Guo WS, Marshall G, 1995. ORDMKV: a computer program fitting proportional odds model for multistate Markov process. *Comput Methods Programs Biomed*;46:257–63.
- Guyatt G, Feeny D, Patrick D, 1991. Issues in quality-of-life measurement in clinical trials. *Control Clin Trials*;12:S81–90.
- Guyatt G, Patrick D, Feeny D, 1991. Proceedings of the international conference on the measurement of quality of life as an outcome in clinical trials: postscript. *Control Clin Trials*;12:266–9S.
- Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A, 1989. Responsiveness and validity in health status measurement: a clarification. *J Clin Epidemiol*;42:403–8.
- Guyatt GH, Eagle DJ, Sackett B, Willan A, Griffith L, McIlroy W, *et al.*, 1993. Measuring quality-of-life in the frail elderly. *J Clin Epidemiol*;46:1433–44.
- Guyatt GH, Kirshner B, Jaeschke R, 1992. Measuring health status: what are the necessary measurement properties? *J Clin Epidemiol*;45:1341–5.
- Guyatt GH, Townsend M, Keller JL, Singer J, 1989. Should study subjects see their previous responses: data from a randomized control trial. *J Clin Epidemiol*;42:913–20.
- Habu, Y, Hayashi K, Watanabe Y, Kawai K, 1992. A clinical decision analysis to assess the value of laparoscopic cholecystectomy in the treatment of gallstone disease. *Endoscopy*;24:709–16.
- Hadorn DC, Hays RD, Uebersax J, Hauber T, 1992. Improving task comprehension in the measurement of health state preferences: a trial of informational cartoon figures and a paired comparison task. *J Clin Epidemiol*;45:233–43.
- Hadorn DC, Sorensen J, Holte J, 1995. Large scale health outcomes evaluation: how should quality-of-life be measured? 2. Questionnaire validation in a cohort of patients with advanced cancer. *J Clin Epidemiol*;48:619–29.
- Hadorn DC, Uebersax J, 1995. Large-scale health outcomes evaluation: how should quality-of-life be measured. 1. Calibration of a brief questionnaire and a search for preference subgroups. *J Clin Epidemiol*;48:607–18.
- Haes JCJM, 1993. The integration of quality-of-life and survival: quality-adjusted life years. *Qual Life Res*;2:60.
- Halford WK, Hayes RL, 1995. Social skills in schizophrenia: assessing the relationship between social skills, psychopathology and community functioning. *Soc Psychiatry Psychiat Epidemiol*;30:14–19.
- Hall J, 1989. Best medical practice in practice: measuring efficiency in mammography screening. *Int J Health Planning Manage*;4:235–46.
- Hall J, Gerard K, Salkeld G, Richardson J, 1992. A cost utility analysis of mammography screening in Australia. *Soc Sci Med*;34:993–1004.
- Hand D, Crowder M, 1996. Practical longitudinal data analysis. London: Chapman and Hall.
- Hansen BE, Thorogood J, Hermans J, Ploeg RJ, van Bockel JH, van Houwelingen JC, 1994. Multistate modeling of liver transplantation data. *Stat Med*;13:2517–29.
- Harris A, Hendrie D, Bower C, Payne J, Deklerk N, Stanley F, 1994. The burden of haemophilus influenzae type-B disease in Australia and an economic appraisal of the vaccine PRP-OMP. *Med J Aust*;160:483–8.
- Harris J, 1987. QALYfying the value of life. *J Med Ethics*;13:117–23.
- Harris J, 1988. Life: quality, value and justice. *Health Policy*;10:259–66.
- Harris J, 1991. Unprincipled QALYs – a response to Cubbon. *J Med Ethics*;17:185–8.
- Harris J, 1992. Unprincipled QALYs. *J Med Ethics*;18:162.
- Harris J, 1995. Double jeopardy and the veil of ignorance – a reply. *J Med Ethics*;21:151–7.
- Harris J, 1996. Would Aristotle have played Russian roulette? *J Med Ethics*;22:209–15.
- Harris JR, Tambs K, Magnus P, 1995. Sex-specific effects for body mass index in the new Norwegian twin panel. *Genet Epidemiol*;12:251–65.
- Harron DWG, 1995. Pharmacoeconomics: quality use of quality medicines and its impact on the Middle East. *Int Pharmacy J*;9:161–3.
- Harwood R, 1992. Reprocessing data to form QALYs. *BMJ*;305:424.
- Hasson HM, 1995. Endoscopic surgery, cost-effectiveness, and the quality-of-life. *J Am Assoc Gynecol Laparosc*;2:115–21.

- Hastie TJ, Botha JL, Schnitzler CM, 1989. Regression with an ordered categorical response. *Stat Med*;8:785–94.
- Hatziandreu EI, Koplan JP, Weinstein MC, Caspersen CJ, Warner KE, 1988. A cost-effectiveness analysis of exercise as a health promotion activity. *Am J Public Health*;78:1417–21.
- Hatziandreu EJ, Brown RE, Revicki DA, Turner R, Martindale J, Levine S, Siegel JE, 1994. Cost–utility of maintenance treatment of recurrent depression with sertraline versus episodic treatment with dothiepin. *Pharmacoeconomics*;5:249–64.
- Hayashi C, Suzuki T, 1984. Changes in belief systems, quality of life issues and social conditions over 25 years in post-war Japan. *Ann Inst Stat Math*;36:135–61.
- Hayry M, 1991. Measuring the quality-of-life: why, how and what. *Theor Med*;12:97–116.
- Hays RD, 1995. Directions for future research. *Qual Life Res*;4:179–80.
- Hays RD, Anderson R, Revicki D, 1993. Psychometric considerations in evaluating health-related quality of life measures. *Qual Life Res*;2:441–9.
- Hays RD, Stewart AL, Sherbourne CD, Marshall GN, 1993. The ‘states versus weights’ dilemma in quality of life measurement. *Qual Life Res*;2:167–8.
- Heckerling PS, Verp MS, 1994. A cost-effectiveness analysis of amniocentesis and chorionic villus sampling for prenatal genetic testing. *Med Care*;32:863–80.
- Heinemann LA, Jones DH, 1993. Assessment of physical activity in East Germany. *Ann Epidemiol*;3:S85–9.
- Heinemann LAJ, Greiser EM, 1993. Blood pressure, hypertension, and other risk factors in East and West Germany. *Ann Epidemiol*;3:S90–5.
- Helderman WHV, Mkasabuni E, 1993. Impact of dental fluorosis on the perception of well-being in an endemic fluorosis area in Tanzania. *Community Dent Oral Epidemiol*;21:243–4.
- Helms RW, Laird NN, Lebowitz MD, Mantel N, Louis TA, Wu M. 1988. Discussion: analysis of data with missing values. *Stat Med*;7:357–60.
- Helms RW, Edwards LJ, Okamoto LJ, Mahajan P, Hansen JE, 1995. Evaluating the validity of linear model analyses of quality of life scales [abstract]. *Qual Life Res*;4:438–9.
- Herndon JE, Fleishman S, Kosty MP, Green MR, 1997. A longitudinal study of quality of life in advanced non-small cell lung cancer: Cancer and Leukemia Group B (CALGB) 8931. *Control Clin Trials*;18:286–300.
- Heyink J, Schaub R, 1986. Denture problems and the quality of life in a Dutch elderly population. *Community Dent Oral Epidemiol*;14:193–4.
- Heyting A, Tolboom JTB, Essers JGA, 1992. Statistical handling of drop-outs in longitudinal clinical trials. *Stat Med*;11:2043–61.
- Hilden J, 1987. Reporting clinical trials from the viewpoint of a patient’s choice of treatment. *Stat Med*;6:745–52.
- Hillis TK, Guyatt GH, Oldridge N, Crowe J, Willan A, Griffith L, et al., 1994. Quality-of-life after myocardial infarction. *J Clin Epidemiol*;47:1287–96.
- Hillis A, Maguire M, Hawkins BS, Newhouse MM, 1986. The Markov process as a general method or non-parametric analysis of right-censored medical data. *J Chronic Dis*;39:595–604.
- Hillner BE, Desch CE, Carlson RW, Smith TJ, Esserman L, Bear HD, 1996. Trade-offs between survival and breast preservation for 3 initial treatments of ductal carcinoma-in-situ of the breast. *J Clin Oncol*;14:70–7.
- Hillner BE, Smith TJ, 1991. Efficacy and cost-effectiveness of adjuvant chemotherapy in women with node–negative breast cancer: a decision analysis model. *N Engl J Med*;324:160–8.
- Hillner BE, Smith TJ, 1992. Should women with node-negative breast cancer receive adjuvant chemotherapy: insights from a decision analysis model. *Breast Cancer Res Treat*;23:17–27.
- Hillner BE, Smith TJ, Desch CE, 1992. Efficacy and cost-effectiveness of autologous bone-marrow transplantation in metastatic breast cancer: estimates using decision analysis while awaiting clinical trial results. *JAMA*;267:2055–61.
- Hillner BE, Smith TJ, Desch CE, 1993. Assessing the cost-effectiveness of adjuvant therapies in early breast cancer using a decision analysis model. *Breast Cancer Res Treat*;25:97–105.
- Hinkley DV, 1988. Bootstrap methods. *J R Stat Soc B*;50:321–37.
- Hioki A, Iguchi T, 1994. Factors affecting the incidence of death at home of residents of Hida district, Japan. *J Clin Epidemiol*;47:1253–8.
- Hjorth M, Holmberg E, Rodger S, Taube A, Westin J, 1996. Physicians attitudes toward clinical trials and their relationship to patient accrual in a Nordic multicenter study on myeloma. *Control Clin Trials*;17:372–86.
- Hoehn JP, Fishelson G, 1987. Quality-adjusted prices and weak complementarity: a new method for valuing environmental services. *Am J Agric Econ*;69:1092.
- Hoehn JP, Fishelson G, 1988. Quality adjusted prices and weak complementarity: a new method for estimating the demand for environmental services. *Resources Energy*;10:337–54.
- Hofman A, Grobbee DE, de Jong PTVM, van den Ouweland FA, 1991. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol*;7:403–22.
- Hogenhuis W, Stevens SK, Wang P, Wong JB, Manolis AS, Estes NAM, et al., 1993. Cost-effectiveness of radio-frequency ablation compared with other strategies in Wolff–Parkinson–White syndrome. *Circulation*;88:437–46.
- Holland JC, 1991. Progress and challenges in psychosocial and behavioral research in cancer in the 20th century. *Cancer*;67:767–73.

- Hollen PJ, Gralla RJ, Cox C, Eberly SW, Kris MG, 1997. A dilemma in analysis: issues in the serial measurement of quality of life in patients with advanced lung cancer. *Lung Cancer*;18:119–36.
- Hollingworth W, MacKenzie R, Todd CJ, Dixon AK, 1995. Measuring changes in quality-of-life following magnetic resonance imaging of the knee SF-36, EuroQol(C) or Rosser Index. *Qual Life Res*;4:325–34.
- Holme I, Ekelund LG, Hjermann I, Leren P, 1994. Quality-adjusted meta analysis of the hypertension/coronary dilemma. *Am J Hypertens*;7:703–12.
- Holmes AM, 1995. A QALY-based societal health statistic for Canada, 1985. *Soc Sci Med*;41:1417–27.
- Holtgrave DR, Kelly JA, 1996. Preventing HIV/AIDS among high-risk urban women: the cost-effectiveness of a behavioral group intervention. *Am J Public Health*;86:1442–5.
- Holtgrave DR, Qualls NL, 1995. Threshold analysis and programs for prevention of HIV infection. *Med Decis Making*;15:311–17.
- Hoover DR, Peng Y, Saah AJ, Detels RR, Day RS, Phair JP, 1996. Using multiple decrement models to estimate risk and morbidity from specific AIDS illnesses. *Stat Med*;15:2307–21.
- Hope T, 1996. QALYs, lotteries and veils: the story so far. *J Med Ethics*;22:195–6.
- Hope T, Sprigings D, Crisp R, 1993. Not clinically indicated: patients interests or resource allocation. *BMJ*;306:379–81.
- Hopwood P, Stephens RJ, Machin D, 1994. Approaches to the analysis of quality of life data: experiences gained from a Medical Research Council lung cancer working party palliative chemotherapy trial. *Qual Life Res*;3:339–52.
- Hornberger JC, 1993. The hemodialysis prescription and cost-effectiveness. *J Am Soc Nephrol*;4:1021–7.
- Hornberger JC, 1993. The hemodialysis prescription and quality-adjusted life expectancy. *J Am Soc Nephrol*;4:1004–20.
- Hornberger JC, Redelmeier DA, Petersen J, 1992. Variability among methods to assess patients well-being and consequent effect on a cost-effectiveness analysis. *J Clin Epidemiol*;45:505–12.
- Hsieh FY, Crowley J, Tormey DC, 1983. Some test statistics for use in multistate survival analysis. *Biometrika*;70:111–19.
- Hubert HB, Fries JF, 1994. Predictors of physical disability after age 50: six-year longitudinal study in a runners club and a university population. *Ann Epidemiol*;4:285–94.
- Hulten CR, 1992. Growth accounting when technical change is embodied in capital. *Am Econ Rev*;82:964–80.
- Hunink MGM, Wong JB, Donaldson MC, Meyerovitz MF, de Vries J, Harrington DP, 1995. Revascularization for femoropopliteal disease: a decision and cost-effectiveness analysis. *JAMA*;274:165–71.
- Hunt LC, Lynk EL, 1995. Privatisation and efficiency in the UK water industry: an empirical analysis. *Oxf Bull Econ Stat*;57:371–88.
- Hurley S, 1990. A review of cost-effectiveness analyses. *Med J Aust*;153:S20–3.
- Hurley SF, Kaldor JM, 1992. The benefits and risks of mammographic screening for breast cancer. *Epidemiol Rev*;14:101–30.
- Hurny C, Bernhard J, Coates AS, Castiglione-Gertsch M, Peterson HF, Gelber RD, *et al.*, 1996. Impact of adjuvant therapy on quality of life in women with node-positive operable breast cancer. *Lancet*;347:1279–84.
- Hurteloup P, Mauriac L, Mathieu A, Montcuquet Ph, Ferriere P, Bastit P, 1987. Phase III chemotherapy in advanced breast cancer. 5-Fluorouracil + doxorubicin (adriamycin) + cyclophosphamide vs. 5-fluorouracil + epirubicin + cyclophosphamide. Preliminary results. *Clin Trials J*;24:123–30.
- Hutchings HC, 1987. Obtaining quality-of-life and cost-effectiveness data by the use of lay outcome assessors: a report from a multicenter clinical trial. *Control Clin Trials*;8:295.
- Hutton J, Brown R, Borowitz M, Abrams K, Rothman M, Shakespeare A, 1996. A new decision model for cost-utility comparisons of chemotherapy in recurrent metastatic breast cancer. *PharmacoEconomics*;9:8–22.
- Hwang JS, Tsao JY, Wang JD, 1996. Estimation of expected quality-adjusted survival by cross-sectional survey. *Stat Med*;15:93–102.
- Irvine EJ, Feagan BG, Wong CJ, 1996. Does self administration of a quality-of-life index for inflammatory bowel disease change the results. *J Clin Epidemiol*;49:1177–85.
- Islam MA, 1994. An alternative to the product limit method: the multistate approach. *Statistician*;43:237–45.
- Islam MA, 1994. Multistate survival models for transitions and reverse transitions: an application to contraceptive use data. *J R Stat Soc A*;157:441–55.
- Islam MA, Sing KP, 1992. Multistate survival models for partially censored data. *Environmetrics*;3:223–34.
- Iwatsubo Y, Derriennic F, Cassou B, Poitrenaud J, 1996. Predictors of life satisfaction amongst retired people in Paris. *Int J Epidemiol*;25:160–70.
- Jacobson JJ, Schweitzer SO, Kowalski CJ, 1991. Chemoprophylaxis of prosthetic joint patients during dental treatment: a decision utility analysis. *Oral Surg Oral Med Oral Pathol*;72:167–77.
- Jaeschke R, Guyatt GH, Keller J, Singer J, 1991. Interpreting changes in quality-of-life score in n of 1 randomized trials. *Control Clin Trials*;12:S226–33.

- Jaeschke R, Singer J, Guyatt GH, 1989. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials*;10:407–15.
- Jaeschke R, Singer J, Guyatt GH, 1990. A comparison of seven-point and visual analogue scales. Data from a randomized trial. *Control Clin Trials*;11:43–51.
- James KW, 1988. The morality of the QALY. *BMJ*;296:213
- Javitt J, Cas RD, Chiang YP, 1993. Cost-effectiveness of screening and cryotherapy for threshold retinopathy of prematurity. *Pediatrics*;91:859–66.
- Javitt JC, Aiello LP, 1996. Cost-effectiveness of detecting and treating diabetic retinopathy. *Ann Intern Med*;124:164–9.
- Jayes RL Jr, Levey AS, 1989. Donor specific transfusions or cyclosporine for related donor kidney transplantation? *Kidney Int*;36:485–96.
- Jayne TS, Khatri Y, Thirtle C, Reardon T, 1994. Determinants of productivity change using a profit function: smallholder agriculture in Zimbabwe. *Am J Agric Econ*;76:613–18.
- Jefferson T, Mugford M, Demicheli V, 1994. QALY league tables. *Health Econ*;3:205
- Jenicek M, Cleroux R, Lamoureux M, 1979. Principal component analysis of four health indicators and construction of a global health index in the aged. *Am J Epidemiol*;110:343–9.
- Jenkinson C, Layte R, Coulter A, Wright L, 1996. Evidence for the sensitivity of the SF-36 health-status measure to inequalities in health: results from the Oxford healthy life-styles survey. *J Epidemiol Community Health*;50:377–80.
- Jenkinson C, Peto V, Coulter A, 1994. Measuring change over time: a comparison of results from a global single item of health status and the multi-dimensional SF-36 health status survey questionnaire in patients presenting with menorrhagia. *Qual Life Res*;3:317–21.
- Jennett B, 1987. Technology assessment for countries at different stages of development. *Health Policy*;8:69–75.
- Jennett B, 1988. Balancing benefits and burdens in surgery. *Br Med Bull*;44:499–513.
- Johannesson M, 1994. QALYs, HYE and individual preferences: a graphical illustration. *Soc Sci Med*;39:1623–32.
- Johannesson M, 1995. Quality-adjusted life-years versus healthy-years equivalents – a comment. *J Health Econ*;14:9–16.
- Johannesson M, 1995. QALYs – a comment. *J Publ Econ*;56:327–8.
- Johannesson M, 1995. The ranking properties of healthy-years equivalents and quality-adjusted life-years under certainty and uncertainty. *Int J Technol Assess Health Care*;11:40–8.
- Johannesson M, Gerdtham UG, 1996. A note on the estimation of the equity efficiency trade-off for QALYs. *J Health Econ*;15:359–68.
- Johannesson M, Jonsson B, 1993. Economic-evaluation of osteoporosis prevention. *Health Policy*;24:103–24.
- Johannesson M, Jonsson B, Karlsson G, 1996. Outcome measurement in economic evaluation. *Health Econ*;5:279–96.
- Johannesson M, Pliskin JS, Weinstein MC, 1993. Are healthy-years equivalents an improvement over quality-adjusted life years. *Med Decis Making*;13:281–6.
- Johannesson M, Pliskin JS, Weinstein MC, 1994. A note on QALYs, time tradeoff, and discounting. *Med Decis Making*;14:188–93.
- Johannesson M, Weinstein MC, 1993. On the decision rules of cost-effectiveness analysis. *J Health Econ*;12:459–67.
- Johnston K, Buxton MJ, Jones DR, Fitzpatrick R, 1999. Assessing the costs of healthcare technologies in clinical trials. *Health Technol Assess*;3(6).
- Jones R, 1988. What price QALYs? *J R Coll Gen Pract*;38:231.
- Jonsson B, 1991. Quality of life: economic aspects. *Scand J Prim Health Care*;9:93–6.
- Jonsson B, Christiansen C, Johnell O, Hedbrandt J, 1995. Cost-effectiveness of fracture prevention in established osteoporosis. *Osteoporos Int*;5:136–42.
- Jonsson B, Christiansen C, Johnell O, Hedbrandt J, 1996. Cost-effectiveness of fracture prevention in established osteoporosis. *Scand J Rheumatol*;25 suppl:30–8.
- Jonsson B, Karlsson G, 1993. Outcome measurement in economic evaluations of growth factors. *Eur J Cancer*;29A:S19–22.
- Jordan JE, Marks MP, Lane B, Steinberg GK, 1996. Cost-effectiveness of endovascular therapy in the surgical management of cerebral arteriovenous malformations. *Am J Neuroradiol*;17:247–54.
- Juniper EF, Guyatt GH, Willan A, Griffith LE, 1994. Determining a minimal important change in a disease-specific quality-of-life questionnaire. *J Clin Epidemiol*;47:81–7.
- Kagan AR, Hintz BL, 1979. Is there a best management for localized adenocarcinoma of the prostate? *Cancer Clin Trials*;2:359–63.
- Kalbfleisch JD, Lawless JF, 1988. Likelihood analysis of multi-state models for disease incidence and mortality. *Stat Med*;7:149–60.
- Kalbfleisch JD, Lawless JF, Vollmer WM, 1983. Estimation in Markov models from aggregate data. *Biometrics*;39:907–19.
- Kalbfleisch JD, Prentice RL, 1980. The statistical analysis of failure time data. New York: John Wiley.



- Kalish SC, Gurwitz JH, Krumholz HM, Avorn J, 1995. A cost-effectiveness model of thrombolytic therapy for acute myocardial infarction. *J Gen Intern Med*;10:321–30.
- Kamlet MS, Paul N, Greenhouse J, Kupfer D, Frank E, Wade M, 1995. Cost–utility analysis of maintenance treatment for recurrent depression. *Control Clin Trials*;16:17–40.
- Kannel WB, Vokonas PS, 1992. Demographics of the prevalence, incidence, and management of coronary heart disease in the elderly and in women. *Ann Epidemiol*;2:5–14.
- Kantoff PW, Block C, Letvak L, George M, 1993. 14-day continuous infusion of mitoxantrone in hormone–refractory metastatic adenocarcinoma of the prostate. *Am J Clin Oncol: Cancer Clin Trials*;16:489–91.
- Kaplan EL, Meier P, 1958. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*;53:457–81.
- Kaplan RM, 1993. Quality of life assessment for cost/utility studies in cancer. *Cancer Treat Rev*;19:85–96.
- Kaplan RM, 1994. Value judgment in the Oregon medicaid experiment. *Med Care*;32:975–88.
- Kaplan RM, 1994. Using quality-of-life information to set priorities in health policy. *Soc Indicators Res*;33:121–63.
- Kaplan RM, Feeny D, Revicki DA, 1993. Methods for assessing relative importance in preference based outcome measures. *Qual Life Res*;2:467–75.
- Kaplan RM, Ganiats TG, 1990. QALYs: their ethical implications. *JAMA*;264:2502–3.
- Kaplan RM, McCutchan JA, Navarro AM, Anderson JP, Atkinson JH, Chandler J, *et al.*, 1994. Quality-adjusted survival analysis: a neglected application of the Quality of Well-being scale. *Psychol Health*;9:131–41.
- Karanci AN, 1995. Caregivers of Turkish schizophrenic patients: causal attributions, burdens and attitudes to help from the health professionals. *Soc Psychiatry Psychiat Epidemiol*;30:261–8.
- Karnofsky D, Burchenal JH, 1949. Clinical evaluation of chemotherapeutic agents in cancer. In: MacLeod CM, editor. Evaluation of chemotherapeutic agents. New York: Columbia University Press.
- Kato H, 1996. History of photodynamic therapy past, present and future. *Jpn J Cancer Chemother*;23:8–15.
- Kattan MW, Inoue Y, Giles FJ, Talpaz M, Ozer H, Guilhot F, *et al.*, 1996. Cost-effectiveness of interferon-alpha and conventional chemotherapy in chronic myelogenous leukemia. *Ann Intern Med*;125:541.
- Katz DA, Cronenwett JL, 1994. The cost-effectiveness of early surgery versus watchful waiting in the management of small abdominal aortic aneurysms. *J Vasc Surg*;19:980–91.
- Katz DA, Littenberg B, Cronenwett JL, 1992. Management of small abdominal aortic aneurysms: early surgery vs watchful waiting. *JAMA*;268:2678–86.
- Kawachi I, 1989. QALYs and justice. *Health Policy*;13:115–20.
- Kawachi I, Bethwaite P, Bethwaite J, 1990. The use of quality-adjusted life years (QALYs) in the economic appraisal of health care. *N Z Med J*;103:46–8.
- Kawachi I, Malcolm LA, 1989. The benefits of treating mild to moderate hypertension. A quantitative estimation of the live expectancy gains from pharmacological reduction of blood pressure. *J Clin Epidemiol*;42:905–12.
- Kawachi I, Malcolm LA, 1990. Treating mild to moderate hypertension: cost-effectiveness and policy implications. *J Cardiovasc Pharmacol*;16:S126–8.
- Kawachi I, Malcolm LA, 1991. The cost-effectiveness of treating mild-to-moderate hypertension: a reappraisal. *J Hypertens*;9:199–208.
- Kawachi I, Purdie G, 1990. Should treatment of mild to moderate hypertension be targeted? Results from a Markov cohort model incorporating multiple risk factors. *J Hum Hypertens*;4:651–8.
- Kay L, Jorgensen T, Schultzlarsen K, 1992. Abdominal pain in a 70-year-old Danish population – an epidemiologic study of the prevalence and importance of abdominal pain. *J Clin Epidemiol*;45:1377–82.
- Kay R, 1982. The analysis of transition times in multistate stochastic processes using proportional hazard regression models. *Comm Stat Theory Methodol*;11:1743–56.
- Kay R, 1984. Multistate survival analysis: an application in breast cancer. *Methods Inf Med*;23:157–62.
- Kay R, 1986. A Markov model for analysing cancer markers and disease states in survival studies. *Biometrics*;42:855–65.
- Keeler E, Bell R, 1992. New DEALs: other approximations of life expectancy. *Med Decis Making*;12:307–11.
- Kellett J, Clarke J, 1995. Comparison of accelerated tissue plasminogen activator with streptokinase for treatment of suspected myocardial infarction. *Med Decis Making*;15:297–310.
- Kellett J, Graham J, 1995. Referral for coronary angiography after exercise testing: traditional decision making versus decision analysis. *QJM*;88:401–8.
- Kellett JG, O’Riordan J, 1993. Whither the rationale for thrombolytic agent administration: a retrospective review of traditional intuitive decision making using a decision analysis model. *Ir J Med Sci*;162:133–9.
- Kent KC, Kuntz KM, Patel MR, Kim D, Klufas RA, Whittemore AD, *et al.*, 1995. Perioperative imaging strategies for carotid endarterectomy: an analysis of morbidity and cost-effectiveness in symptomatic patients. *JAMA*;274:888–93.
- Kerridge RK, Glasziou PP, Hillman KM, 1995. The use of quality-adjusted life years (QALYs) to evaluate treatment in intensive care. *Anaesth Intensive Care*;23:322–31.

- Kerschbaum T, Haastert B, Marinello CP, 1996. Risk of debonding in 3-unit resin bonded fixed partial dentures. *J Prosthet Dent*;75:248–53.
- Kianifard F, 1994. Evaluation of clinimetric scales: basic principles and methods. *Statistician*;43:475–82.
- Kiberd BA, 1994. Should hepatitis C infected kidneys be transplanted in the United States. *Transplantation*;57:1068–72.
- Kiberd BA, Jindal KK, 1995. Screening to prevent renal failure in insulin-dependent diabetic patients: an economic evaluation. *BMJ*;311:1595–9.
- Kiebert GM, Stiggelbout AM, Kievit J, Leer JWH, van de Velde CJH, de Haes HJCJM, 1994. Choices in oncology: factors that influence patients treatment preference. *Qual Life Res*;3:175–82.
- Kievit J, Bruinvels DJ, 1995. Detection of recurrence after surgery for colorectal cancer. *Eur J Cancer*;31A:1222–5.
- Kievit J, van de Velde CJH, 1990. Utility and cost of carcinoembryonic antigen monitoring in colon cancer follow-up evaluation. A Markov analysis. *Cancer*;65:2580–7.
- Kind P, Gudex CM, 1994. Measuring health status in the community: a comparison of methods. *J Epidemiol Community Health*;48:86–91.
- King JT, Glick HA, Mason TJ, Flamm ES, 1995. Elective surgery for asymptomatic, unruptured, intracranial aneurysms: a cost-effectiveness analysis. *J Neurosurg*;83:403–12.
- King MT, Dobson AJ, Harnett PR, 1996. A comparison of 2 quality-of-life questionnaires for cancer clinical trials: the functional living index-cancer (FLIC) the quality-of-life questionnaire core module (QLQ-C30). *J Clin Epidemiol*;49:21–9.
- King RB, 1996. Quality-of-life after stroke. *Stroke*;27:1467–72.
- Kitchen WH, Bowman E, Callanan C, Campbell NT, Carse EA, Charlton M, *et al.*, 1993. The cost of improving the outcome for infants of birth weight 500–999 g in Victoria. *J Paediatr Child Health*;29:56–62.
- Klaukka T, Peura S, Martikainen J, 1991. Why has the utilization of antiasthmatics increased in Finland? *J Clin Epidemiol*;44:859–63.
- Klein JP, Keiding N, Copelan EA, 1993. Plotting summary predictions in multistate survival models: probabilities of relapse and death in remission for bone-marrow transplantation patients. *Stat Med*;12:2315–32.
- Kliemt H, 1995. Life: what is worth maintaining? *Cardiovasc Risk Factors*;5:249–54.
- Klioze SS, Miller MF, Spiro, TP, 1988. A randomized, comparative study of buserelin with DES/orchiectomy in the treatment of stage D2 prostatic cancer patients. *Am J Clin Oncol: Cancer Clin Trials*;11:S176–82.
- Knapp MRJ, 1996. Cost-effectiveness in the treatment of schizophrenia. *Eur Psychiat*;11:137–43.
- Koch GG, Landis JR, Freeman JL, Freeman DH, Lehnen RG, 1977. A general methodology for the analysis of experiments with repeated measurement of categorical data. *Biometrics*;33:133–58.
- Korn EL, 1993. On estimating the distribution function for quality-of-life in cancer clinical trials. *Biometrika*;80:535–42.
- Korn EL, 1997. Assessing quality of life in comparative cancer clinical trials: does it make a difference? *Control Clin Trials*;18:275–6.
- Krahn M, Naylor CD, Basinski AS, Detsky AS, 1991. Comparison of an aggressive (United States) and a less aggressive (Canadian) policy for cholesterol screening and treatment. *Ann Intern Med*;115:248–55.
- Krahn MD, Mahoney JE, Eckman MH, Trachtenberg J, Pauker SG, Detsky AS, 1994. Screening for prostate cancer: a decision-analytic view. *JAMA*;272:773–80.
- Kramer BS, editor, 1996. Quality of life in cancer clinical trials: proceedings of a workshop held at the National Institutes of Health, Bethesda, Maryland, March 1–2 1995. *J Natl Cancer Inst*;20.
- Kristiansen IS, Eggen AE, Thelle DS, 1991. Cost-effectiveness of incremental programs for lowering serum cholesterol concentration: is individual intervention worth while. *BMJ*;302:1119–22.
- Kuate Defo B, 1995. Epidemiology and control of infant and early childhood malaria: a competing risks analysis. *Int J Epidemiol*;24:204–17.
- Kuntz KM, Kent KC, 1996. Is carotid endarterectomy cost-effective? An analysis of symptomatic and asymptomatic patients. *Circulation*;94:194–8.
- Kuntz KM, Tsevat J, Goldman L, Weinstein MC, 1996. Cost-effectiveness of routine coronary angiography after acute myocardial infarction. *Circulation*;94:957–65.
- Kuntz KM, Weinstein MC, 1995. Life expectancy biases in clinical decision modelling. *Med Decis Making*;15:158–69.
- Kusek JW, Agodoa LY, Glasscock R, Wright JT, Kirk KA, Lee JY, 1996. The African American Study of Kidney Disease and Hypertension (AASK) Pilot Study: overview. *Control Clin Trials*;17:S1.
- Kusek JW, Lee JY, Smith DE, Milligan S, Faulkner M, Cornell CE, *et al.*, 1996. Effect of blood pressure control and antihypertensive drug regimen on quality-of-life: The African American Study of Kidney Disease and Hypertension (AASK) Pilot Study. *Control Clin Trials*;17:S40–6.
- Kuskey KP, 1980. Health, cost, and hospital services: a new approach to decision making. [Journal not specified by BIOS].
- Labelle RJ, Hurley JE, 1992. Implications of basing health care resource allocations on cost-utility analysis in the presence of externalities. *J Health Econ*;11:259–77.
- Lachs MS, 1993. The more things change. *J Clin Epidemiol*;46:1091–2.

- Lagakos SW, 1976. A stochastic model for censored survival data in the presence of an auxiliary variable. *Biometrics*;32:551–9.
- Lagakos SW, Sommer CJ, Zelen M, 1978. Semi-Markov models for partially censored data. *Biometrika*;65:311–17.
- Laird NM, 1988. Missing data in longitudinal studies. *Stat Med*;7:305–15.
- Laird NM, Ware JH, 1982. Random-effects models for longitudinal data. *Biometrics*;38:963–74.
- Lambert CM, Hurst NP, 1995. Health economics as an aspect of health outcome: basic principles and application in rheumatoid arthritis. *Br J Rheumatol*;34:774–80.
- Landfeld CS, Phillips RS, Bergner M, 1990. Patient characteristics in support: functional status. *J Clin Epidemiol*;43:37–9S.
- Landis JR, Miller ME, Davis CS, Koch GG, 1988. Some general methods for the analysis of categorical data in longitudinal studies. *Stat Med*;7:109–37.
- La Puma J, 1990. QALYs: their ethical implications – reply. *JAMA*;264:2503–4.
- La Puma J, 1992. Quality-adjusted life years: ethical implications and the Oregon plan. *Issues Law Med*;7:429–41.
- La Puma J, Lawlor EF, 1990. Quality-adjusted life-years: ethical implications for physicians and policy makers. *JAMA*;263:2917–21.
- Launois R, 1994. Integrating individual preferences into therapeutic and collective choices. *Rev Epidemiol Sante Publique*;42:246–62.
- Launois R, Henry B, Marty JR, Gersberg M, Lassale C, Benoist M, *et al.*, 1994. Chemonucleolysis versus surgical discectomy for sciatica secondary to lumbar disc herniation: a cost and quality-of-life evaluation. *PharmacoEconomics*;6:453–63.
- Launois R, Orvain J, Ounis I, 1992. The interest of utility measurement: recurrent respiratory diseases. *Rev Epidemiol Sante Publique*;40:46–55.
- Launois R, Reboulmarty J, Henry B, Bonnetterre J, 1996. A cost–utility analysis of 2nd line chemotherapy in metastatic breast cancer: docetaxel versus paclitaxel versus vinorelbine. *PharmacoEconomics*;10:504–21.
- Laupacis A, Bourne R, Rorabeck C, Feeny D, Wong C, Tugwell P, *et al.*, 1994. Costs of elective total hip arthroplasty during the first year. *J Arthroplasty*;9:481–7.
- Laupacis A, Feeny D, Detsky AS, Tugwell PX, 1992. How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations. *Can Med Assoc J*;146:473–81.
- Laupacis A, Wong C, Churchill D, 1991. The use of generic and specific quality-of-life measures in hemodialysis patients treated with erythropoietin. *Control Clin Trials*;12:S168–79.
- Lavori PW, Dawson R, Shera D, 1995. A multiple imputation strategy for clinical trials with truncation of patient data. *Stat Med*;14:1913–25.
- Lawler FH, Bissonni RS, Holtgrave DR, 1991. Circumcision: a decision analysis of its medical value. *Fam Med*;23:587–93.
- Lawrence WF, Fryback DG, Martin PA, Klein R, Klein BEK, 1996. Health status and hypertension: a population-based study. *J Clin Epidemiol*;49:1239–45.
- Lawrence WF, Grist TM, Brazy PC, Fryback DG, 1995. Magnetic resonance angiography in progressive renal failure: a technology assessment. *Am J Kidney Dis*;25:701–9.
- Lazzaroni F, Bonassi S, Magnani M, Calvi A, Repetto E, Serra G, *et al.*, 1993. Moderate maternal drinking and outcome of pregnancy. *Eur J Epidemiol*;9:599–606.
- Lea AR, Hailey DM, 1995. The cochlear implant: a technology for the profoundly deaf. *Med Prog Technol*;21:47–52.
- Lee JY, Oberman A, Fletcher GF, Raczynski JM, Fletcher BJ, Nanda NC, *et al.*, 1994. Design of the training levels comparison trial. *Control Clin Trials*;15:59–76.
- Lee KH, Angus DC, Abramson NS, 1996. Cardio-pulmonary resuscitation: what cost to cheat death. *Crit Care Med*;24:2046–52.
- Le Floch JP, Charles MA, Philippon C, Perlemuter L, 1994. Cost-effectiveness of screening for microalbuminuria using immunochemical dipstick tests or laboratory assays in diabetic patients. *Diabet Med*;11:349–56.
- Lehman AF, 1996. Measures of quality of life among persons with severe and persistent mental disorders. *Soc Psychiatry Psychiat Epidemiol*;31:78–88.
- Lenderking WR, 1995. Quality of life evaluation of zidovudine in asymptomatic HIV infection: a Q-TWiST analysis. *Biomed Pharmacother*;49:431.
- Lenderking WR, Gelber RD, Cotton DJ, Cole BF, Goldhirsch A, Volberding PA, *et al.*, 1994. Evaluation of the quality-of-life associated with zidovudine treatment in asymptomatic Human Immunodeficiency Virus infection. *N Engl J Med*;330:738–43.
- Leon FE, Meneghini F, Rocca WA, Anderson DW, 1993. Instruments for neuroepidemiological surveys. *J Clin Epidemiol*;46:313–14.
- Leplege A, Mesbah M, Marquis P, 1995. Preliminary psychometric analysis of the French version of an international quality-of-life questionnaire: the MOS SF-36 (version 1.1). *Rev Epidemiol Sante Publique*;43:371–9.
- Lesaffre E, Scheys I, Frohlich J, Bluhmki E, 1993. Calculation of power and sample size with bounded outcome scores. *Stat Med*;12:1063–78.

- Levin LA, Jonsson B, 1992. Cost-effectiveness of thrombolysis: a randomized study of intravenous rt-PA in suspected myocardial infarction. *Eur Heart J*;13:2–8.
- Leviton A, Schulman J, Kammerman L, Porter D, Slack W, Graham JR, 1980. A probability model of headache recurrence. *J Chronic Dis*;33:407–12.
- Lewis JM, 1996. Improving dental health status indicators for evaluation. *Community Dent Oral Epidemiol*;24:32–6.
- Liang J, Edward XL, Whitelaw N, 1996. Probabilities and lifetime durations of short stay hospital and nursing home use in the United States, 1985. *Med Care*;34:1018–36.
- Liang K, Zeger SL, 1986. Longitudinal data analysis using generalised linear models. *Biometrika*;73:13–22.
- Liang W, Brown MB, 1997. A median based test under informative dropout: the one sample case. *Control Clin Trials*;18:445–59.
- Lim LLY, Valenti LA, Knapp JC, Dobson AJ, Plotnikoff R, Higginbotham N, *et al.*, 1993. A self-administered quality-of-life questionnaire after acute myocardial infarction. *J Clin Epidemiol*;46:1249–56.
- Lind NC, 1993. A compound index of national development. *Soc Indicators Res*;28:267–84.
- Lind NC, 1995. Policy goals for health and safety. *Risk Anal*;15:639–44.
- Lindberg MS, Sedinger JS, Rexstad EA, 1995. Estimating nest site fidelity of adult female black brant with multi-state modeling and geographic information systems. *J Appl Stat*;22:725–35.
- Lindley CM, Hirsch JD, O'Neill V, Transau MC, Gilbert CS, Osterhaus JT, 1992. Quality of life consequences of chemotherapy-induced emesis. *Qual Life Res*;1:331–40.
- Lindsey JK, 1994. Models for repeated measurements. New York: Oxford University Press.
- Lindsey JK, 1997. Modelling longitudinal measurement dropouts as a survival process with time-varying covariates. Technical Report. Belgium: Limburgs Universitair Centrum.
- Lindstrom B, 1994. Quality-of-life for children and disabled children based on health as a resource concept. *J Epidemiol Community Health*;48:529–30.
- Little RJA, 1988. Analysis of data with missing values – commentary. *Stat Med*;7:347–55.
- Little RJA, 1995. Modelling the drop-out mechanism in repeated measures studies. *J Am Stat Assoc*;90:1112–21.
- Little RJA, Rubin DB, 1987. Statistical analysis with missing data. New York: John Wiley.
- Livartowski A, Boucher J, Detournay B, Reinert P, 1996. Cost-effectiveness evaluation of vaccination against haemophilus influenzae invasive diseases in France. *Vaccine*;14:495–500.
- Livartowski A, Pouillart P, 1992. Cost–utility analysis in oncology: application of adjuvant chemotherapy in the treatment of node positive premenopausal breast cancer. *Med Sci*;8:1073–8.
- Ljungqvist L, 1993. A unified approach to measures of privacy in randomized response models: a utilitarian perspective. *J Am Stat Assoc*;88:97–103.
- Locker D, 1995. Health outcomes of oral disorders. *Int J Epidemiol*;24:S85–9.
- Locker D, 1996. Work stress, job satisfaction and emotional well-being among Canadian dental assistants. *Community Dent Oral Epidemiol*;24:133–7.
- Loewy JW, Kapadia AS, Davis BR, 1992. Statistical methods that distinguish between attributes of assessment: prolongation of life versus quality of life. *Med Decis Making*;12:83–92.
- Longini IM, Byers RH, Hessel NA, Tan WY, 1992. Estimating the stage-specific numbers of HIV infection using a Markov model and back calculation. *Stat Med*;11:831–43.
- Longini IM, Clark WS, Byers RH, Ward JW, Darrow WW, Lemp GF, *et al.*, 1989. Statistical analysis of the stages of HIV infection using a Markov model. *Stat Med*;8:831–43.
- Loomes G, 1995. The myth of the HYE. *J Health Econ*;14:1–7.
- Loomes G, McKenzie L, 1989. The use of QALYs in health care decision making. *Soc Sci Med*;28:299–308.
- Louis TA, 1988. General methods for analysing repeated measures. *Stat Med*;7:29–45.
- Lu Y, Stitt FW, 1994. Using Markov processes to describe the prognosis of HIV-1 infection. *Med Decis Making*;14:266–72.
- Lubeck DP, Fries JF, 1992. Changes in quality of life among persons with HIV infection. *Qual Life Res*;1:359–66.
- Luh YH, Stefanou SE, 1991. Productivity growth in United States agriculture under dynamic adjustment. *Am J Agric Econ*;73:1116–25.
- Lydick E, Epstein RS, Himmelberger D, White CJ, 1995. Area under the curve: a metric for patient subjective responses in episodic diseases. *Qual Life Res*;4:41–5.
- Lyttkens CH, 1994. QALY league tables. *Health Econ*;3:57–8.
- Maasoumi E, Nickelsburg G, 1988. Multivariate measures of well-being and an analysis of inequality in the Michigan data. *J Business Econ Stat*;6:327–34.
- MacDonald LD, Anderson HR, 1984. Stigma in patients with rectal cancer: a community study. *J Epidemiol Community Health*;38:284–90.
- Macheledt JE, Vernon SW, 1992. Diabetes and disability among Mexican Americans: the effect of different measures of diabetes on its association with disability. *J Clin Epidemiol*;45:519–28.

- Machin D, 1997. Assessing quality of life in comparative cancer clinical trials: does it make a difference? *Control Clin Trials*;18:301–5.
- Mackenbach JP, van den Bos J, Joung IMA, van de Mheen H, Stronks K, 1994. The determinants of excellent health: different from the determinants of ill health. *Int J Epidemiol*;23:1273–81.
- MacNeill AB, Sutton MS, Poolewilson PA, Lubsen J, 1993. Doppler flow, echocardiography and functional improvement. Assessment of nisoldipine therapy (DEFIANT II study): objectives and design. *Clin Trials Meta-Anal*;28:267–80.
- Maetani S, Onodera H, Imamura M, 1994. How to incorporate quality of life into survival analysis to select the best treatment for cancer patients: colostomy or no colostomy for rectal cancer [abstract]. *Qual Life Res*;3:53.
- Maisey MN, Ellam SV, 1984. Investigating the adenocarcinoma of unknown origin (ACUP): a cost benefit analysis. *Rev Epidemiol Sante Publique*;32:57–61.
- Malan AF, Ryan E, van der Elst CW, Pelteret R, 1992. The cost of neonatal care. *S Afr Med J*;82:417–19.
- Malani HM, van Ryzin J, 1988. Comparison of two treatments in animal carcinogenicity experiments. *J Am Stat Assoc*;83:1171–7.
- Malone M, 1994. Quality-of-life of patients receiving home parenteral or enteral nutrition support. *PharmacoEconomics*;5:101–8.
- Malter AD, Larson EB, Urban N, Deyo RA, 1996. Cost-effectiveness of lumbar discectomy for the treatment of herniated intervertebral disc. *Spine*;21:1048–54.
- Mangtani P, Hall AJ, Normand CEM, 1995. Hepatitis B vaccination the cost-effectiveness of alternative strategies in England and Wales. *J Epidemiol Community Health*;49:238–44.
- Marks GB, Dunn SM, Woolcock AJ, 1992. A scale for the measurement of quality-of-life in adults with asthma. *J Clin Epidemiol*;45:461–72.
- Marks GB, Dunn SM, Woolcock AJ, 1993. An evaluation of an asthma quality-of-life questionnaire as a measure of change in adults with asthma. *J Clin Epidemiol*;46:1103–11.
- Marshall G, Guo WS, Jones RH, 1995. MARKOV: a computer program for multistate Markov models with covariables. *Comput Methods Programs Biomed*;47:147–56.
- Marshall G, Jones RH, 1995. Multistate models and diabetic retinopathy. *Stat Med*;14:1975–83.
- Maschewsky-Schneider U, Heinemann E, Mager A, Lang P, Lusebrink K, Scheuermann W, et al., 1993. Process evaluation measures and blood pressure results in the German Cardiovascular Prevention Study. *Ann Epidemiol*;3:S44–50.
- Mason JM, 1994. Cost-per-QALY league tables: their role in pharmacoeconomic analysis. *PharmacoEconomics*;5:472–81.
- Matthews JNS, 1993. A refinement to the analysis of serial data using summary measures. *Stat Med*;12:27–37.
- Matthews JNS, Altman DG, Campbell MJ, Royston P, 1990. Analysis of serial measurements in medical research. *BMJ*;300:230–5.
- Mauldin PD, Weintraub WS, 1993. Determining patient utility following coronary revascularization. *J Invas Cardiol*;5:171–8.
- Mauskopf JA, French MT, 1991. Estimating the value of avoiding morbidity and mortality from foodborne illnesses. *Risk Anal*;11:619–31.
- Maynard A, 1987. Incentives for cost-effective physician behaviour. *Health Policy*;7:189–204.
- McAlister D, 1994. Putting health economics into quality. *Publ Money Manage*;14:15–22.
- McClellan WM, Anson C, Birkeli K, Tuttle E, 1991. Functional status and quality-of-life: predictors of early mortality among patients entering treatment for end stage renal disease. *J Clin Epidemiol*;44:83–9.
- McDowell I, Praught E, 1982. On the measurement of happiness. An examination of the Bradburn scale in the Canada Health Survey. *Am J Epidemiol*;116:949–58.
- McGrady GA, Daling JR, Peterson DR, 1985. Maternal urinary tract infection and adverse fetal outcomes. *Am J Epidemiol*;121:377–81.
- McIntyre P, Hall J, Leeder S, 1994. An economic analysis of alternatives for childhood immunization against haemophilus influenzae type B disease. *Aust J Publ Health*;18:394–400.
- McKie J, Kuhse H, Richardson J, Singer P, 1996. Another peep behind the veil. *J Med Ethics*;22:216–21.
- McKie J, Kuhse H, Richardson J, Singer P, 1996. Double jeopardy, the equal value of lives and the veil of ignorance – a rejoinder. *J Med Ethics*;22:204–8.
- McTurk L, 1991. A methodological quibble about QALYs. *BMJ*;302:1601.
- Mehrez A, Gafni A, 1989. Quality-adjusted life years, utility theory, and healthy-years equivalents. *Med Decis Making*;9:142–9.
- Mehrez A, Gafni A, 1991. The healthy-years equivalents: how to measure them using the standard gamble approach. *Med Decis Making*;11:140–6.
- Mehrez A, Gafni A, 1993. Healthy-years equivalents versus quality-adjusted life years: in pursuit of progress. *Med Decis Making*;13:287–92.
- Mendeloff J, 1983. Measuring elusive benefits: on the value of health. *J Health Polit Policy Law*;8:554–80.
- Menzel PT, 1995. QALYs: maximization, distribution and consent – a response. *Health Care Anal*;3:226–9.
- Michel JP, Proust J, Rapin CH, Staehelin H, 1996. The geriatrician in 1996: a viewpoint from Switzerland. *Aging (Milano)*;8:135–40.

- Miller DK, Homan SM, 1994. Determining transition probabilities: confusion and suggestions. *Med Decis Making*;14:52–8.
- Miller P, Mulvey C, Martin N, 1996. Multiple regression analysis of the occupational status of twins: a comparison of economic and behavioral genetics models. *Oxf Bull Econ Stat*;58:227.
- Minet P, Bartsch P, Chevalier PH, Raets D, Gras A, Dejardinclonson MT, *et al.*, 1987. Quality of life of inoperable non-small cell lung carcinoma. A randomized phase II clinical study comparing radiotherapy alone and combined radio-chemotherapy. *Radiother Oncol*;8:217–30.
- Mitchell BD, Stern MP, Haffner SM, Hazuda HP, Patterson JK, 1990. Functional impairment in Mexican Americans and non-Hispanic whites with diabetes. *J Clin Epidemiol*;43:319–27.
- Miyamoto JM, Eraker SA, 1985. Parameter estimates for a QALY utility model. *Med Decis Making*;5:191–213.
- Moatti JP, Auquier P, le Coroller AG, Macquartmoulin G, 1995. QALYs or not QALYs: that is the question. *Rev Epidemiol Sante Publique*;43:573–83.
- Moatti JP, le Coroller AG, 1996. Some economists' thoughts about medical ethics. *J Econ Med*;14:67–78.
- Mocan HN, 1995. Quality-adjusted cost functions for child care centers. *Am Econ Rev*;85:409–13.
- Mohide EA, Torrance GW, Streiner DL, Pringle DM, Gilbert R, 1988. Measuring the wellbeing of family caregivers using the time trade-off technique. *J Clin Epidemiol*;41:475–82.
- Moinpour CM, 1997. Do quality of life assessments make a difference in the evaluation of cancer treatments? *Control Clin Trials*;18:311–17.
- Mold JW, Holtgrave DR, Bissonni RS, Marley DS, Wright RA, Spann SJ, 1992. The evaluation and treatment of men with asymptomatic prostate nodules in primary care: a decision analysis. *J Fam Pract*;34:561–8.
- Mooney G, 1989. QALYs: are they enough? A health economist's perspective. *J Med Ethics*;15:148–52.
- Mor V, Guadagnoli E, 1988. Quality of life measurement: a psychometric tower of Babel. *J Clin Epidemiol*;41:1055–8.
- Moret L, Chwalow J, Baudoinballeur C, 1993. Evaluating quality-of-life: construction of a scale. *Rev Epidemiol Sante Publique*;41:65–71.
- Mori M, Woodworth GG, Woolson RF, 1992. Application of empirical Bayes inference to estimation of rate of change in the presence of informative right censoring. *Stat Med*;11:621–31.
- Morris JN, Suissa S, Sherwood S, Wright SM, Greer D, 1986. Last days: a study of the quality of life of terminally ill cancer patients. *J Chronic Dis*;39:47–62.
- Morrow RH, Bryant JH, 1995. Health policy approaches to measuring and valuing human life: conceptual and ethical issues. *Am J Public Health*;85:1356–60.
- MRC Lung Cancer Working Party, 1991. Inoperable non-small cell lung cancer (NSCLC): a Medical Research Council randomised trial of palliative radiotherapy with two fractions or ten fractions. *Br J Cancer*;63:265–70.
- MRC Lung Cancer Working Party, 1991. Assessment of quality of life in small cell lung cancer using a daily diary card developed by the Medical Research Council Lung Cancer Working Party. *Br J Cancer*;64:299–306.
- MRC Lung Cancer Working Party, 1992. A Medical Research Council (MRC) randomised trial of palliative radiotherapy with two fractions or a single fraction in patients with inoperable non-small cell lung cancer (NSCLC), poor performance status. *Br J Cancer*;65:934–41.
- MRC Lung Cancer Working Party, 1995. Symptoms at presentation for treatment in patients with lung cancer: implications for the evaluation of palliative treatment. *Br J Cancer*;71:633–6.
- MRC Lung Cancer Working Party, 1996. Randomized trial of palliative two-fraction versus more intensive 13-fraction radiotherapy for patients with inoperable non-small cell lung cancer and good performance status. *Clin Oncol*;8:167–75.
- Murphy DJ, Cluff LE, 1990. Introduction: the support study. *J Clin Epidemiol*;43:V–VIII.
- Murray JR, 1995. Do you CEA the need for CMAs or have a CUA about extra QALY ratios? *J Dermatol Treat*;6:3–4.
- Murray KJ, Nelson DF, Isaacson S, Scott C, Fischbach JA, Porter A, *et al.*, 1993. Quality-adjusted survival analysis of malignant glioma patients treated with twice daily radiation (RT) carmustine: a report of the Radiation Therapy Oncology Group (RTOG) 83-02. *Int J Radiat Oncol Biol Phys*;27:207.
- Murray KJ, Nelson DF, Scott C, Fischbach AJ, Porter A, Farnan N, *et al.*, 1995. Quality-adjusted survival analysis of malignant glioma, patients treated with twice daily radiation (RT) carmustine: a report of the Radiation Therapy Oncology Group (RTOG) 83-02. *Int J Radiat Oncol Biol Phys*;31:453–9.
- Mushlin AI, Mooney C, Grow V, Phelps CE, 1994. The value of diagnostic information to patients with suspected multiple sclerosis. *Arch Neurol*;51:67–72.
- Myers LE, Paulson DF, Berry WR, Cox EB, Laszlo J, Stanley W, 1980. A time-dependent statistical model which relates current clinical status to prognosis: application to advanced prostatic cancer. *J Chronic Dis*;33:491–9.
- Myerson RJ, Shapiro SJ, Lacey D, Lopez M, Birnbaum E, Fleshman J, *et al.*, 1995. Carcinoma of the anal canal. *Am J Clin Oncol: Cancer Clin Trials*;18:32–9.
- Nabholtz JM, Gelmon K, Bontenbal M, Spielmann M, Catimel G, Conte P, *et al.*, 1996. Multicenter, randomized comparative study of 2 doses of paclitaxel in patients with metastatic breast cancer. *J Clin Oncol*;14:1858–67.

- Naglie IG, Detsky AS, 1992. Treatment of chronic nonvalvular atrial fibrillation in the elderly: a decision analysis. *Med Decis Making*;12:239–49.
- Naji SA, Sheldon TA, 1993. The measurement of patient outcomes. *Proc R Soc Edinb B Biol Sci*;101:55–75.
- Nayfield SG, Ganz PA, Moinpour CM, Cella DF, Hailey BJ, 1992. Report from a National Cancer Institute (USA) workshop on quality of life assessment in cancer clinical trials. *Qual Life Res*;1:203–10.
- Nease RF Jr, Tsai R, Hynes LM, Littenberg B, 1996. Automated utility assessment of global health. *Qual Life Res*;5:175–82.
- Nease RF, Owens DK, 1994. A method for estimating the cost effectiveness or incorporating patient preferences into practice guidelines. *Med Decis Making*;14:382–92.
- Neaton JD, Wentworth DN, Rhame F, Hogan C, Abrams DI, Deyton L, 1994. Considerations in choice of a clinical end-point for AIDS clinical trials. *Stat Med*;13:2107–25.
- Nelson RA, Tanguay TL, Patterson CD, 1994. A quality-adjusted price index for personal computers. *J Business Econ Stat*;12:23–31.
- Nelson RM, Drought T, 1992. Justice and the moral acceptability of rationing medical care: the Oregon experiment. *J Med Philos*;17:97–117.
- Netterstrom B, Suadicanani P, 1993. Self-assessed job satisfaction and ischemic heart disease mortality: a 10 year follow-up of urban bus drivers. *Int J Epidemiol*;22:51–6.
- Nettleman MD, 1989. Outcome measurements in decision analysis: life versus quality of life. *Infect Control Hosp Epidemiol*;10:521–4.
- Nichol G, Laupacis A, Stiell IG, Orourke K, Anis A, Bolley H, Detsky AS, 1996. Cost-effectiveness analysis of potential improvements to emergency medical services for victims of out-of-hospital cardiac arrest. *Ann Emerg Med*;27:711–20.
- Nichols JD, Hines JE, Pollock KH, Hinz RL, Link WA, 1994. Estimating breeding proportions and testing hypotheses about costs of reproduction with capture-recapture data. *Ecology*;75:2052–65.
- Nichols JD, Kendall WL, 1995. The use of multistate capture-recapture models to address questions in evolutionary ecology. *J Appl Stat*;22:835–46.
- Nichols JD, Sauer JR, Pollock KH, Hestbeck JB, 1992. Estimating transition probabilities for stage-based population projection matrices using capture recapture data. *Ecology*;73:306–12.
- Niiranen A, Kajanti M, Tammilehto L, Mattson K, 1990. The clinical effect of medroxyprogesterone (MPA) in elderly patients with lung cancer. *Am J Clin Oncol: Cancer Clin Trials*;13:113–16.
- Nissinen A, Wiklund I, Lahti T, Akkila J, Wilson A, Wahl M, *et al.*, 1991. Antianginal therapy and quality-of-life: a comparison of the effects of transdermal nitroglycerin and long acting oral nitrates. *J Clin Epidemiol*;44:989–97.
- Noble S, Goa KL, 1996. Drug evaluation: ticlopidine – a review of its pharmacology, clinical efficacy and tolerability in the prevention of cerebral ischemia and stroke. *Drugs Aging*;8:214–32.
- Nolan JP, Tarsa NJ, Dibenedetto G, 1985. Case-finding for unsuspected thyroid disease: costs and health benefits. *Am J Clin Pathol*;83:346–55.
- Nord E, 1989. The significance of contextual factors in valuing health states. *Health Policy*;13:189–98.
- Nord E, 1991. EuroQol: health-related quality-of-life measurement – valuations of health states by the general public in Norway. *Health Policy*;18:25–36.
- Nord E, 1992. Methods for quality adjustment of life years. *Soc Sci Med*;34:559–69.
- Nord E, 1992. Alternative to QALYs: saved young life equivalent (SAVE) [reply]. *BMJ*;305:1365–6.
- Nord E, 1992. An alternative to QALYs: the saved young life equivalent (SAVE). *BMJ*;305:875–7.
- Nord E, 1993. The trade-off between severity of illness and treatment effect in cost-value analysis of health care. *Health Policy*;24:227–38.
- Nord E, 1993. The relevance of health state after treatment in prioritizing between different patients. *J Med Ethics*;19:37–42.
- Nord E, 1993. Unjustified use of the Quality of Well Being Scale in priority setting in Oregon. *Health Policy*;24:45–53.
- Nord E, 1993. Toward quality assurance in QALY calculations. *Int J Technol Assess Health Care*;9:37–45.
- Nord E, 1994. The QALY: a measure of social value rather than individual utility? *Health Econ*;3:89–93.
- Nord E, 1995. The person-trade-off approach to valuing health care programs. *Med Decis Making*;15:201–8.
- Nord E, 1996. Health status index models for use in resource allocation decisions: a critical review in the light of observed preferences for social choice. *Int J Technol Assess Health Care*;12:31–44.
- Nord E, Richardson J, Macarounas-Kirchmann K, 1993. Social evaluation of health care versus personal evaluation of health states: evidence on the validity of four health state scaling instruments using Norwegian and Australian surveys. *Int J Technol Assess Health Care*;9:463–78.
- Nord E, Richardson J, Street A, Kuhse H, Singer P, 1995. Who cares about cost: does economic analysis impose or reflect social values? *Health Policy*;34:79–94.

- Nord E, Richardson J, Street A, Kuhse H, Singer P, 1995. Maximizing health benefits vs egalitarianism: an Australian survey of health issues. *Soc Sci Med*;41:1429–37.
- Nord E, Street A, Richardson J, Kuhse H, Singer P, 1996. The significance of age and duration of effect in social evaluation of health care. *Health Care Anal*;4:103–11.
- Nordentoft M, Knudsen HC, Jessenpetersen B, Krasnik A, Saelan H, Treufeldt P, *et al.*, 1996. CCPP – Copenhagen Community Psychiatric Project. Implementation of community mental health centers in Copenhagen: effects of service utilization, social integration, quality-of-life and positive and negative symptoms. *Soc Psychiatry Psychiat Epidemiol*;31:336–44.
- Normantaylor FH, Palmer CR, Villar RN, 1996. Quality-of-life improvement compared after hip and knee replacement. *J Bone Joint Surg Br*;78:74–7.
- Norum J, 1996. Quality-of-life (QOL) measurement in economical analysis in cancer: a comparison of the EuroQol questionnaire, a simple QOL scale and the global QOL measure of the EORTC QLQ-C30. *Oncol Rep*;3:787–91.
- Norum J, Angelsen V, 1995. Chemotherapy in gastric cancer: an economic evaluation of the FAM (5-fluorouracil, adriamycin, mitomycin-C) versus ELF (etoposide, leucovorin, 5-fluorouracil) regimens. *J Chemother*;7:455–9.
- Norum J, Angelsen V, Wist E, Olsen JA, 1996. Treatment costs in Hodgkins disease: a cost–utility analysis. *Eur J Cancer*;32A:1510–17.
- Nou E, Aberg T, 1980. Quality of survival in patients with surgically treated bronchial carcinoma. *Thorax*;35:255–63.
- Nussbaum ES, Heros RC, Erickson DL, 1996. Cost-effectiveness of carotid endarterectomy. *Neurosurg*;38:237–43.
- Nussbaum H, O’Connell TX, 1980. Minimal breast cancer. Is excisional biopsy alone adequate treatment? *Cancer Clin Trials*;3:19–22.
- Nusselder WJ, van der Velden K, van Sonsbeek JLA, Lenior ME, van den Bos GAM, 1996. The elimination of selected chronic diseases in a population: the compression and expansion of morbidity. *Am J Public Health*;86:187–94.
- O’Brien B, 1994. Measurement of health-related quality of life in the economic evaluation of medicines. *Drug Inf J*;28:45–53.
- O’Brien B, Rushby J, 1990. Outcome assessment in cardiovascular cost–benefit studies. *Am Heart J*;119:740–8.
- O’Brien BJ, Banner NR, Gibson S, Yacoub MH, 1988. The Nottingham Health Profile as a measure of quality of life following combined heart and lung transplantation. *J Epidemiol Community Health*;42:232–4.
- O’Brien BJ, Buxton MJ, Ferguson BA, 1987. Measuring the effectiveness of heart transplant programs: quality of life data and their relationship to survival analysis. *J Chronic Dis*;40:S137–53.
- O’Brien PC, 1984. Procedures for comparing samples with multiple endpoints. *Biometrics*;40:1079–87.
- O’Brien PC, Geller NL, 1997. Interpreting tests for efficacy in clinical trials with multiple endpoints. *Control Clin Trials*;18:222–7.
- O’Boyle CA, 1992. Assessment of quality-of-life in surgery. *Br J Surg*;79:395–8.
- O’Boyle CA, 1994. Quality-of-life and cardiovascular medication. *Ir J Psychol*;15:126–47.
- Ochs J, Mulhern R, Kun L, 1988. Quality-of-life assessment in cancer patients. *Am J Clin Oncol: Cancer Clin Trials*;11:415–21.
- O’Connor AM, 1989. Effects of framing and level of probability on patients’ preferences for cancer chemotherapy. *J Clin Epidemiol*;42:119–26.
- Oddone EZ, Cowper P, Hamilton JD, Matchar DB, Hartigan P, Samsa G, *et al.*, 1993. Cost-effectiveness analysis of early zidovudine treatment of HIV infected patients. *BMJ*;307:1322–5.
- Oh PI, Maerov P, Pritchard D, Knowles SR, Einarson TR, Shear NH, 1996. A cost–utility analysis of 2nd line antibiotics in the treatment of acute otitis media in children. *Clin Ther*;18:160–82.
- O’Kelly TJ, Westaby S, 1990. Trauma centres and the efficient use of financial resources. *Br J Surg*;77:1142–4.
- Oldenburg B, MacDonald GJ, Perkins RJ, 1988. Prediction of quality of life in a cohort of end-stage renal disease patients. *J Clin Epidemiol*;41:555–64.
- Oldridge N, Furlong W, Feeny D, Torrance G, Guyatt G, Crowe J, *et al.*, 1993. Economic evaluation of cardiac rehabilitation soon after acute myocardial infarction. *Am J Cardiol*;72:154–61.
- Olliaro P, Dautzenberg B, 1994. Control of the body burden of mycobacterium avium complex is associated with improved quality-of-life and prolonged survival of patients with AIDS: a prospective trial with rifabutin combined with isoniazid, clofazimine, ethambutol. *J Chemother*;6:189–96.
- Olschewski M, 1984. Methods of evaluating and analyzing quality of life in clinical trials. *Biometrics*;40:264.
- Olschewski M, Schulgen G, Schumacher M, 1992. In discussion of Cox *et al.* Quality-of-life assessment: can we keep it simple? *J R Stat Soc A*;155:353–93.
- Olschewski M, Schumacher M, 1990. Statistical analysis of quality of life data in cancer clinical trials. *Stat Med*;9:749–63.
- Olsen JA, 1994. Persons vs years: 2 ways of eliciting implicit weights. *Health Econ*;3:39–46.
- O’Neill C, Normand C, Cupples M, McKnight A, 1996. Cost-effectiveness of personal health education in primary care for people with angina in the Greater Belfast area of Northern Ireland. *J Epidemiol Community Health*;50:538–40.



- O'Neill C, Normand C, Cupples M, McKnight A, 1996. A comparison of 3 measures of perceived distress: results from a study of angina patients in general practice in Northern Ireland. *J Epidemiol Community Health*;50:202–6.
- Oster G, Huse DM, Lacey MJ, Epstein AM, 1994. Cost-effectiveness of ticlopidine in preventing stroke in high risk patients. *Stroke*;25:1149–56.
- Owens DK, Nease RF, 1992. Occupational exposure to Human Immunodeficiency Virus and hepatitis B virus: a comparative analysis of risk. *Am J Med*;92:503–12.
- Paccagnella A, Pappagallo GL, Segati R, Zorat P, Cavaniglia G, Fosser VP, *et al.*, 1990. Response and toxicity of cisplatin and 120-h 5-fluorouracil infusion in pretreated and untreated patients with advanced epidermoid cancer of the head and neck. *Am J Clin Oncol: Cancer Clin Trials*;13:194–8.
- Pal DK, 1996. Quality-of-life assessment in children: a review of conceptual and methodological issues in multidimensional health status measures. *J Epidemiol Community Health*;50:391–6.
- Parker J, 1990. Images of health, urban design and human well-being. *Statistician*;39:191–7.
- Parker MJ, 1993. Hip fractures and QALYs [reply]. *J Bone Joint Surg Br*;75:164.
- Parker MJ, Myles JW, Anand JK, Drewett R, 1992. Cost-benefit analysis of hip fracture treatment. *J Bone Joint Surg Br*;74:261–4.
- Parkerson GR, Broadhead WE, Tse CKJ, 1992. Quality-of-life and functional health of primary care patients. *J Clin Epidemiol*;45:1303–13.
- Parkerson GR, Broadhead WE, Tse CKJ, 1996. Anxiety and depressive symptom identification using the Duke health profile. *J Clin Epidemiol*;49:85–93.
- Parmar MKB, Machin D, 1995. *Survival analysis: a practical approach*. London: John Wiley.
- Parmenter TR, 1992. Quality-of-life of people with developmental disabilities. *Int Rev Res Mental Retard*;18:247–87.
- Parmigiani G, 1993. On optimal screening ages. *J Am Stat Assoc*;88:622–8.
- Pascal N, 1979. An economic analysis of the utilization of medical services by Medicare patients. Unpublished.
- Paterson AHG, Szafran O, Hanson J, Cyr M, Lees AW, 1985. Response to treatment and its influence on survival in metastatic breast cancer. *Am J Clin Oncol: Cancer Clin Trials*;8:283–92.
- Patterson RE, Eisner RL, Horowitz SF, 1995. Comparison of cost-effectiveness and utility of exercise ecg, single photon emission computed tomography, positron emission tomography, and coronary angiography for diagnosis of coronary artery disease. *Circulation*;91:54–65.
- Patterson RE, Eng C, Horowitz SF, *et al.*, 1984. Bayesian comparison of cost-effectiveness of different clinical approaches to diagnose coronary artery disease. *J Am Coll Cardiol*;4:278–89.
- Paul JE, Mauskopf JA, Bell L, 1995. Cost-consequence models for varicella zoster virus infections. *Pharmacother*;15:S49–58.
- Pawitan Y, Self S, 1993. Modelling disease marker processes in AIDS. *J Am Stat Assoc*;88:719–26.
- Pederson LL, Wanklin JM, Lefcoe NM, 1988. Self-reported long-term smoking cessation in patients with respiratory disease: prediction of success and perception of health effects. *Int J Epidemiol*;17:804–9.
- Pepe MS, Longton G, Thornquist M, 1991. A qualifier q for the survival function to describe the prevalence of a transient condition. *Stat Med*;10:413–21.
- Pereira J, 1993. What does equity in health mean? *J Soc Policy*;22:19–48.
- Permanyermiralda G, Alonso J, Anto JM, Aljardeguimera M, Solersoler J, 1991. Comparison of perceived health status and conventional functional evaluation in stable patients with coronary artery disease. *J Clin Epidemiol*;44:779–86.
- Perneger TV, Leplege A, Etter JF, Rougemont A, 1995. Validation of a French language version of the MOS 36-item short-form health survey (SF-36) in young healthy adults. *J Clin Epidemiol*;48:1051–60.
- Petrou S, Renton A, 1993. The QALY: a guide for the public health physician. *Public Health*;107:327–36.
- Phillips RS, Goldman L, Bergner M, 1990. Patient characteristics in support: activity status and cognitive function. *J Clin Epidemiol*;43:33–6S.
- Phillips RS, Knaus WA, 1990. Patient characteristics in support: sociodemographics, admission diagnosis, co-morbidities and acute physiology score. *J Clin Epidemiol*;43:29–31S.
- Picavet HSJ, van den Bos GAM, 1996. Comparing survey data on functional disability: the impact of some methodological differences. *J Epidemiol Community Health*;50:86–93.
- Piccoli A, Pillon L, Passerini P, Ponticelli C, 1994. Therapy for idiopathic membranous nephropathy: tailoring the choice by decision analysis. *Kidney Int*;45:1193–202.
- Pickard JD, Bailey S, Sanderson H, Rees M, Garfield JS, 1990. Steps towards cost-benefit analysis of regional neurosurgical care. *BMJ*;301:629–35.
- Plasswilm L, Seegenschmiedt H, Ganssaue F, Sauer R, 1996. Simultaneous radiochemotherapy for recurrent and metastatic breast carcinoma: evaluation of 2 treatment concepts. *Am J Clin Oncol: Cancer Clin Trials*;19:403–7.

- Pliskin JS, Stason WB, Weinstein MC, Johnson RA, Cohn PF, McEnany MT, *et al.*, 1981. Coronary artery bypass graft surgery: clinical decision making and cost-effectiveness analysis. *Med Decis Making*;1:10–28.
- Pocock SJ, 1991. A perspective on the role of quality-of-life assessment in clinical trials. *Control Clin Trials*;12:S257–65.
- Pocock SJ, Geller NL, Tsiatis AA, 1987. The analysis of multiple endpoints in clinical trials. *Biometrics*;43:487–98.
- Poitrenaud J, Vallerymasson J, Costagliola D, Darcet P, Lion MR, 1983. Retirement, health and psychological well-being: a longitudinal study of a French managerial population. *Rev Epidemiol Sante Publique*;31:469–82.
- Prentice RL, Williams BJ, Peterson AV, 1981. On the regression analysis of multivariate failure time data. *Biometrika*;68:373–9.
- Presant CA, 1984. Quality of life in cancer patients: who measures what. *Am J Clin Oncol: Cancer Clin Trials*;7:571–3.
- Prieto L, Alonso J, Viladrich MC, Anto JM, 1996. Scaling the Spanish version of the Nottingham Health Profile: evidence of limited value of item weights. *J Clin Epidemiol*;49:31–8.
- Pritchard CW, 1994. Depression and smoking in pregnancy in Scotland. *J Epidemiol Community Health*;48:377–82.
- Probstfield JL, 1988. Rationale and design of a randomized clinical trial on prevention of stroke in isolated systolic hypertension. *J Clin Epidemiol*;41:1197–208.
- Prorok PC, Connor RJ, Baker SG, 1990. Statistical considerations in cancer screening programs. *Urol Clin North Am*;17:699–708.
- Provenzale D, Kemp JA, Arora S, Wong JB, 1994. A guide for surveillance of patients with Barretts-esophagus. *Am J Gastroenterol*;89:670–80.
- Rabin R, Rosser RM, Butler C, 1993. Impact of diagnosis on utilities assigned to states of illness. *J R Soc Med*;86:444–8.
- Ramsey SD, Patrick DL, Albert RK, Larson EB, Wood DE, Raghu G, 1995. The cost-effectiveness of lung transplantation: a pilot study. *Chest*;108:1594–1601.
- Ranaboldo CJ, Chant ADB, 1992. Reprocessing data to form QALYs. *BMJ*;305:424.
- Randolph AG, Washington AE, Prober CG, 1993. Cesarean delivery for women presenting with genital herpes lesions: efficacy, risks, and costs. *JAMA*;270:77–82.
- Rasbash J, Woodhouse G, 1995. MLn command reference. London: Institute of Education.
- Rawles J, 1989. Castigating QALYs. *J Med Ethics*;15:143–7.
- Rawles J, Light J, Watt M, 1992. Quality-of-life in the 1st 100 days after suspected acute myocardial infarction: a suitable trial end-point. *J Epidemiol Community Health*;46:612–16.
- Rawles J, Light J, Watt M, 1993. Loss of quality-adjusted days as a trial end-point: effect of early thrombolytic treatment in suspected myocardial infarction. *J Epidemiol Community Health*;47:377–81.
- Rawles J, Rawles K, 1990. The QALY argument: a physicians and a philosophers view. *J Med Ethics*;16:93–4.
- Re MA, Budde CE, Caceres MO, 1996. Survival probability in the presence of a dynamic trap. *Phys Rev E*;54:4427–30.
- Reckless JPD, 1993. Cost-effectiveness of hypolipemic drugs. *Postgrad Med J*;69:S30–3.
- Redelmeier DA, Guyatt GH, Goldstein RS, 1996. Assessing the minimal important difference in symptoms: a comparison of 2 techniques. *J Clin Epidemiol*;49:1215–19.
- Reeves RA, Chen E, 1994. Who wants to eliminate heart disease? *J Clin Epidemiol*;47:667–70.
- Reisine ST, Fertig J, Weber J, Leder S, 1989. Impact of dental conditions on patients' quality of life. *Community Dent Oral Epidemiol*;17:7–10.
- Research Information Systems, 1995. Reference manager, version 7. USA: Research Information Systems.
- Revicki DA, 1995. Measuring health outcomes for cost-effectiveness studies: are all quality adjusted life years created equal? *Drug Inf J*;29:1459–67.
- Revicki DA, Allen H, Bungay K, Williams GH, Weinstein MC, 1994. Responsiveness and calibration of the general well-being adjustment scale in patients with hypertension. *J Clin Epidemiol*;47:1333–42.
- Revicki DA, Brown RE, Palmer W, Bakish D, Rosser WW, Anton SF, *et al.*, 1995. Modeling the cost-effectiveness of antidepressant treatment in primary care. *Pharmacoeconomics*;8:524–40.
- Revicki DA, Kaplan RM, 1993. Relationship between psychometric and utility-based approaches to the measurement of health-related quality of life. *Qual Life Res*;2:477–87.
- Revicki DA, Simpson KN, Wu AW, LaVallee RL, 1995. Evaluating the quality of life associated with rifabutin prophylaxis for Mycobacterium avium complex in persons with AIDS: combining Q-TWiST and multiattribute utility techniques. *Qual Life Res*;4:309–18.
- Reynolds DL, Torrance GW, Badley EM, Bennett KJ, Chambers LW, Goldsmith CH, *et al.*, 1993. Modeling the population health impact of musculoskeletal diseases: arthritis. *J Rheumatol*;20:1037–47.
- Reynolds P, Kaplan GA, 1986. Psychologic well-being and cancer: prospective evidence from the Alameda County Study. *Am J Epidemiol*;124:503–4.
- Rhodd RG, 1993. QALY as an estimator of social benefits or anti-drug strategies. *J Policy Model*;15:329–33.
- Rice DP, Hodgson TA, 1980. Social and economic implications of cancer in the United States of America. *World Health Stat Q*;33:56–100.

- Richards DM, Irving MH, 1996. Cost–utility analysis of home parenteral nutrition. *Br J Surg*;83:1226–9.
- Richardson J, 1994. Cost–utility analysis: what should be measured. *Soc Sci Med*;39:7–21.
- Richardson J, Hall J, Salkeld G, 1996. The measurement of utility in multiphase health states. *Int J Technol Assess Health Care*;12:151–62.
- Richardson J, Schwartz S, 1994. Quality-adjusted life years [reply]. *Aust J Publ Health*;18:227–8.
- Richardson J, Schwartz S, Glasziou PP, 1993. QALYs for resource allocation – a reply to Burrows and Brown. *Aust J Publ Health*;17:394–6.
- Ridley S, Biggam M, Stone P, 1994. A cost–utility analysis of intensive therapy. 2. Quality-of-life in survivors. *Anaesthesia*;49:192–6.
- Robine JM, Colvez A, 1990. Expectation of life in good health an indicator of future to measure state of wellbeing. *Rev Epidemiol Sante Publique*;38:373–5.
- Robine JM, Labbe M, Seroussi MC, Colvez A, 1989. The Upper Normandy longitudinal survey on disability among the aged 1978–1985. *Rev Epidemiol Sante Publique*;37:37–48.
- Robine JM, Mathers CD, Bucquet D, 1993. Distinguishing health expectancies and health-adjusted life expectancies from quality-adjusted life years. *Am J Public Health*;83:797–8.
- Robine JM, Michel JP, Branch LG, 1992. Measurement and utilization of healthy life expectancy: conceptual issues. *Bull World Health Organ*;70:791–800.
- Robine JM, Ritchie K, 1991. Healthy life expectancy: evaluation of global indicator of change in population health. *BMJ*;302:457–60.
- Robinson R, 1993. Economic evaluation and health care. 4. Cost–utility analysis. *BMJ*;307:859–62.
- Robinson R, 1995. Economics, QALYs and medical ethics: a practical agenda. *Health Care Anal*;3:229–32.
- Rockwood K, Stolee P, Fox RA, 1993. Use of goal attainment scaling in measuring clinically important change in the frail elderly. *J Clin Epidemiol*;46:1113–18.
- Rogers A, Rogers RG, Branch LG, 1989. A multistate analysis of active life expectancy. *Public Health Rep*;104:222–6.
- Romelsjo A, Kaplan GA, Cohen RD, Allebeck P, Andreasson S, 1992. Protective factors and social risk factors for hospitalization and mortality among young men. *Am J Epidemiol*;135:649–58.
- Romney DM, Evans DR, 1996. Toward a general model of health-related quality of life. *Qual Life Res*;5:235–41.
- Rosenman J, Choi NC, 1982. Improved quality of life of patients with small-cell carcinoma of the lung by elective irradiation of the brain. *Int J Radiat Oncol Biol Phys*;8:1041–3.
- Rosenthal MA, Webster PJ, Gebiski VJ, Stuartharris RC, Langlands AO, Boyages J, 1992. The cost of treating small-cell lung cancer. *Med J Aust*;156:605–10.
- Ross CE, Hayes D, 1988. Exercise and psychologic well-being in the community. *Am J Epidemiol*;127:762–71.
- Rossdegnan D, Soumerai SB, Avorn J, Bohn RL, Bright R, Aledort LM, 1995. Hemophilia home treatment: economic analysis and implications for health policy. *Int J Technol Assess Health Care*;11:327–44.
- Rosser R, 1993. The history of health-related quality-of-life in 10<sup>1/2</sup> paragraphs. *J R Soc Med*;86:315–18.
- Rosser R, Kind P, 1978. A scale of valuations of states of illness: is there a social consensus? *Int J Epidemiol*;7:347–58.
- Rosser RM, Watts VC, 1972. The measurement of hospital output. *Int J Epidemiol*;1:361–8.
- Rossouw JE, Harlan WR, 1994. Postmenopausal estrogen and the risk of breast cancer: the need for randomized trials. *Ann Epidemiol*;4:255–6.
- Rouse DJ, Owen J, Goldenberg RL, Vermund SH, 1995. Zidovudine for the prevention of vertical HIV transmission: a decision-analytic approach. *J Acquir Immune Defic Syndr Hum Retrovirol*;9:401–7.
- Rucker G, Messerer D, 1988. Remission duration: an example of interval-censored observations. *Stat Med*;7:1139–45.
- Russell JGB, 1990. Is screening for abdominal aortic aneurysm worthwhile? *Clin Radiol*;41:182–4.
- Ryan LS, Moritz TE, Zacharski LR, 1986. Reliability and validity of a quality of life measure for lung and colon cancer patients. *Control Clin Trials*;7:258.
- Sackett DL, Torrance GW, 1978. The utility of different health states as perceived by the general public. *J Chronic Dis*;31:697–704.
- Sailly JC, 1996. Medico-economic evaluation of medical strategies in prevention diagnosis and therapy. *Med Nucl*;20:532–8.
- Salvioli G, 1994. Therapeutic problems in geriatrics: osteoporosis. *G Gerontol*;42:653–60.
- Sandler RS, Jordan MC, Kupper LL, 1988. Development of a Crohn's index for survey research. *J Clin Epidemiol*;41:451–8.
- Sarasin FP, Eckman MH, 1993. Management and prevention of thromboembolic events in patients with cancer-related hypercoagulable states: a risky business. *J Gen Intern Med*;8:476–86.
- Sarasin FP, Schifferli JA, 1994. Prophylactic oral anticoagulation in nephrotic patients with idiopathic membranous nephropathy. *Kidney Int*;45:578–85.
- Sartorius N, Janca A, 1996. Psychiatric assessment instruments developed by the World Health Organization. *Soc Psychiatry Psychiat Epidemiol*;31:55–69.

- SAS Institute Inc, 1989. SAS/STAT users guide, v. 6, vol. 2. 4th ed. Cary, NC: SAS Institute Inc.
- SAS Institute Inc, 1992. SAS/STAT software: changes and enhancements; release 6.07. SAS technical report P-229. Cary, NC: SAS Institute Inc.
- Sassine MP, Mergler D, Larribe F, Belanger S, 1996. Mental health deterioration among workers exposed to styrene. *Rev Epidemiol Sante Publique*;44:14–24.
- Savitz DA, Hamman RF, Grace C, Stroo K, 1986. Respondents' attitudes regarding participation in an epidemiologic study. *Am J Epidemiol*;123:362–6.
- Sayers G, 1992. Alternative to QALYs: saved young life equivalent (SAVE). *BMJ*;305:1365.
- Schaapveld K, van de Water HPA, van Ginneken JKS, 1991. Healthy life expectancy, a contemporary indicator for public health. *Ned Tijdsch Geneesk*;135:1703–6.
- Schechter CB, Rose DN, Fahs MC, Silver AL, 1990. Tuberculin screening: cost-effectiveness analysis of various testing schedules. *Am J Prev Med*;6:167–75.
- Schluchter MD, 1988. Analysis of incomplete multivariate data using linear models with structured covariance matrices. *Stat Med*;7:317–24.
- Schluchter MD, 1992. Methods for the analysis of informatively censored longitudinal data. *Stat Med*;11:1861–70.
- Schmid GP, 1995. Understanding the essentials of economic evaluation. *J Acquir Immune Defic Syndr Hum Retrovirol*;10S6–13.
- Schramm W, Offner A, Szucs T, 1995. Socioeconomic evaluation of hemostaseologic therapies and anticoagulation. *Zentralbl Chir*;120:593–7.
- Schraub S, Mercier M, Eschwege F, Le Febvre JL, Vrousos C, Barthod L, 1996. A new quality-of-life questionnaire for head and neck cancers. *Rev Epidemiol Sante Publique*;44:346–57.
- Schulgen G, 1989. A measure of similarity for response curves based on ranks. *Stat Med*;8:1401–11.
- Schulman KA, Glick H, Buxton M, Sculpher M, Backhouse M, Bell L, *et al.*, 1996. The economic evaluation of the FIRST study: design of a prospective analysis alongside a multinational phase III clinical trial. *Control Clin Trials*;17:304–15.
- Schumacher M, Olschewski M, Schulgen G, 1991. Assessment of quality-of-life in clinical trials. *Stat Med*;10:1915–30.
- Schwartz CE, Cole BF, Gelber RD, 1994. Measuring patient-centered outcomes: adapting the Q-TWiST for multiple sclerosis. *Qual Life Res*;3:62.
- Schwartz CE, Cole BF, Gelber RD, 1995. Measuring patient-centered outcomes in neurologic disease: extending the Q-TWiST method. *Arch Neurol*;52:754–62.
- Schwartz CE, Cole BF, Vickrey BG, Gelber RD, 1995. The Q-TWiST approach to assessing health-related quality of life in epilepsy. *Qual Life Res*;4:135–41.
- Schwartz CE, Lipshutz M, Kshalal P, Zeng Q, Kaplan R, Rapkin B, 1995. The extended Q-TWiST method for neurologic disease: new method and measures [abstract]. *Qual Life Res*;4:482–3.
- Schwartz JS, 1995. Clinical economics and noncoronary vascular disease. *J Vasc Interv Radiol*;6:S116–24.
- Schwartz S, Richardson J, Glasziou PP, 1993. Quality-adjusted life years: origins, measurements, applications, objections. *Aust J Publ Health*;17:272–8.
- Schwarz RP Jr, 1991. Maintaining integrity and credibility in industry-sponsored clinical research. *Control Clin Trials*;12:753–60.
- Scott CB, 1997. Quality-adjusted survival analysis of malignant glioma patients. *Control Clin Trials*;18:277–85.
- Scott CB, Stetz J, Bruner DW, Wasserman TH, 1994. Radiation Therapy Oncology Group quality of life assessment: design, analysis, and data management issues. *Qual Life Res*;3:199–206.
- Selai CE, Rosser RM, 1993. Good quality quality: some methodological issues. *J R Soc Med*;86:440–3.
- Selai CE, Trimble MR, 1994. The role of quality-of-life measures in psychopharmacology. *Hum Psychopharmacol Clin Exp*;9:211–14.
- Sen A, 1981. Public action and the quality of life in developing countries. *Oxf Bull Econ Stat*;43:287–319.
- Shaar K, McCarthy M, 1994. Definitions and determinants of handicap in people with disabilities. *Epidemiol Rev*;16:228–42.
- Shaar KH, McCarthy M, 1992. Disadvantage as a measure of handicap: a paired sibling study of disabled adults in Lebanon. *Int J Epidemiol*;21:101–7.
- Shackley P, Cairns J, 1996. Evaluating the benefits of antenatal screening: an alternative approach. *Health Policy*;36:103–15.
- Shapiro SH, Macklem PT, Graydonald K, Martin JG, Ernst PP, Wooddauphinee S, *et al.*, 1991. A randomized clinical trial of negative pressure ventilation in severe chronic obstructive pulmonary disease: design and methods. *J Clin Epidemiol*;44:483–96.
- Shepard DS, Walsh JA, Kleinau E, Stansfield S, Bhalotra S, 1995. Setting priorities for the childrens vaccine initiative: a cost-effectiveness approach. *Vaccine*;13:707–14.
- Shephard RJ, 1989. Nutritional benefits of exercise. *J Sports Med Phys Fitness*;29:83–90.
- Shephard RJ, 1989. Exercise in secondary and tertiary rehabilitation: costs and benefits. *J Cardiopulm Rehabil*;9:188–94.
- Shephard RJ, 1994. Physiological basis of training in the elderly. *Sci Sports*;9:189–96.

- Shephard RJ, Montelpare W, 1988. Geriatric benefits of exercise as an adult. *J Gerontol*;43:M86–90.
- Siegrist J, Junge A, 1989. Background material for the workshop on QALYs: conceptual and methodological problems in research on the quality of life in clinical medicine. *Soc Sci Med*;29:463–8.
- Sigurdardottir V, Bolund C, Brandberg Y, Sullivan M, 1993. The impact of generalized malignant melanoma on quality of life evaluated by the EORTC questionnaire technique. *Qual Life Res*;2:193–203.
- Silverstein MD, Albert DA, Hadler NM, Ropes MW, 1988. Prognosis in SLE: comparisons of Markov model to life table analysis. *J Clin Epidemiol*;41:623–33.
- Simpson KN, 1995. Design and assessment of cost-effectiveness studies in AIDS populations. *J Acquir Immune Defic Syndr Hum Retrovirus*;10:S28–32.
- Simpson NB, 1993. Social and economic aspects of acne and the cost-effectiveness of isotretinoin. *J Dermatol Treat*;4:S6–9.
- Singer P, McKie J, Kuhse H, Richardson J, 1995. Double jeopardy and the use of QALYs in health care allocation. *J Med Ethics*;21:144–50.
- Siu AL, Ouslander JG, Osterweil D, Reuben DB, Hays RD, 1993. Change in self-reported functioning in older persons entering a residential care facility. *J Clin Epidemiol*;46:1093–101.
- Slotje DJ, 1991. Measuring the quality-of-life across countries. *Rev Econ Stat*;73:684–93.
- Smith A, 1987. Qualms about QALYs. *Lancet*;i:1134–6.
- Smith A, 1987. Quality-adjusted life-years. *Lancet*;ii:47.
- Smith GT, 1990. The economics of hypertension and stroke. *Am Heart J*;119:725–8.
- Smith K, 1995. Quality adjusted life years in the NHS: health care for all? *Care Crit Ill*;11:82–5.
- Smith KJ, Pesce RR, 1994. Pulmonary artery catheterization in exacerbations of COPD requiring mechanical ventilation: a cost-effectiveness analysis. *Respiratory Care*;39:961–7.
- Smith RD, Hall J, Gurney H, Harnett PR, 1993. A cost-utility approach to the use of 5-fluorouracil and levamisole as adjuvant chemotherapy for Duke C colonic carcinoma. *Med J Aust*;158:319–22.
- Smith TJ, Hillner BE, 1993. The efficacy and cost-effectiveness of adjuvant therapy of early breast cancer in premenopausal women. *J Clin Oncol*;11:771–6.
- Sonnenberg FA, Beck JR, 1993. Markov models in medical decision making: a practical guide. *Med Decis Making*;13:322–8.
- Spiegelhalter DJ, Gore SM, Fitzpatrick R, Fletcher AE, Jones DR, Cox DR, 1992. Quality-of-life measures in health care. 3. Resource allocation. *BMJ*;305:1205–9.
- Spitzer WO, 1991. Drugs as determinants of health and disease in the population: an orientation to the bridge science of pharmacoepidemiology. *J Clin Epidemiol*;44:823–30.
- Sprangers MAG, Aaronson NK, 1992. The role of health care providers and significant others in evaluating the quality-of-life of patients with chronic disease: a review. *J Clin Epidemiol*;45:743–60.
- St Leger AS, Spencely M, McCollum CN, Mossa M, 1996. Screening for abdominal aortic aneurysm: a computer assisted cost-utility analysis. *Eur J Vasc Endovasc Surg*;11:183–90.
- Stalmeier PFM, Bezembinder TGG, Unic IJ, 1996. Proportional heuristics in time tradeoff and conjoint measurement. *Med Decis Making*;16:36–44.
- Stansfeld SA, North FM, White I, Marmot MG, 1995. Work characteristics and psychiatric disorder in civil servants in London. *J Epidemiol Community Health*;49:48–53.
- Staquet MJ, Hays RD, Fayers PM, editors, 1998. Quality of life assessment in clinical trials: methods and practice. New York: Oxford University Press.
- Statistical Science Inc, 1991. S-PLUS user's manual. Seattle: Statistical Sciences Inc.
- Stengele U, Baumgartner BR, Chezmar JL, Slaker DP, 1991. Biliary lithotripsy versus cholecystectomy: a cost-utility analysis. *J Lithotripsy Stone Dis*;3:133–40.
- Stephens RJ, Hopwood P, 1995. Defining and analysing the palliation of physical symptoms: a deceptively difficult exercise. *Palliat Med*;9:63–75.
- Stevenson RC, Pharoah POD, Cooke RWI, Sandhu B, 1991. Predicting costs and outcomes of neonatal intensive care for very low birth-weight infants. *Public Health*;105:121–6.
- Stewart AL, Hays RD, Wells KB, Rogers WH, Spritzer KL, Greenfield S, 1994. Long-term functioning and well-being outcomes associated with physical activity and exercise in patients with chronic conditions in the medical outcomes study. *J Clin Epidemiol*;47:719–30.
- Stiggelbout AM, Kiebert GM, Kievit J, Leer JWH, Habbema JDF, de Haes JCJM, 1995. The utility of the time trade-off method in cancer patients: feasibility and proportional trade-off. *J Clin Epidemiol*;48:1207–14.
- Stiggelbout AM, Kiebert GM, Kievit J, Leer JWH, Stoter G, de Haes JCJM, 1994. Utility assessment in cancer patients: adjustment or time tradeoff scores for the utility of life years and comparison with standard gamble scores. *Med Decis Making*;14:82–90.
- StLeger AS, Spencely M, McCollum CN, Mossa M, 1996. Screening for abdominal aortic aneurysm: a computer assisted cost-utility analysis. *Eur J Vasc Endovasc Surg*; 11:183–90.

- Stucki G, Daltroy L, Katz JN, Johannesson M, Liang MH, 1996. Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. *J Clin Epidemiol*;49:711–17.
- Su L, Guess HA, Girman CJ, Jacobsen SJ, Oesterling JE, Panser LA, *et al.*, 1996. Adverse effects of medications on urinary symptoms and flow rate: a community-based study. *J Clin Epidemiol*;49:483–7.
- Suarezvarela MMM, Castello LS, Benajas AP, Gonzalez AL, 1995. Mortality, morbidity and drug consumption in a rural area (Spain). *Eur J Epidemiol*;11:403–14.
- Sullivan M, 1994. Cross-cultural validation for quality-of-life. *J Clin Epidemiol*;47:1465.
- Sutherland HJ, Llewellyn-Thomas H, Boyd NF, Till JE, 1982. Attitudes toward quality of survival: the concept of 'maximum endurable time'. *Med Decis Making*;2:299–309.
- Sutherland HJ, Lockwood GA, Boyd NF, 1990. Ratings of the importance of quality of life variables: therapeutic implications for patients with metastatic breast cancer. *J Clin Epidemiol*;43:661–6.
- Sutton AJ, Abrams KR, Jones DJ, Sheldon TA, Song F, 1999. Systematic reviews of trials and other studies. *Health Technol Assess*;2(19).
- Szczepura A, 1992. Routine low-cost pathology tests: measuring the value in use of bacteriology tests in hospital and primary care. *Health Serv Manage Res*;5:225–37.
- Szmukler GI, Burgess P, Herrman H, Benson A, Colusa S, Bloch S, 1996. Caring for relatives with serious mental illness: the development of the experience of caregiving inventory. *Soc Psychiatry Psychiat Epidemiol*;31:137–48.
- Tandon PK, 1990. Applications of global statistics in analyzing quality of life data. *Stat Med*;9:819–27.
- Tandon PK, Stander H, Schwarz RP, 1987. Assessment of quality of life in congestive heart-failure patients. *Control Clin Trials*;8:294–5.
- Tandon PK, Stander H, Schwarz RP, 1989. Analysis of quality of life data from a randomized, placebo-controlled heart failure trial. *J Clin Epidemiol*;42:955–62.
- Tarter RE, Switala J, Arria A, Vanthiel DH, 1991. Impact of liver disease on daily living in transplantation candidates. *J Clin Epidemiol*;44:1079–83.
- Tate DG, Dijkers M, Johnson-Greene, L, 1996. Outcome measures in quality of life. *Top Stroke Rehabil*;2:1–17.
- Tauber R, Groos E, 1986. Palliative treatment in terminally ill cancer patients: quality-adjusted life expectancy and the ethical problems of nephrostomy in patients with cancer. *Theor Surg*;1:148–54.
- Taylor WM, 1983. The estimation of quality-adjusted rates of return in stamp auctions. *J Finance*;38:1095–110.
- Taylor WM, 1992. The estimation of quality-adjusted auction returns with varying transaction intervals. *J Fin Quantit Anal*;27:131–42.
- Temkin N, McLean A Jr, Dikmen S, Gale J, Bergner M, Almes MJ, 1988. Development and evaluation of modifications to the Sickness Impact Profile for head injury. *J Clin Epidemiol*;41:47–57.
- Tenhaven TR, Chinchilli VM, 1994. Bayesian hierarchical analysis of within-units variances in repeated measures experiments. *Stat Med*;13:1841–52.
- Ter Berg HWM, Dippel DWJ, Habbema JDF, Bijlsma JB, van Gijn J, Tulleken CAF, *et al.*, 1988. Treatment of intact familial intracranial aneurysms: a decision-analytical approach. *Neurosurgery*;23:329–34.
- Testa MA, Nackley JF, 1994. Methods for quality-of-life studies. *Annu Rev Public Health*;15:535–59.
- Thomas DC, 1987. Use of computer simulation to explore analytical issues in nested case-control studies of cancer involving extended exposures: methods and preliminary findings. *J Chronic Dis*;40:201–8S.
- Thomas G, 1986. QALYs or short straws? *BMJ*;293:338.
- Thornquist MD, 1993. Proportional hazards model for repeated measures with monotonic ordinal response. *Biometrics*;49:721–30.
- Till JE, Dehaes JCJM, 1991. Quality-adjusted survival analysis. *J Clin Oncol*;9:525–6.
- Till JE, Osoba D, Pater JL, Young JR, 1994. Research on health-related quality of life: dissemination into practical applications. *Qual Life Res*;3:279–83.
- Tolias CM, Choksey MS, 1996. Will increased awareness among physicians of the significance of sudden agonizing headache affect the outcome of subarachnoid hemorrhage? Coventry and Warwickshire study: audit of subarachnoid hemorrhage (establishing historical controls), hypothesis, campaign layout, and cost estimation. *Stroke*;27:807–12.
- Torrance GW, 1976. Social preferences for health states. An empirical evaluation of three measurement techniques. *Socio-econ Plan Sci*;10:129–36.
- Torrance GW, 1986. Measurement of health state utilities for economic appraisal. *J Health Econ*;5:1–30.
- Torrance GW, 1987. Utility approach to measuring health-related quality of life. *J Chronic Dis*;40:593–600.
- Torrance GW, Blaker D, Detsky A, Kennedy W, Schubert F, Menon D, *et al.*, 1996. Canadian guidelines for economic evaluation of pharmaceuticals. *PharmacoEconomics*;9:535–59.
- Torrance GW, Feeny D, 1989. Utilities and quality-adjusted life years. *Int J Technol Assess Health Care*;5:559–75.
- Torrance GW, Furlong W, Feeny D, Boyle M, 1995. Multiattribute preference functions: health utilities index. *PharmacoEconomics*;7:503–20.
- Torrance GW, Thomas WH, Sackett DL, 1972. A utility maximisation model for the evaluations of health care programs. *Health Serv Res*;7:118–33.

- Tousignant P, Cosio M, Levy R, Groome P, 1990. Quality-adjusted life years added by treatment of obstructive sleep apnea. *Med Decis Making*;10:327.
- Tousignant P, Cosio MG, Levy RD, Groome PA, 1994. Quality-adjusted life years added by treatment of obstructive sleep apnea. *Sleep*;17:52–60.
- Townsend M, Guyatt GH, Berman LB, Taylor DW, 1985. Assessing disease-specific quality of life in clinical trials. *Control Clin Trials*;6:243.
- Trachtenberg J, 1996. Appraisal of current therapies for prostate cancer. *Eur Urol*;29:110–13.
- Tsevat J, Dawson NV, Matchar DB, 1990. Assessing quality-of-life and preferences in the seriously ill using utility theory. *J Clin Epidemiol*;43:S73–7.
- Tsevat J, Duke D, Goldman L, Pfeffer MA, Lamas GA, Soukup JR, *et al.*, 1995. Cost-effectiveness of captopril therapy after myocardial infarction. *J Am Coll Cardiol*;26:914–19.
- Tsevat J, DurandZaleski I, Pauker SG, 1989. Cost-effectiveness of antibiotic prophylaxis for dental procedures in patients with artificial joints. *Am J Public Health*;79:739–43.
- Tsiantis J, Dragonas T, Cox A, Smith M, Ispanovic V, Sampaiofaria J, 1996. Promotion of childrens early psychosocial development through primary health care services. *Paediatr Perinatal Epidemiol*;10:339–54.
- Tsiatis AA, de Gruttola V, Wulfsohn MS, 1995. Modelling the relationship of survival to longitudinal data measured with error. Applications to survival and CD4 counts in patients with AIDS. *J Am Stat Assoc*;90:27–37.
- Tubman TRJ, Halliday HL, Normand C, 1990. Cost of surfactant replacement treatment for severe neonatal respiratory distress syndrome: a randomised controlled trial. *BMJ*;301:842–5.
- Tugwell P, Bombardier C, Bell M, Bennett K, Bensen W, Grace E, *et al.*, 1991. Current quality-of-life research challenges in arthritis relevant to the issue of clinical significance. *Control Clin Trials*;12:S217–25.
- Tuley MR, Mulrow CD, McMahan CA, 1991. Estimating and testing an index of responsiveness and the relationship of the index to power. *J Clin Epidemiol*;44:417–21.
- Tuomilehto J, Sarti C, Narva EV, Salmi K, Sivenius J, Kaarsalo E, *et al.*, 1992. The FINMONICA stroke register: community-based stroke registration and analysis of stroke incidence in Finland, 1983–1985. *Am J Epidemiol*;135:1259–70.
- Turner R, 1994. Quality of life: experience with sertraline. *J Brasil Psiquiat*;43:81–6S.
- Turner R, 1994. Quality-of-life: experience with sertraline. *Int Clin Psychopharmacol*;9:27–31.
- Turner RP, Lustig SP, 1994. Using quality-of-life measures to produce QALYs in clinical trial: a comparison of 3 methods [abstract]. *Qual Life Res*;3:55–6.
- Uchino J, Une Y, Sato Y, Gondo H, Nakajima Y, Sato N, 1993. Chemohormonal therapy of unresectable hepatocellular carcinoma. *Am J Clin Oncol: Cancer Clin Trials*;16:206–9.
- Ulin PR, Hardee K, Bailey P, Williamson N, 1994. The impact of family planning on women's lives: expanding the research agenda. *World Health Stat Q*;47:6–8.
- Unsworthwhite J, Kallis P, Treasure T, Pepper J.R, 1994. Quality-of-life after cardiac surgery in patients over 70 years of age. *Cardiol Elderly*;2:133–8.
- Urdal P, Haaland A, Hjerermann I, Gjesdal K, Christian C, Sorensen M, *et al.*, 1993. The Oslo Diet and Exercise Study (ODES): design and objectives. *Control Clin Trials*;14:229–43.
- Urquhart DS, McLatchie GR, Maffulli N, Hutchison JD, 1996. Quality adjusted sporting years: a method of outcome measurement. *Sports Exercise Injury*;2:43–5.
- Uther JFB, 1992. The automatic implantable defibrillator is the most realistic and cost-effective way of preventing sudden cardiac death. *Aust N Z J Med*;22:636–8.
- Uylde Groot CA, Hagenbeek A, Verdonck LF, Lowenberg B, Rutten FFH, 1995. Cost-effectiveness of ABMT in comparison with CHOP chemotherapy in patients with intermediate and high grade malignant non-Hodgkin's lymphoma (NHL). *Bone Marrow Transplant*;16:463–70.
- van der Donk J, Levendag PC, Kuijpers AJ, Roest FHJ, Habbema JDF, Meeuwis CA, *et al.*, 1995. Patient participation in clinical decision-making for treatment of T3 laryngeal cancer: a comparison of state and process utilities. *J Clin Oncol*;13:2369–78.
- van der Meulen JHP, Steyerberg EW, van der Graaf Y, van Herwerden LA, Verbaan CJ, de Fauw JJAMT, *et al.*, 1993. Age thresholds for prophylactic replacement of Bjork–Shiley convexo-concave heart valves: a clinical and economic evaluation. *Circulation*;88:156–64.
- van Hout B, Bonsel G, Habbema D, van der Maas P, de Charro F, 1993. Heart transplantation in The Netherlands: costs, effects and scenarios. *J Health Econ*;12:73–93.
- van Zonneveld RJ, 1987. The role of longitudinal surveys among the elderly in predicting disability and longevity. *Rev Epidemiol Sante Publique*;35:274–81.
- van Bruggen AC, Dippel DWJ, Habbema JDF, Mooij JJA, 1996. Acoustic detection of intracranial aneurysms: a decision analysis. *Acta Neurochir*;138:1148–56.
- van Damme P, Beutels P, 1996. Economic evaluation of vaccination. *PharmacoEconomics*;9:8–15.
- van Es GA, Lubsen J, Olsson G, Rehnqvist N, 1986. Quality of life as an outcome in long-term clinical trials. *Control Clin Trials*;7:253.
- van Hems P, Toma E, Pineault R, 1996. Quality-of-life assessment and HIV infection: a review. *Eur J Epidemiol*;12:221–8.

- van Knippenberg FCE, de Haes JCJM, 1988. Measuring the quality of life of cancer patients: psychometric properties of instruments. *J Clin Epidemiol*;41:1043–53.
- van Oostrom, MA, Tjihuis MAR, de Haes JCJM, Tempelaar R, Kromhout D, 1995. A measurement of social support in epidemiologic research: the social experiences checklist tested in a general population in the Netherlands. *J Epidemiol Community Health*;49:518–24.
- van Oyen H, Tafforeau J, Roelands M, 1996. Regional inequities in health expectancy in Belgium. *Soc Sci Med*;43:1673–8.
- van Zanten SJOV, Tytgat KMAJ, Pollak PT, Goldie J, Goodacre RL, Riddell RH, *et al.*, 1993. Can severity of symptoms be used as an outcome measure in trials of nonulcer dyspepsia and *Helicobacter pylori* associated gastritis? *J Clin Epidemiol*;46:273–9.
- Veatch RM, 1994. Health care rationing through global budgeting – the ethical choices. *J Clin Ethics*;5:291–6.
- Velanovich CPTV, 1990. Choice of treatment for stage I floor-of-mouth cancer. A decision analysis. *Arch Otolaryngol Head Neck Surg*;116:951–6.
- Velanovich V, 1995. Blunt splenic injury in adults: a decision analysis comparing options for treatment. *Eur J Surg*;161:463–70.
- Velanovich V, 1995. Immediate biopsy versus observation for abnormal findings on mammograms: an analysis of potential outcomes and costs. *Am J Surg*;170:327–32.
- Velanovich V, Tapper D, 1993. Decision analysis in children with blunt splenic trauma: the effects of observation, splenorrhaphy, or splenectomy on quality-adjusted life expectancy. *J Pediatr Surg*;28:179–85.
- Veldhuyzen SJO, van Zanten, 1991. Quality-of-life as outcome measures in randomized clinical trials: an overview of 3 general medical journals (reprinted from MA 1989; 140:1441-S). *Control Clin Trials*;12:S234–42.
- Verhoef LCG, Stalpers LJA, Verbeek ALM, Wobbes T, van Daal WAJ, 1991. Breast conserving treatment or mastectomy in early breast cancer: a clinical decision analysis with special reference to the risk of local recurrence. *Eur J Cancer*;27:1132–7.
- Vermeer F, Simoons ML, de Zwaan C, van Es GA, Verheugt FWA, van der Laarse A, *et al.*, 1988. Cost benefit analysis of early thrombolytic treatment with intracoronary streptokinase: twelve month follow up report of the randomised multicentre trial conducted by the Interuniversity Cardiology Institute of The Netherlands. *Br Heart J*;59:527–34.
- Visser MC, Fletcher AE, Parr G, Simpson A, Bulpitt CJ, 1994. A comparison of three quality of life instruments in subjects with angina pectoris: the Sickness Impact Profile, the Nottingham Health Profile, and the Quality of Well Being Scale. *J Clin Epidemiol*;47:157–63.
- Visser MC, Koudstaal PJ, Erdman RAM, Deckers JW, Passchier J, van Gijn J, *et al.*, 1995. Measuring quality-of-life in patients with myocardial infarction or stroke: a feasibility study of 4 questionnaires in The Netherlands. *J Epidemiol Community Health*;49:513–17.
- von Neumann J, Morgenstern O, 1953. Theory of games and economic behaviour. 3rd ed. New York: Wiley.
- von Korff M, Ustun TB, Ormel J, Kaplan I, Simon GE, 1996. Self-reported disability in an international primary care study of psychological illness. *J Clin Epidemiol*;49:297–303.
- Vorobeve SI, 1980. Autoimmune reactions as indicators of a harmful effect of industrial poisons (a review). *J Hygiene Epidemiol Microbiol Immunol*;24:489–94.
- Wade AG, Crawford RJ, Allman S, 1988. A comparative study of Frumil and Frusemide plus Slow K in the treatment of patients with congestive cardiac failure. *Clin Trials J*;25:37–42.
- Wade DT, 1991. The Q in QALYs. *BMJ*;303:1136–7.
- Wagstaff A, 1991. QALYs and the equity-efficiency trade-off [published erratum appears in *J Health Econ* 1993;12:237]. *J Health Econ*;10:21–41.
- Wakker P, 1996. A criticism of healthy-years equivalents. *Med Decis Making*;16:207–14.
- Wakker P, 1996. HYE: rejoinder. *Med Decis Making*;16:216.
- Wakker P, Stiggelbout A, 1995. Explaining distortions in utility elicitation through the rank-dependent model for risky choices. *Med Decis Making*;15:180–6.
- Wall BF, Russell JGB, 1988. The application of cost–utility analysis to radiological protection in diagnostic radiology. *J Radiol Prot*;8:221–9.
- Walpole I, Zubrick S, Pontre J, 1990. Is there a fetal effect with low to moderate alcohol use before or during pregnancy? *J Epidemiol Community Health*;44:297–301.
- Walsh P, 1993. Tetraplegics and the justice of resource allocation. *Paraplegia*;31:143–6.
- Wang-Clow F, Lange M, Laird NM, Ware JH, 1995. A simulation study of estimators for rates of change in longitudinal studies with attrition. *Stat Med*;14:283–97.
- Ware JH, Lipsitz S, Speizer FE, 1988. Issues in the analysis of repeated categorical outcomes. *Stat Med*;7:95–107.
- Warren JM, 1994. Rationing health care resources. *Can Fam Physician*;40:123.
- Weeks JC, Tierney MR, Weinstein MC, 1991. Cost-effectiveness of prophylactic intravenous immune globulin in chronic lymphocytic leukemia. *N Engl J Med*;325:81–6.
- Wei LJ, Stram DO, 1988. Analysing repeated measurements with possibly missing observations by modelling marginal distributions. *Stat Med*;7:139–48.



- Weinberger M, Oddone EZ, Samsa GP, Landsman PB, 1996. Are health-related quality-of-life measures affected by the mode of administration? *J Clin Epidemiol*; **49**:135–40.
- Weinstein MC, 1988. A QALY is a QALY is a QALY – or is it? *J Health Econ*; **7**:289–90.
- Weinstein MC, Coley CM, Richter JM, 1990. Medical management of gallstones: a cost-effectiveness analysis. *J Gen Intern Med*; **5**:277–84.
- Weinstein MC, Fineberg HV, 1980. Clinical decision analysis. Philadelphia: WB Saunders.
- Weinstein MC, Pliskin JS, 1996. Perspectives on healthy-years equivalents. HYE: what are the issues? *Med Decis Making*; **16**:205–6.
- Weinstein MC, Stason WB, 1982. Cost-effectiveness of coronary artery bypass surgery. *Circulation*; **66**:III-56–66.
- Weisbrod BA, 1992. Productivity and incentives in the medical care sector. *Scand J Econ*; **94**:S131–45.
- Weissfeld JL, Holloway JJ, 1992. Precision of blood cholesterol measurement and high blood cholesterol case finding and treatment. *J Clin Epidemiol*; **45**:971–84.
- Wentraub WS, Mauldin PD, Becker ER, 1994. An introduction to clinical microeconomic analysis: purposes and analytic methods. *J Interventional Cardiol*; **7**:281–9.
- Whalen CC, Antani M, Carey J, Landefeld CS, 1994. An index of symptoms for infection with human immunodeficiency virus: reliability and validity. *J Clin Epidemiol*; **47**:537–46.
- Whittington R, 1995. Alglucerase: a pharmacoeconomic appraisal of its use in the treatment of Gaucher disease. *PharmacoEconomics*; **7**:63–90.
- Whittington R, Faulds D, 1994. Hormone replacement therapy. 1. Pharmacoeconomic appraisal of its therapeutic use in menopausal symptoms and urogenital estrogen deficiency. *PharmacoEconomics*; **5**:419–45.
- Whittington R, Faulds D, 1994. Hormone replacement therapy. 2. A pharmacoeconomic appraisal of its role in the prevention of postmenopausal osteoporosis and ischemic heart disease. *PharmacoEconomics*; **5**:513–54.
- Whynes DK, Neilson AR, Robinson MHE, Hardcastle JD, 1994. Colorectal cancer screening and quality of life. *Qual Life Res*; **3**:191–8.
- Whynes DK, Walker AR, Hardcastle JD, 1992. Effect of subject age on costs of screening for colorectal cancer. *J Epidemiol Community Health*; **46**:577–81.
- Wiener Z, Abelson A, Herman J, 1989. The diagnostic process: some unwitting uses. *J Clin Epidemiol*; **42**:377–9.
- Wiersma D, 1996. Measuring social disabilities in mental health. *Soc Psychiatry Psychiat Epidemiol*; **31**:101–8.
- Wiesner PJ, 1993. Four diseases of disarray in public health. *Ann Epidemiol*; **3**:196–8.
- Wiklund I, Dimenas E, Wahl M, 1990. Factors of importance when evaluating quality of life in clinical trials. *Control Clin Trials*; **11**:169–79.
- Wiklund I, Karlberg J, 1991. Evaluation of quality-of-life in clinical trials: selecting quality-of-life measures. *Control Clin Trials*; **12**:S204–16.
- Wilkins R, Adams OB, 1983. Health expectancy in Canada, late 1970s: demographic, regional, and social dimensions. *Am J Public Health*; **73**:1073–80.
- Wilkins R, Adams OB, 1987. Changes in the healthfulness of life of the elderly population: an empirical approach. *Rev Epidemiol Sante Publique*; **35**:225–35.
- Wilkinson G, Williams B, Krekorian H, McLees S, Falloon I, 1992. QALYs in mental health: a case study. *Psychol Med*; **22**:725–31.
- Willemse PHB, de Vries EGE, Kloppenburg M, Fontein DL, Aalders JG, Boonstra H, et al., 1992. Carboplatin with cyclophosphamide in patients with advanced ovarian cancer: an efficacy and quality-adjusted survival analysis. *Int J Gynecol Cancer*; **2**:236–43.
- Willemse PHB, van Lith J, Mulder NH, Aalders JG, Bouma J, de Vries EGE, et al., 1990. Risks and benefits of cisplatin in ovarian cancer: a quality-adjusted survival analysis. *Eur J Cancer*; **26**:345–52.
- Williams A, 1985. Economics of coronary artery bypass grafting. *BMJ*; **291**:326–9.
- Williams A, 1986. QALYs or short straws? *BMJ*; **293**:337–8.
- Williams A, 1987. Quality-adjusted life-years. *Lancet*; **i**:1372.
- Williams A, 1987. QALYfying the value of life [response]. *J Med Ethics*; **13**:123.
- Williams A, 1990. Should QALYs be program specific? *J Health Econ*; **8**:485–7.
- Williams A, 1990. 'Should QALYs be programme specific?' by Donaldson, Atkinson, Bond and Wright. *J Health Econ*; **8**:485–7.
- Williams A, 1991. Is the QALY a technical solution to a political problem? – of course not. *Int J Health Serv*; **21**:365–9.
- Williams A, 1991. The role of health economics in clinical decision making: is it ethical? *Respir Med*; **85**:3–5.
- Williams A, 1992. Cost-effectiveness analysis: is it ethical? *J Med Ethics*; **18**:7–11.
- Williams A, 1994. Economics, QALYs and medical ethics: a health economist's perspective. Discussion Paper 121. York: Centre for Health Economics, University of York.
- Williams A, 1995. The role of the EuroQol instrument in QALY calculations. Discussion Paper 130. York: Centre for Health Economics, University of York.
- Williams A, 1995. Economics, QALYs and medical ethics: a health economist's perspective. *Health Care Anal*; **3**:221–6.

- Williams A, 1996. QALYs and ethics: a health economist's perspective. *Soc Sci Med*;43:1795–804.
- Williams JBW, Rabkin JG, 1991. The concurrent validity of items in the quality-of-life index in a cohort of HIV positive and HIV negative gay men. *Control Clin Trials*;12:S129–41.
- Williams PL, Portier CJ, 1992. Explicit solutions for constrained maximum-likelihood estimators in survival sacrifice experiments. *Biometrika*;79:717–29.
- Wilson A, Wiklund I, Lahti T, Wahl M, 1991. A summary index for the assessment of quality-of-life in angina pectoris. *J Clin Epidemiol*;44:981–8.
- Wilson TW, Chockalingam A, Quest DW, 1996. Pharmacoeconomics of hypertension control: basic principles of economic evaluation. *J Hum Hypertens*;10:S19–22.
- Wimo A, Mattson B, Krakau I, Eriksson T, Nelvig A, Karlsson G, 1995. Cost–utility analysis of group living in dementia care. *Int J Technol Assess Health Care*;11:49–65.
- Winefield AH, Tiggemann M, Goldney RD, 1988. Psychological concomitants of satisfactory employment and unemployment in young people. *Soc Psychiatry Psychiat Epidemiol*;23:149–57.
- Winer BJ, 1971. Statistical principles in experimental design. New York: McGraw-Hill.
- Winkler R, 1981. Preoperative irradiation from the surgeon's viewpoint. *Strahlentherapie*;157:159–65.
- Wittmann MM, Wittmann A, Wittmann DH, 1996. AIDS, emergency operations, and infection control. *Infect Control Hosp Epidemiol*;17:532–8.
- Wolf FM, Jacober SJ, Wolf LL, Cornell RG, Floyd JC, 1989. Quality of life activities associated with adherence to insulin infusion pump therapy in the treatment of insulin dependent diabetes mellitus. *J Clin Epidemiol*;42:1129–36.
- Wong JB, Koff RS, Tine F, Pauker SG, 1995. Cost-effectiveness of interferon-alpha-2b treatment for hepatitis B antigen-positive chronic hepatitis B. *Ann Intern Med*;122:664–75.
- Wong JB, Sonnenberg FA, Salem DN, Pauker SG, 1990. Myocardial revascularization for chronic stable angina. Analysis of the role of percutaneous transluminal coronary angioplasty based on data available in 1989. *Ann Intern Med*;113:852–71.
- Woodbury MA, Manton KG, Yashin AI, 1988. Estimating hidden morbidity via its effect on mortality and disability. *Stat Med*;7:325–36.
- World Health Organization, 1947. The constitution of the World Health Organization. *WHO Chron*;1:29.
- Wortman PM, Yeaton WH, 1985. Cumulating quality of life results in controlled trials of coronary artery bypass graft surgery. *Control Clin Trials*;6:289–305.
- Wright J, Jones G, Whelan T, Lukka H, 1994. Patient preference for high or low dose rate brachytherapy in carcinoma of the cervix. *Radiother Oncol*;33:187–94.
- Wu MC, Bailey K, 1988. Analysing changes in the presence of informative right censoring caused by death and withdrawal. *Stat Med*;7:337–46.
- Wu MC, Bailey KR, 1989. Estimation and comparison of changes in the presence of informative right censoring: conditional linear model. *Biometrics*;45:939–55.
- Wu MC, Carroll RJ, 1988. Estimation and comparison of changes in the presence of informative right censoring by modelling the censoring process. *Biometrics*;44:175–88.
- Wu SC, 1982. A semi-Markov model for survival data with covariates. *Math Biosci*;60:197–206.
- Wyatt GKH, Friedman LL, 1996. Development and testing of a quality of life model for long-term female cancer survivors. *Qual Life Res*;5:387–94.
- Wyatt JR, Niparko JK, Rothman M, deLissovoy G, 1996. Cost–utility of the multichannel cochlear implant in 258 profoundly deaf individuals. *Laryngoscope*;106:816–21.
- Wyatt JR, Niparko JK, Rothman ML, deLissovoy G, 1995. Cost-effectiveness of the multichannel cochlear implant. *Am J Otol*;16:52–62.
- Yin DP, Baum RA, Carpenter JP, Langlotz CP, Pentecost MJ, 1995. Cost-effectiveness of MR angiography in cases of limb threatening peripheral vascular disease. *Radiology*;194:757–64.
- Yin DP, Forman HP, Langlotz CP, 1995. Evaluating health services: the importance of patients preferences and quality-of-life. *Am J Roentgenol*;165:1323–8.
- Zbrozek AS, Cantor SB, Cardenas MP, Hill DP, 1994. Pharmacoeconomic analysis of ondansetron versus metoclopramide for cisplatin-induced nausea and vomiting. *Am J Hosp Pharmacy*;51:1555–63.
- Zbrozek AS, Cantor SB, Hill DP, 1994. Reexamination of clinical aspects of pharmacoeconomic analysis. *Am J Hosp Pharmacy*;51:2972–3.
- Zeger SL, 1988. Repeated categorical response – commentary. *Stat Med*;7:161–8.
- Zeger SL, Liang K, 1986. Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*;42:121–30.
- Zeger SL, Liang K, 1992. An overview of methods for the analysis of longitudinal data. *Stat Med*;11:1825–39.
- Zeger SL, Liang K, Albert PS, 1988. Models for longitudinal data: a generalised estimating equation approach. *Biometrics*;44:1049–60.
- Zesch A, 1989. Disinfectants as registered drugs. *Hygiene Med*;14:135–40.
- Zhang J, Quan H, Ng J, Stepanavage ME, 1997. Some statistical methods for multiple endpoints in clinical trials. *Control Clin Trials*;18:204–21.
- Zhang ML, Owen RR, Pope SK, Smith GR, 1996. Cost-effectiveness of clozapine monitoring after the first 6 months. *Arch Gen Psychiatry*;53:954–8.

Ziegler F, 1996. Cost-utility in CNS drug trials. *Eur Psychiatry*;11:159-64.

Zuckerman AJ, 1995. Occupational exposure to hepatitis B virus and Human Immunodeficiency Virus: a comparative risk analysis. *Am J Infect Control*;23:286-9.

Zuckerman DM, Kasl SV, Ostfeld AM, 1984. Psychosocial predictors of mortality among the elderly poor: the role of religion, well-being, and social contacts. *Am J Epidemiol*;119:410-23.

Zwetsloot-Schonk JHM, Hermans J, van Oosterom AT, Pliskin JS, Zwartendijk J, Leer JWH, 1988. Decision analysis in carcinoma of the testis. *Theor Surg*;2:176-84.

Zwinderman AH, 1990. The measurement of change of quality of life in clinical trials. *Stat Med*;9:931-42.

Zwinderman AH, 1992. Statistical analysis of longitudinal quality of life data with missing measurements. *Qual Life Res*;1:219-24.



# Appendix I

## Details of search terms used in electronic database searching

### Notes on searching in BIDS

- Putting an asterisk at the end of a search term allows any ending to the word to be searched for on the database.
- Search terms consisting of hyphenated words need to be enclosed in double quotes otherwise the hyphen will be interpreted as an exclusion sign.
- Using a comma between two search terms allows the database to be searched for articles containing either term.

### Selecting type of journal

The following criteria were used in BIDS to restrict the search to just the statistical, biometry, clinical trial and epidemiology journals.

journal = statistic\*, biometri\*, clinical trial\*, epidemiol\*

### Quality-of-life search terms

The search terms used in BIDS to select articles containing some information of interest on quality of life are shown in *Table 30*.

### Survival search terms

The search terms used in BIDS to select articles with some interest in survival are shown in *Table 31*.

**TABLE 30** Search terms: quality of life

Search terms	Comments
quality of life, "quality-of-life"	Also captures more specific phrases such as health-related quality-of-life.
life quality, "life-quality"	
quality life	Also captures articles where a non-standard hyphenated version has been used, e.g. quality-of life, quality-of-life, quality- of-life, quality of-life.
"well-being", wellbeing, well being	

**TABLE 31** Search terms: survival

Search terms	Comments
survival	Captures a whole range of survival-type phrases, e.g. survival time, survival analysis, survival date, survival endpoints, survival rate, survival curve, survival probability, disease-free survival, event-free survival.
length of life	
relapse	Also captures more specific terms such as time to relapse.
failure time*, "failure-time*"	* captures both time and times and also hyphenated phrases such as failure-time-data.
time to death, time to recurrence, time to progression, time to disease, time to failure	Also captures more specific phrases such as time to disease progression, time to disease recurrence.
life table*, "life-table*"	* captures both table and tables and also hyphenated phrases such as life-table-analysis.
Kaplan Meier, "Kaplan-Meier"	
log rank, "log-rank"	
proportional hazard*, "proportional-hazard*"	* captures both hazard and hazards and also hyphenated phrases such as proportional-hazards-analysis.
Cox* + (model*, analys*, regression*, test*, method*, proportional*)	Cox* allows for various versions such as Cox, Coxs or Cox's. Cox* can not be used as a search term on its own because it lacks precision. The specification captures phrases such as Cox multi-variate regression or multiple regression model of Cox.
event histor*, "event-histor*"	* captures both both history and histories and also hyphenated phrases such as event-history-analysis.
censor*	* captures censor, censored and censoring and thus captures more specific phrases such as informative censoring.

## Methodology search terms

The search terms used in BIDS to select articles which used a known methodology for the simultaneous analysis of quality-of-life and survival data are shown in *Table 32*.

**TABLE 32** Search terms: known methodology for simultaneous analysis of quality-of-life and survival data

Search terms	Comments
quality adjusted*, "quality-adjusted*"	* captures hyphenated phrases such as ...-life, ...-life-years, ...-years. Captures other phrases such as ... life, ... life years, ... years of life, ... life expectancy, ... loss in life expectancy, ... survival, ... days, ... time, ... tooth years.
quality of life adjusted*, "quality-of-life adjusted*", "quality-of-life-adjusted*"	* captures hyphenated phrases.
QALY*, QALD*, QALE*, QTIME, "Q-TIME", QTWIST, "Q-TWIST", HYE, HYES, HYE'S	* allows for plurals. Could not use TWIST as a search term because it lacked precision. Had to specify HYE, HYES, HYE'S rather than HYE* because HYE* lacked precision.
quality survival time, "quality-survival time", "quality survival-time", "quality-survival-time"	
healthy life expectancy, "healthy-life expectancy", "healthy life-expectancy", "healthy-life-expectancy"	
healthy years equivalent*, "healthy-years equivalent*", "healthy year-equivalent*", "healthy-years-equivalent*"	* captures equivalent and equivalence.
healthy year equivalent*, "healthy-year equivalent*", "healthy year-equivalent*", "healthy-year-equivalent*"	* captures equivalent and equivalence.
("multi-state", multistate) + (survival, censor*)	Captures articles using multistate survival analysis.

## Longitudinal search terms

The search terms used in BIDS to select articles describing longitudinal studies (which were more relevant to this review) are shown in *Table 33*.

**TABLE 33** Search terms: longitudinal studies

Search terms	Comments
longitudinal	
repeated measure*, "repeated-measure*", repeated assessment*, "repeated-assessment*"	Measure* captures measure, measures, measurement or measurements. Assessment* captures assessment or assessments.
serial measure*, "serial-measure*", serial assessment*, "serial-assessment*"	Measure* captures measure, measures, measurement or measurements. Assessment* captures assessment or assessments.
measure* over time, assess* over time	Measure* captures measure, measures, measurement or measurements. Assess* captures assessment or assessments.
profile, profiles	Profile* could not be used because it had a lack of precision.
dynamic	

# Health Technology Assessment panel membership

This report was identified as a priority by the Methodology Panel.

## Acute Sector Panel

### Current members

<b>Chair:</b> <b>Professor Francis H Creed,</b> University of Manchester	Dr Katherine Darton, M.I.N.D. Mr John Dunning, Papworth Hospital, Cambridge	Ms Grace Gibbs, West Middlesex University Hospital NHS Trust	Dr Duncan Keeley, General Practitioner, Thame
Professor Clifford Bailey, University of Leeds	Mr Jonathan Earnshaw, Gloucester Royal Hospital	Dr Neville Goodman, Southmead Hospital Services Trust, Bristol	Dr Rajan Madhok, East Riding Health Authority
Ms Tracy Bury, Chartered Society of Physiotherapy	Mr Leonard Fenwick, Freeman Group of Hospitals, Newcastle-upon-Tyne	Professor Mark P Haggard, MRC	Dr John Pounsford, Frenchay Hospital, Bristol
Professor Collette Clifford, University of Birmingham	Professor David Field, Leicester Royal Infirmary	Professor Robert Hawkins, University of Manchester	Dr Mark Sculpher, University of York
			Dr Iqbal Sram, NHS Executive, North West Region

### Past members

Professor John Farndon, University of Bristol*	Professor Cam Donaldson, University of Aberdeen	Mrs Wilma MacPherson, St Thomas's & Guy's Hospitals, London	Professor Michael Sheppard, Queen Elizabeth Hospital, Birmingham
Professor Senga Bond, University of Newcastle- upon-Tyne	Professor Richard Ellis, St James's University Hospital, Leeds	Dr Chris McCall, General Practitioner, Dorset	Professor Gordon Stirrat, St Michael's Hospital, Bristol
Professor Ian Cameron, Southeast Thames Regional Health Authority	Mr Ian Hammond, Bedford & Shires Health & Care NHS Trust	Professor Alan McGregor, St Thomas's Hospital, London	Dr William Tarnow-Mordi, University of Dundee
Ms Lynne Clemence, Mid-Kent Health Care Trust	Professor Adrian Harris, Churchill Hospital, Oxford	Professor Jon Nicholl, University of Sheffield	Professor Kenneth Taylor, Hammersmith Hospital, London
	Dr Gwyneth Lewis, Department of Health	Professor John Norman, University of Southampton	

## Diagnostics and Imaging Panel

### Current members

<b>Chair:</b> <b>Professor Mike Smith,</b> University of Leeds	Dr Barry Cookson, Public Health Laboratory Service, Colindale	Mrs Maggie Fitchett, Association of Cytogeneticists, Oxford	Professor Chris Price, London Hospital Medical School
Dr Philip J Ayres, Leeds Teaching Hospitals NHS Trust	Professor David C Cumberland, University of Sheffield	Dr Peter Howlett, Portsmouth Hospitals NHS Trust	Dr William Rosenberg, University of Southampton
Dr Paul Collinson, Mayday University Hospital, Thornton Heath	Professor Adrian Dixon, University of Cambridge	Professor Alistair McGuire, City University, London	Dr Gillian Vivian, Royal Cornwall Hospitals Trust
	Mr Steve Ebdon-Jackson, Department of Health	Dr Andrew Moore, Editor, <i>Bandolier</i>	Dr Greg Warner, General Practitioner, Hampshire
		Dr Peter Moore, Science Writer, Ashtead	

### Past members

Professor Michael Maisey, Guy's & St Thomas's Hospitals, London*	Professor MA Ferguson-Smith, University of Cambridge	Professor Donald Jeffries, St Bartholomew's Hospital, London	Professor John Stuart, University of Birmingham
Professor Andrew Adam, Guy's, King's & St Thomas's School of Medicine & Dentistry, London	Dr Mansel Hacney, University of Manchester	Dr Ian Reynolds, Nottingham Health Authority	Dr Ala Szczepura, University of Warwick
Dr Pat Cooke, RDRD, Trent Regional Health Authority	Professor Sean Hilton, St George's Hospital Medical School, London	Professor Colin Roberts, University of Wales College of Medicine	Mr Stephen Thornton, Cambridge & Huntingdon Health Commission
Ms Julia Davison, St Bartholomew's Hospital, London	Mr John Hutton, MEDTAP International Inc., London	Miss Annette Sergeant, Chase Farm Hospital, Enfield	Dr Jo Walsworth-Bell, South Staffordshire Health Authority

\* Previous Chair  
*continued*

continued

## Methodology Panel

### Current members

<b>Chair:</b> <b>Professor Martin Buxton,</b> Brunel University	Professor Ann Bowling, University College London Medical School	Professor Jeremy Grimshaw, University of Aberdeen	Dr Nick Payne, University of Sheffield
Professor Doug Altman, Institute of Health Sciences, Oxford	Dr Mike Clarke, University of Oxford	Dr Stephen Harrison, University of Leeds	Professor Margaret Pearson, NHS Executive North West
Dr David Armstrong, Guy's, King's & St Thomas's School of Medicine & Dentistry, London	Professor Michael Drummond, University of York	Mr John Henderson, Department of Health	Professor David Sackett, Centre for Evidence Based Medicine, Oxford
Professor Nick Black, London School of Hygiene & Tropical Medicine	Dr Vikki Entwistle, University of Aberdeen	Professor Richard Lilford, Regional Director, R&D, West Midlands	Dr PAG Sandercock, University of Edinburgh
	Professor Ewan Ferlie, Imperial College, London	Professor Theresa Marteau, Guy's, King's & St Thomas's School of Medicine & Dentistry, London	Dr David Spiegelhalter, Institute of Public Health, Cambridge
	Professor Ray Fitzpatrick, University of Oxford	Dr Henry McQuay, University of Oxford	Professor Joy Townsend, University of Hertfordshire

### Past members

Professor Anthony Culyer, University of York*	Professor George Davey-Smith, University of Bristol	Mr Nick Mays, King's Fund, London	Professor Charles Warlow, Western General Hospital, Edinburgh
Professor Michael Baum, Royal Marsden Hospital	Professor Stephen Frankel, University of Bristol	Professor Ian Russell, University of York	
Dr Rory Collins, University of Oxford	Mr Philip Hewitson, Leeds FHSA	Dr Maurice Slevin, St Bartholomew's Hospital, London	

## Pharmaceutical Panel

### Current members

<b>Chair:</b> <b>Professor Tom Walley,</b> University of Liverpool	Professor Rod Griffiths, NHS Executive West Midlands	Dr Andrew Mortimore, Southampton & SW Hants Health Authority	Dr Frances Rotblat, Medicines Control Agency
Dr Felicity Gabbay, Transcrip Ltd	Mrs Jeanette Howe, Department of Health	Mr Nigel Offen, Essex Rivers Healthcare, Colchester	Dr Eamonn Sheridan, St James's University Hospital, Leeds
Mr Peter Golightly, Leicester Royal Infirmary	Professor Trevor Jones, ABPI, London	Mrs Marianne Rigge, The College of Health, London	Mrs Katrina Simister, Liverpool Health Authority
Dr Alastair Gray, Health Economics Research Unit, University of Oxford	Ms Sally Knight, Lister Hospital, Stevenage	Mr Simon Robbins, Camden & Islington Health Authority, London	Dr Ross Taylor, University of Aberdeen

### Past members

Professor Michael Rawlins, University of Newcastle- upon-Tyne*	Ms Christine Clark, Hope Hospital, Salford	Dr Tim Elliott, Department of Health	Dr John Posnett, University of York
Dr Colin Bradley, University of Birmingham	Mrs Julie Dent, Ealing, Hammersmith & Hounslow Health Authority, London	Dr Desmond Fitzgerald, Mere, Bucklow Hill, Cheshire	Dr Tim van Zwanenberg, Northern Regional Health Authority
Professor Alasdair Breckenridge, RDRD, Northwest Regional Health Authority	Mr Barrie Dowdeswell, Royal Victoria Infirmary, Newcastle-upon-Tyne	Professor Keith Gull, University of Manchester	Dr Kent Woods, RDRD, Trent RO, Sheffield
		Dr Keith Jones, Medicines Control Agency	



## Population Screening Panel

### Current members

<b>Chair:</b> <b>Professor Sir John Grimley Evans,</b> Radcliffe Infirmary, Oxford	Professor Howard Cuckle, University of Leeds	Professor Dian Donnai, St Mary's Hospital, Manchester	Professor Alexander Markham, St James's University Hospital, Leeds
Ms Stella Burnside, Altnagelvin Hospitals Trust, Londonderry	Dr Carol Dezateux, Institute of Child Health, London	Dr Tom Fahey, University of Bristol	Dr Ann McPherson, General Practitioner, Oxford
Mr John Cairns, University of Aberdeen	Dr Anne Dixon Brown, NHS Executive, Anglia & Oxford	Mrs Gillian Fletcher, National Childbirth Trust	Dr Susan Moss, Institute of Cancer Research
		Dr JA Muir Gray, Institute of Health Sciences, Oxford	Dr Sarah Stewart-Brown, University of Oxford

### Past members

Dr Sheila Adam, Department of Health*	Dr Anne Ludbrook, University of Aberdeen	Professor Catherine Peckham, Institute of Child Health, London	Professor Nick Wald, University of London
Professor George Freeman, Charing Cross & Westminster Medical School, London	Professor Theresa Marteau, Guy's, King's & St Thomas's School of Medicine & Dentistry, London	Dr Connie Smith, Parkside NHS Trust, London	Professor Ciaran Woodman, Centre for Cancer Epidemiology, Manchester
Dr Mike Gill, Brent & Harrow Health Authority		Ms Polly Toynbee, Journalist	

## Primary and Community Care Panel

### Current members

<b>Chair:</b> <b>Dr John Tripp,</b> Royal Devon & Exeter Healthcare NHS Trust	Ms Judith Brodie, Age Concern, London	Mr Andrew Farmer, Institute of Health Sciences, Oxford	Dr Chris McCall, General Practitioner, Dorset
Mr Kevin Barton, East London & City Health Authority	Mr Shaun Brogan, Daventry & South Northants Primary Care Alliance	Professor Richard Hobbs, University of Birmingham	Dr Robert Peveler, University of Southampton
Professor John Bond, University of Newcastle- upon-Tyne	Mr Joe Corkill, National Association for Patient Participation	Professor Allen Hutchinson, University of Sheffield	Professor Jennie Popay, University of Salford
Dr John Brazier, University of Sheffield	Dr Nicky Cullum, University of York	Dr Phillip Leech, Department of Health	Ms Hilary Scott, Tower Hamlets Healthcare NHS Trust, London
	Professor Pam Enderby, University of Sheffield	Dr Aidan Macfarlane, Oxfordshire Health Authority	Dr Ken Stein, North & East Devon Health Authority

### Past members

Professor Angela Coulter, King's Fund, London*	Professor Andrew Haines, RDRD, North Thames Regional Health Authority	Mr Lionel Joyce, Chief Executive, Newcastle City Health NHS Trust	Professor Dianne Newham, King's College London
Professor Martin Roland, University of Manchester*	Dr Nicholas Hicks, Oxfordshire Health Authority	Professor Martin Knapp, London School of Economics & Political Science	Professor Gillian Parker, University of Leicester
Dr Simon Allison, University of Nottingham	Mr Edward Jones, Rochdale FHSA	Professor Karen Luker, University of Liverpool	Dr Mary Renfrew, University of Oxford
Professor Shah Ebrahim, Royal Free Hospital, London	Professor Roger Jones, Guy's, King's & St Thomas's School of Medicine & Dentistry, London	Dr Fiona Moss, Thames Postgraduate Medical & Dental Education	

## National Coordinating Centre for Health Technology Assessment, Advisory Group

### Current members

---

**Chair:**

**Professor John Gabbay,**  
Wessex Institute  
for Health Research  
& Development

Ms Lynn Kerridge,  
Wessex Institute for Health  
Research & Development

Professor James Raftery,  
Health Economics Unit,  
University of Birmingham

Professor Andrew  
Stevens,  
Department of Public  
Health & Epidemiology,  
University of Birmingham

Dr Ruairidh Milne,  
Wessex Institute for Health  
Research & Development

Professor Ian Russell,  
Department of Health Sciences  
& Clinical Evaluation,  
University of York

Professor Mike  
Drummond,  
Centre for Health Economics,  
University of York

Ms Kay Pattison,  
Research & Development  
Directorate, NHS Executive

Dr Ken Stein,  
North & East Devon  
Health Authority

### Past member

---

Dr Paul Roderick,  
Wessex Institute for Health  
Research & Development

# HTA Commissioning Board

## Current members

---

**Chair:**

**Professor Charles Florey**,  
Department of Epidemiology &  
Public Health, Ninewells  
Hospital & Medical School,  
University of Dundee

**Professor Doug Altman**,  
Director of ICRF/NHS Centre  
for Statistics in Medicine,  
Oxford

**Professor John Bond**,  
Professor of Health Services  
Research, University of  
Newcastle-upon-Tyne

**Mr Peter Bower**,  
Independent Health Advisor,  
Newcastle-upon-Tyne

**Ms Christine Clark**,  
Honorary Research Pharmacist,  
Hope Hospital, Salford

**Professor Shah Ebrahim**,  
Professor of Epidemiology  
of Ageing, University of Bristol

**Professor Martin Eccles**,  
Professor of  
Clinical Effectiveness,  
University of Newcastle-  
upon-Tyne

**Dr Mike Gill**,  
Director of Public Health &  
Health Policy, Brent & Harrow  
Health Authority

**Dr Alastair Gray**,  
Director, Health Economics  
Research Centre,  
University of Oxford

**Professor Mark Haggard**,  
MRC Institute of  
Hearing Research

**Dr Jenny Hewison**,  
Senior Lecturer,  
Department of Psychology,  
University of Leeds

**Professor Sir Miles Irving**  
(Programme Director),  
Professor of Surgery,  
University of Manchester,  
Hope Hospital, Salford

**Professor Alison Kitson**,  
Director, Royal College of  
Nursing Institute

**Dr Donna Lamping**,  
Senior Lecturer, Department of  
Public Health, London School  
of Hygiene & Tropical Medicine

**Professor Alan Maynard**,  
Professor of Economics,  
University of York

**Professor Jon Nicholl**,  
Director, Medical Care  
Research Unit,  
University of Sheffield

**Professor Gillian Parker**,  
Nuffield Professor of  
Community Care,  
University of Leicester

**Dr Tim Peters**,  
Reader in Medical Statistics,  
Department of Social Medicine,  
University of Bristol

**Professor Martin Severs**,  
Professor in Elderly  
Health Care,  
Portsmouth University

**Dr Sarah Stewart-Brown**,  
Director, Institute of  
Health Sciences,  
University of Oxford

**Professor Ala Szczepura**,  
Director, Centre for  
Health Services Studies,  
University of Warwick

**Dr Gillian Vivian**,  
Consultant, Royal Cornwall  
Hospitals Trust

**Professor Graham Watt**,  
Department of General Practice,  
Woodside Health Centre,  
Glasgow

**Professor Kent Woods**,  
Regional Director of R&D  
NHS Executive, Trent

**Dr Jeremy Wyatt**,  
Senior Fellow, Health &  
Public Policy, School of Public  
Policy, University College,  
London

## Past members

---

**Professor Ian Russell**,  
Department of Health  
Sciences & Clinical Evaluation,  
University of York\*

**Professor David Cohen**,  
Professor of Health Economics,  
University of Glamorgan

**Mr Barrie Dowdeswell**,  
Chief Executive,  
Royal Victoria Infirmary,  
Newcastle-upon-Tyne

**Dr Michael Horlington**,  
Head of Corporate Licensing,  
Smith & Nephew Group  
Research Centre

**Professor Martin Knapp**,  
Director, Personal Social  
Services Research Unit,  
London School of Economics  
& Political Science

**Professor Theresa Marteau**,  
Director, Psychology & Genetics  
Research Group, Guy's, King's  
& St Thomas's School of  
Medicine & Dentistry,  
London

**Professor Sally McIntyre**,  
MRC Medical Sociology Unit,  
Glasgow

**Professor David Sackett**,  
Centre for Evidence Based  
Medicine, Oxford

**Dr David Spiegelhalter**,  
MRC Biostatistics Unit,  
Institute of Public Health,  
Cambridge

**Professor David Williams**,  
Department of  
Clinical Engineering,  
University of Liverpool

**Dr Mark Williams**,  
Public Health Physician,  
Bristol

\* Previous Chair

---

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,  
Mailpoint 728, Boldrewood,  
University of Southampton,  
Southampton, SO16 7PX, UK.  
Fax: +44 (0) 1703 595 639 Email: [hta@soton.ac.uk](mailto:hta@soton.ac.uk)  
<http://www.hta.nhsweb.nhs.uk>

ISSN 1366-5278