Supplemental Content

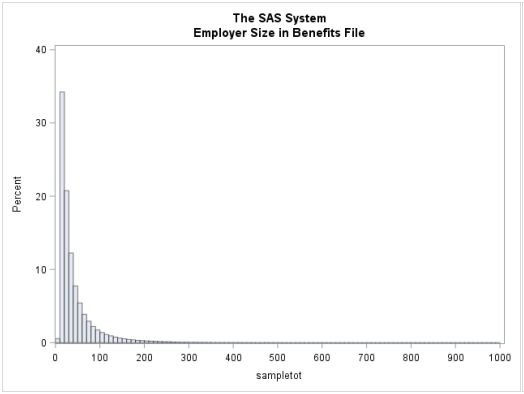
Approaches

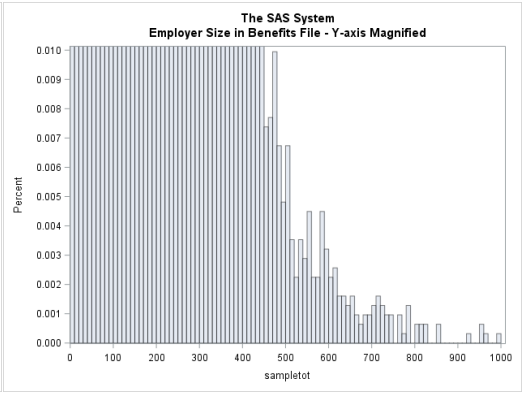
A. Deductible Imputation Algorithm

To determine employer deductible levels, we used a benefits type variable that we had for most smaller employers (with approximately 100 or fewer employees). For larger employers, we took advantage of the fact that health insurance claims data are the most accurate source for assessing out-of-pocket obligations among patients who utilize health services. Our claims data contained an in-network/out-of-network individual deductible payment field. For patients who use expensive or frequent services, the sum of their yearly deductible payments adds up to clearly identifiable exact amounts such as $500.00, $1000.00, $2000.00, etc. When even several members have these same amounts, it provides strong evidence that the employer offered such an annual deductible level. It is also possible to detect employers that offer choices of deductible levels when multiple employees have deductibles at two or more levels, such as 20 employees with an exact annual amount of $1000.00 and 12 employees with $500.00. For employer accounts with at least 10 enrollees, we therefore summed each member's in-network (individual-level) deductible payments and number of claims over the enrollment year and assessed other key characteristics such as percentage with Health Savings Accounts. We randomly selected half of the employer account data set that contained both our calculated employer characteristics (independent variables, below) and actual annual deductible levels from the benefits table (dependent variable, after categorization; below). We then used a multinomial logistic model that predicted the 4-level outcome of individual-level deductible ≤$500/$500-$999/$1000-$2499/≥$2500 (again, dependent variable) based on multiple aggregate employer characteristics (independent variables) such as the percentage with Health Savings Accounts and Health Reimbursement Arrangements, the deductible payment per employer in the 75 percentile of payments, the percentage of employees reaching exact deductible levels or with deductible payments but not reaching an exact deductible level, the employer account size, the percentage of enrollees per account with summed whole dollar annual deductible amounts (from claims data) between $0 to <$100, ≥$100 to ≤$500, >$500 to <$1000, ≥$1000 to <$2500, ≥$2500, etc.

This predictive model output the probability that employers had deductibles in the four categories (summing to 1.0) and we assigned the employer to the level that had the highest probability. We overwrote this assignment with the most common whole number deductible amount per year if it was not zero, and with the second most common whole number deductible amount if the most common amount was zero and at least 10 members had the value of the second most common whole number deductible amount. If an employer had members with both enrollment and evidence of utilization, but never had any amounts in the deductible field, we assigned that employer to <$500 deductible level. If an employer had only members that reached a whole number annual deductible amount such as $1000.00 or $2000.00, we assigned the most common deductible amount as the employer's deductible if that amount was greater than or equal to $1000 and to the 95% percentile value if that number was less than $1000. If at least 99% of employees had Health Savings Accounts or Health Reimbursement Arrangements, we also overwrote any previous assignment to classify the employer as a high-deductible employer. We assigned employers to have a choice between deductible levels of $1000 to $2499 and ≥$2500 when both were common and one accounted for at least 85% of $1000-$2499 or ≥$2500 deductible levels reached per employer. If we detected employers that had sufficient enrollees with whole number deductible levels both above and below $1000 (e.g. $250.00 and $1500.00), we assigned the employers' category as "choice," applying a similar 85% rule. Finally, for any employer that had gold standard deductible level information in our benefits file, we overwrote any previous imputed deductible level.

Our file that contains actual deductible amounts per employer covers the "small employer" segment of the insurer's business, a segment that generally includes employers with fewer than 100 or so enrollees. However, it does include a modest number of employers with more than 100 enrollees, even up to approximately 1000 enrollees. The histograms below, in which the x-axis represents employer size and the y-axis shows the percentage of employers that are that size, demonstrate the distribution of employer sizes. The second plot "magnifies" the y-axis to display the smaller number of large employers.





To demonstrate the robustness of our imputation algorithm, and its predictive value as employer size increases (given that we do not have benefits information on most large employers), we took advantage of the fact that although this file mostly covers employers with 100 enrollees or fewer, there is some overlap with larger employers (i.e., those with ~100 to 1000 enrollees). A random half of our imputation sample had the actual deductible levels of employers of all sizes “hidden” from the imputation. Thus, this random half included a modest number of employers with 75 to 1000 enrollees. We tested the sensitivity and specificity of the imputation in this overlap zone, categorizing employer sizes as 75-100, 101-400, 401-700, and 701-1000 enrollees (Table 1). At employers with 75-100 enrollees, we found sensitivity of 95.4% and specificity of 98.3% (Table 1a). Sensitivity and specificity increased across employer size to 100%, and Table Ss 1b-1d display these for employers of sizes 101-400, 401-700, and 701-1000.

We used an employer ID and an algorithm that determined linked employer subaccounts to identify an employer’s subaccounts per benefit year, and removed benefit years when employers offered both low and high deductible levels.

Rationale for High-Deductible Cutoffs: When Health Savings Account-eligible high-deductible health plans came to market in 2005-2006, the Internal Revenue Service set the minimum deductible level for qualifying high-deductible health plans at $1050 (which could be adjusted upward for inflation annually). The range of this minimum deductible during our study period was $1050-$1250. For these reasons, we defined high-deductible health plans as annual individual deductibles of at least $1000 (otherwise some health savings account plans would be excluded). In addition, choosing this cutoff (as opposed to, e.g., $2000) improves the sensitivity and specificity of the imputation because this is common deductible level and more enrollees per employer meet this threshold. This cutoff is also a “real-world” deductible minimum that allows the most generalizable results. It should also be noted that $1000 was the *minimum* annual deductible level we included and not the mean deductible level. We cannot precisely calculate the mean deductible level of the high-deductible health plan group, but we estimate, using the most common non-zero deductible levels per employer account, an approximate mean deductible of $1900. We defined traditional plans as having deductible levels of ≤$500 after determining that a threshold of ≤$250 would lead to an inadequate sample size for the control group. Again, the mean deductible level of the control group members would be lower than $500.

Our high-deductible health plan group comprised the enrollment years of employers that had a year-on-year transition from low- to high-deductible coverage (i.e., from $500 or less to $1000 or more). Some members had multiple eligible index dates (e.g., multiple low-to-low deductible years or both low-to-low and low-to-high deductible years). In the cases of members with both low-to-low and low-to-high deductible years, we randomly assigned enrollees to the high-deductible health plan or control pool. Then, for members assigned to the control pool that had multiple low-to-low deductible spans, we randomly selected one of their potential index dates (and their corresponding before-after enrollment years).

There were 1,799,404 members with 3,908,743 unique potential patient index date combinations within the overall cohort. Among them, 1,600,531 members had only low-to-low deductible years, 117,217 members had only low-to-high deductible enrollment, and 81,656 members had both low-to-low and low-to-high deductible enrollment. Among those 81,656 members, we first randomly assigned members with statistical “coin flip” to a study group, resulting in 40,962 members for whom we randomly assigned their low-to-high deductible plan switch date as their index date. A very small number of these members had more than one low-to-high deductible plan switch date, so we also randomly chose one of those. For the members remaining from the initial 81,656 sample, we randomly assigned one of their low-to-low deductible plan switches as their index date. Next, among the 1,600,531 members with multiple potential low-to-low deductible index dates, we randomly selected one of the index dates. The final cohort included 158,179 members who underwent a switch from a low-deductible health plan to a high-deductible health plan and 1,641,225 members who remained in a low-deductible health plan throughout the baseline and follow-up periods.

B. Study Group Construction

Individuals in the study were commercially insured members in a large national health insurance database enrolled between 1/1/2003-12/31/2014. This database includes approximately 48 million members with commercial insurance along with their enrollment information and all medical, pharmacy, and hospitalization claims. We included only members with employer-sponsored insurance.

We defined the index date for employers that switched to HDHPs as the beginning of the month when the switch occurred. For employers that did not switch plans, the index date was the beginning of the month when their yearly account renewed. Some members had multiple eligible index dates (eg, multiple low-to-low deductible years or both low-to-low and low-to-HDHP years). In the cases of members with both low-to-low and low-to-HDHP years, we randomly assigned enrollees to the HDHP pool or the control pool. For members assigned to the control pool that had multiple low-to-low deductible spans, we randomly selected one of their potential index dates (and their corresponding before-after enrollment years). Employers had index dates between January 1, 2004, and December 1, 2014.

Individuals were eligible for the study if their employers were present in the database for at least 1 year before after the index date. In addition, we included only employers that provided a single deductible level each year so that subjects were unable to choose a low-deductible versus a high-deductible level. That is, some employers offered a low-deductible plan throughout the study, and others that had offered a low-deductible plan during an earlier time then switched all their enrollees to an HDHP for the remainder of the study. We flagged members who met criteria for diabetes before the index date based on version 11.1 of the Johns Hopkins ACG System.[43,44]

In the above-defined employers, we limited the qualifying population to a pre-match sample of 53,247 diabetes patients whose employers switched to high-deductible plans and 440,359 whose employers kept low deductible plans, were continuously enrolled for at least 1 year before and at least 1 month after the index date, and were aged 40-64 years at the time of the index date. People entered the study at different times because their employers had different index dates. For all individuals in the study, the beginning of study time (time zero) was 12 months before their employers' index date and we defined this 12-month period as the baseline year (Figure 1). The employer's index date was the beginning of the follow-up period.

C. Variables Used in Matching, Adjusting, and Creating Subgroups

Person-level variables: We used version 11.1 of the Johns Hopkins ACG® System[2,3] to calculate members' baseline period morbidity score. The algorithm uses age, gender, and ICD-9-CM codes to calculate a morbidity score and the average of the reference population is 1.0.[3] Researchers have validated the index against premature mortality.[2] To derive proxy demographic measures, the data vendor linked members' most recent residential street addresses to their 2010 US Census tract.[4] Census-based measures of socioeconomic status have been validated [5,6] and used in multiple studies to examine the impact of policy changes on disadvantaged populations.[7-9] Using 2008-2012 American Community Survey[10] census tract-level data and validated cut-points,[5,6] we created categories that defined residence in neighborhoods with below-poverty levels of <5%, 5%-9.9%, 10%-19.9%, and ≥20%. Similarly, we defined categories of residence in neighborhoods with below-high-school education levels of <15%, 15%-24.9%, 25%-39.9%, ≥40%.[5,6] We classified members as from predominantly white, black, or Hispanic neighborhoods if they lived in a census tract with at least 75% of members of the respective race/ethnicity. We then applied a superseding ethnicity assignment using flags created by the E-Tech system (Ethnic Technologies), which analyzes full names and geographic locations of individuals.[11] We classified remaining members as from mixed race/ethnicity neighborhoods. This validated approach of combining surname analysis and census data has positive and negative predictive values of approximately 80 and 90 percent, respectively.[12] Other variables included age as a category (40-49 and 50-64 years) or continuous value; sex; US region (West, Midwest, South, Northeast); employer size used as either categories of 1-99, 100-999, or 1000+ individuals or a continuous value; calendar year of first identified diagnosis; calendar month of the index date; calendar year of the index date; baseline out-of-pocket spending category ($0-<$500), ($500-<$1000), ($1000-<$2500), $2500+); count of continuous enrollment during follow up (capped at 48); tertile of total baseline total standard costs; and, using codes from Table 2, baseline diagnosis of mild-to-moderate retinopathy, mild-to-moderate chronic kidney disease, severe retinopathy, and severe chronic kidney disease, baseline encounter for vision loss; and definitive procedures to treat end-stage renal disease at baseline.

Employer-level variables: These included employer size (number of enrollees insured through a given employer) used as either a continuous variable or with categories of 1-99, 100-999, 1000+ individuals; baseline proportion of members in age category (0-19, 20-29, 30-39, 40-49, 50-64, and 65+ years); proportion of female members; proportion of members in race category (white, black, mixed, Asian, Hispanic); proportion of members in census tract education category (below-high-

school education levels of <15%, 15%-24.9%, 25%-39.9%, ≥40%); proportion of members in census tract poverty category (below-poverty levels of <5%, 5%-9.9%, 10%-19.9%, and ≥20%); baseline intercept and monthly trend of employer total standardized costs; baseline mean ACG morbidity score among members with 12 months of baseline enrollment; baseline median outpatient primary care visit copay; proportion of members in US region (West, Midwest, South, Northeast); calendar month of index date; baseline-to-follow-up network change type (e.g., HMO to PPO, HMO to HMO, HMO to POS, POS to PPO, etc.); and quartile of employer baseline total out-of-pocket expenditure:total spending ratio.

D. Coarsened Exact Matching Description

We conducted pre-post with comparison group study. The intervention group consisted of individuals who were in low-deductible insurance plans for 1 year and then were switched to high-deductible plans for at least an additional month. The control group consisted of matched individuals who remained in low-deductible plans throughout the study. We matched based on calendar year of the index date; the propensity of the employer to mandate high-deductible insurance and the propensity of individuals to work for such employers (each divided into tertiles; component variables described below);[45,46] follow-up duration; and the following member-level baseline variables: out-of-pocket spending category, total health care cost tertile, mild-to-moderate retinopathy diagnosis, mild-to-moderate chronic kidney disease diagnosis, severe retinopathy diagnosis, severe chronic kidney disease diagnosis, encounter for loss of vision, and definitive procedure to treat end-stage renal disease.[13,14]

Components of our member-level propensity score (again, used as a coarsened exact match variable) were age category, employer size category, US region, poverty category, and ACG score. The employer-level propensity score included baseline employer size category, proportion of members in age category; proportion of female members; proportion of members in race category; proportion of members in census tract education category; proportion of members in census tract poverty category; intercept and monthly trend of employer total standardized costs; mean ACG morbidity score among members with 12 months of baseline enrollment; baseline median outpatient primary care visit copay; proportion of members in US region; calendar month of index date; baseline-to-follow-up network change type; and baseline total out-of-pocket expenditure:total spending ratio.

We used coarsened exact matching[15-17] to match the study groups. Coarsened exact matching is a straightforward approach to balancing certain characteristics of study groups and is only a slight modification of exact matching. In exact matching, investigators determine population characteristics that they believe should be balanced between study groups. The exact matching process chooses control group members who have the exact same characteristics as intervention group members (for example, the same gender, age, and income). “Coarsening” denotes that investigators have leeway, based on their understanding of the clinical or research situation, to categorize a given matching variable into the number or type of categories that are most clinically meaningful or relevant to the research question (e.g., categorize age by 5 year versus 10 year intervals or by other meaningful strata such as 20-39, 40-49, 50-64, etc). Coarsened exact matching is then simply an exact match on the selected variables including by their investigator-defined categories.

Because the percentages of subjects with the exact same characteristics will differ between the intervention and control groups, coarsened exact matching software packages generate simple weights so that control members in the stratum represent the same percentage of the control group as intervention members in the stratum represent of the intervention group.[18] For example, if there are 10 intervention group members who are female, high income, and residing in predominantly black neighborhoods, and they represent 2% of the intervention group, and there are 50 control group members with these same characteristics yet they represent 4% of the intervention group population, the exact matching algorithm would assign the control group members in this bin weights of 0.5 (2%/4%). This effectively “shrinks” this overrepresented control group bin from 50 to 25 members. The following is an example from our data showing a single very small stratum where the Stata coarsened exact matching algorithm upweighted the 3 controls relative to the 2 high-deductible members because the particular strata of characteristics is underrepresented in the control pool:

Observation #	Coarsened Exact Match Strata Number	De-identified Member ID	HDHP Group	Coarsened Exact Match Weights
71219	14	2143080659	0	5.50593
119039	14	3262715568	1	1.00000
182213	14	5325447707	0	5.50593
198330	14	2344191836	0	5.50593
231020	14	6484083692	1	1.00000

Coarsened exact matching tries to mimic stratification by population characteristics and then randomization within the defined strata (i.e., fully blocked randomization). Proponents cite numerous advantages of coarsened exact matching over other approaches such as propensity score matching, matching based on Mahalanobis distance, nearest neighbor matching, and optimal matching. These advantages include better balance, lower root mean squared error, less model dependence, being robust to measurement error, ease of implementation, and compatibility with any subsequent statistical estimation model.[16,17,19]

Results.

Table 1. Validation of Deductible Imputation Algorithm, Stratified by Employer Size

Table 1a. Validation of deductible imputation algorithm, using employer accounts of size 75-100 enrollees.

	Gold Standard[a]=high-deductible (n)	**Gold Standard=low-deductible (n)**
We imputed high-deductible	882,588	24,786
We imputed low-deductible	15,612	511,770

	High-deductible	**Low-deductible**
Sensitivity	98.3%	95.4%
Specificity	95.4%	98.3%

[a]Gold standard was a benefits variable specific to each employer derived from a benefits table and obtained from the health insurer via the data vendor.

Table 1b. Validation of deductible imputation algorithm, using employer accounts of size 101-400 enrollees.

	Gold Standard[a]=high-deductible (n)	**Gold Standard=low-deductible (n)**
We imputed high-deductible	1,998,885	42,655
We imputed low-deductible	20,302	1,748,826

	High-deductible	**Low-deductible**
Sensitivity	99.0%	97.6%
Specificity	97.6%	99.0%

[a]Gold standard was a benefits variable specific to each employer derived from a benefits table and obtained from the health insurer via the data vendor.

Table 1c. Validation of deductible imputation algorithm, using employer accounts of size 401-700 enrollees.

	Gold Standard[a]=high-deductible (n)	Gold Standard=low-deductible (n)
We imputed high-deductible	83,393	485
We imputed low-deductible	2,017	122,983

	High-deductible	Low-deductible
Sensitivity	97.6%	99.6%
Specificity	99.6%	97.6%

[a]Gold standard was a benefits variable specific to each employer derived from a benefits table and obtained from the health insurer via the data vendor.

Table 1d. Validation of deductible imputation algorithm, using employer accounts of size 701-1000 enrollees.

	Gold Standard[a]=high-deductible (n)	Gold Standard=low-deductible (n)
We imputed high-deductible	9950	0
We imputed low-deductible	0	19,664

	High-deductible	Low-deductible
Sensitivity	100.0%	100.0%
Specificity	100.0%	100.0%

[a]Gold standard was a benefits variable specific to each employer derived from a benefits table and obtained from the health insurer via the data vendor.

Figure 1. Study Design Showing Example Members of the High-Deductible Health Plan Group and Matched Control Group

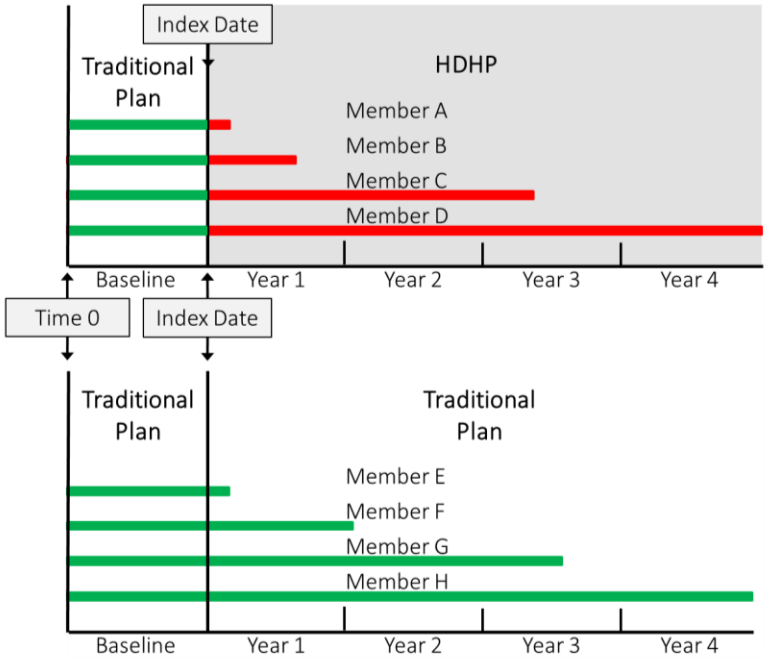
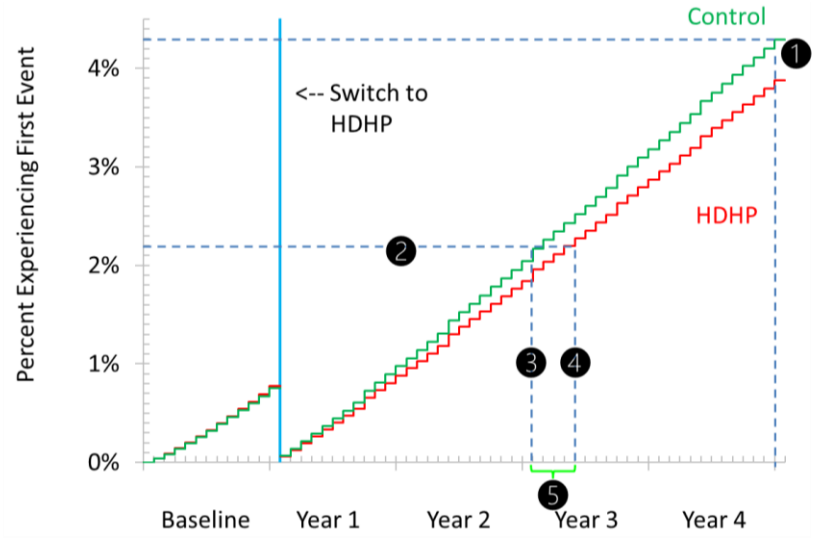


Figure 2. Steps in Estimating the Delay for the High-deductible Group to Reach Half the Final Event Rate of Controls at Follow Up.



1. Determine control group's final event rate using a parametric regression survival-time model with a Weibull distribution and marginal effects methods (e.g., 4.3%)
2. Calculate half of this final event rate (e.g., 2.2%)
3. Estimate time for the control group to reach this 2.2% event rate using same methods as in step 1
4. Use the same approach to estimate time for the high-deductible group to reach this 2.2% event rate
5. Use nonlinear combinations of estimators to determine the difference in months between the high-deductible and control groups reaching the 2.2% event rate (e.g., approximately 4 months)

Table 2. Billing Codes Used to Create Measures of Microvascular Complications

Category	Diagnosis and Procedure Codes
Retinopathy	
Screening	CPT-4: 92227, 92228, 2021F, 2019F, 2022F, 2026F, 92250
Mild-to-moderate disease	ICD-9: 362.01, 362.03, 362.04, 362.05, 362.06
Severe disease	ICD-9: 361.81, 362.02, 362.07, 362.16, 362.35, 362.53, 364.42, 365.63, 379.23
Vision loss diagnosis/treatment	CPT-4: 67025, 67027, 67028, 67040, 67113, 67210, 67220, 67228, 67500, 67515 HCPCS: C9128, C9233, C9291, C9399, J0178, J2503, J2778, J3490, J3590, J9035 ICD-9: 369, 369.0, 369.00, 369.01, 369.02, 369.03, 369.04, 369.05, 369.06, 369.07, 369.08, 369.1, 369.10, 369.11, 369.12, 369.13 369.14, 369.15, 369.16, 369.17, 369.18, 369.2, 369.20, 369.21, 369.22, 369.23, 369.24, 369.25, 369.3, 369.4, 369.6, 369.60, 369.61, 369.62, 369.63, 369.64, 369.65, 369.66, 369.67, 369.68, 369.69, 369.7, 369.71, 369.72, 369.73, 369.74, 369.75, 369.76, 369.8, 369.9, V19.00
Non-specific ophthalmologic disease codes	ICD-9: 249.5, 250.5, 250.50, 250.51, 250.52, 250.53, 361.00, 361.01, 361.02, 361.03, 361.04, 361.05, 361.06, 361.07, 361.2, 361.33, 361.89, 361.9, 362, 362.0, 362.10, 362.13, 362.81, 362.82, 362.83, 362.84
Nephropathy	
Screening	CPT-4: 81000, 81002, 82042, 82043, 82044, 82045, 82565, 82570, 82575
Mild-to-moderate disease	ICD-9: 585.1, 585.2, 585.3
Severe disease	ICD-9: 404.02, 404.03, 404.12, 404.13, 404.92, 404.93, 585.4, 585.5, 585.6 CPT-4: 90970
Loss of renal function diagnosis/treatment	CPT-4: 36556, 36558, 36575, 36580, 36581, 36589, 36832, 36833, 50323, 50325, 50327, 50328, 50329, 50360, 50365, 50370, 50380, 76776, 76778, 90935, 90937, 90939, 90940, 90945, 90947, 90960, 90961, 90962, 99512, 4052F, 4053F, 4054F, 4055F HCPCS: C1750, C1752, S9339, ICD-9: 38.95, 39.27, 39.42, 39.43, 39.95, 54.98, 55.53, 55.6, 55.61, 55.69, 996.68, 996.73, 996.81, V42.0, V56, V56.0, V56.2, V56.8
Non-specific renal disease codes	ICD-9: 250.4, 250.40, 250.41, 250.42, 250.43, 404.00, 404.01, 404.10, 404.11, 404.90, 404.91, 583, 583.8, 583.81, 583.89, 583.9, 585, 585.9, 586

Table 3. Out-of-Pocket Total (Medical and Pharmacy), Medical, and Pharmacy Expenditure Changes in the HDHP and Control Groups in Follow-up Years 1-4 Compared with the Baseline Year

	Mean Out-of-pocket Expenditure, $[1]										Relative Change vs Baseline, HDHP versus Control Group, % (95% Confidence Interval)[1]													
	HDHP Group					Control Group					Year 1			Year 2			Year 3			Year 4				
	Baseline	Year 1	Year 2	Year 3	Year 4	Baseline	Year 1	Year 2	Year 3	Year 4														
Total	1470.0	1963.9	1977.5	2088.9	2248.5	1454.5	1546.3	1587.6	1643.0	1702.8	**25.7**	**(22.0,**	**29.3)**	**23.3**	**(20.6,**	**25.9)**	**25.8**	**(22.2,**	**29.4)**	**30.7**	**(25.7,**	**35.6)**
Medical	755.2	1152.0	1124.7	1202.1	1317.7	741.9	780.8	780.1	795.2	827.3	**44.9**	**(37.8,**	**52.0)**	**41.6**	**(36.5,**	**46.8)**	**48.5**	**(41.4,**	**55.6)**	**56.5**	**(46.7,**	**66.2)**
Pharmacy	715.2	809.6	851.7	886.2	930.4	712.9	765.6	807.6	847.3	875.8	**5.4**	**(4.6,**	**6.2)**	**5.1**	**(3.6,**	**6.6)**	**4.3**	**(1.9,**	**6.6)**	**5.9**	**(2.6,**	**9.2)**

Abbreviations: HDHP, high-deductible health plan. [1]Values derived using generalized estimating equations with a zero-inflated negative binomial distribution and adjusted for employer size.

Table 4. Pre-to-Post Out-of-Pocket Expenditures and Changes among the HDHP and Control Groups for Microvascular Screening Services

	Out-of-Pocket Spending per Service, $[1]				Absolute Change, HDHP vs. Control (95% 95% Confidence Interval), $[1]
	HDHP		Control		
	Baseline	Follow Up	Baseline	Follow Up	
Retinopathy Screening	29.4	43.3	28.3	29.9	**12.2 (11.1, 13.3)**
Nephropathy Screening	1.0	2.2	1.3	1.3	**1.2 (1.0, 1.3)**

Abbreviations: HDHP, high-deductible health plan. [1]Values derived using generalized estimating equations with a zero-inflated negative binomial distribution and adjusted for employer size.

Table 5. Adjusted hazard ratios at baseline and follow up in the high-deductible health plan group relative to controls

	Baseline Year[1]		Follow-up Year[1]	
	Hazard Ratio	(95% Confidence Interval)	Hazard Ratio	(95% Confidence Interval)
Retinopathy-related measures				
Screening	**1.03**	**(1.01, 1.05)**	**0.96**	**(0.95, 0.98)**
Diagnosis and treatment				
Diagnosis of mild-to-moderate disease	1.01	(0.91, 1.12)	0.94	(0.87, 1.01)
Diagnosis of severe disease	1.02	(0.84, 1.25)	**0.89**	**(0.81, 0.98)**
Vision loss diagnosis/treatment	1.03	(0.86, 1.24)	**0.86**	**(0.78, 0.95)**
Diagnosis of any ophthalmologic disease[2]	1.04	(0.999, 1.08)	**0.91**	**(0.88, 0.95)**
Chronic kidney disease-related measures				
Screening	**1.02**	**(1.01, 1.04)**	0.99	(0.98, 1.01)
Diagnosis and treatment				
Diagnosis of mild-to-moderate disease	1.02	(0.88, 1.19)	0.93	(0.86, 1.01)
Diagnosis of severe disease	1.00	(0.68, 1.45)	0.93	(0.81, 1.07)
Loss of renal function diagnosis/treatment	0.99	(0.89, 1.10)	0.94	(0.86, 1.02)
Diagnosis of any renal disease[3]	1.02	(0.98, 1.07)	0.98	(0.93, 1.02)

[1]Baseline and follow-up values derived from separate models using parametric survival time regression with a Weibull distribution and adjusting for employer size category. [2]Includes all specific retinopathy diagnosis/treatment codes as well as non-specific ophthalmologic codes from Supplement Table 2. [3]Includes all specific chronic kidney disease diagnosis/treatment codes as well as non-specific renal disease codes from Supplement Table 2.