

Methods for the Analysis of HEDIS Measures for CAD and Diabetes

Contact: David Eddy

August 31, 2006

Introduction

The value of improving performance on interventions for diabetes and coronary artery disease has been recognized for several decades. Beginning in the early 1990's NCQA developed a set of performance measures called HEDIS. Since then they have been widely used by health plans and others to measure and compare performance. The process for developing performance measures includes a search for evidence to insure that each activity that is promoted by a performance measure improves health outcomes. However it has been very difficult to determine the magnitude of the benefit caused by improved performance, either to a person who is a candidate for a performance measure, or to the population as a whole. Without this information it is not possible to determine the relative importance of improving performance on the measures, or the priority each measure should receive. This paper outlines the methods used to analyze the benefits of improving performance on the HEDIS measures relating to coronary artery disease and diabetes.

Methods – general

The analysis was conducted using data from the Third National Health and Nutrition Evaluation Survey (NHANES III) and the Archimedes model.

NHANES III. NHANES III is a detailed survey and examination of a representative sample of the U.S. population conducted from 1988 to 1994 (1) It provides person-specific data on approximately 40,000 persons. It over sampled certain populations, but includes weights to enable reconstruction of a true random sample of the U.S. population.

Archimedes. The Archimedes model is a person-by-person, object-by-object simulation written at a relatively high level of biological, clinical and administrative detail using object-oriented programming and run on a distributed computing network (2, 3, 4). The core of the model is a set of differential equations that represent the physiological pathways pertinent to diseases and their complications. Currently the model includes diabetes and its complications, congestive heart failure, coronary artery disease, stroke, hypertension, obesity, and the metabolic syndrome in a single integrated model. Use of a single model enables Archimedes to address co-morbidities, syndromes that span multiple organ systems, drugs that have multiple effects, and combinations of drugs. Examples of biological variables and outcomes in the model relating to cardiovascular disease (CVD) and diabetes include cholesterol (LDL, HDL, total), triglycerides, blood pressure (systolic and diastolic), gradual stenosis of coronary and cerebral arteries,

sudden occlusion (e.g. plaque rupture and thrombosis) of coronary and cerebral arteries, myocardial viability, arrhythmias, stroke volume, pulse pressure, arterial compliance, peripheral resistance, basal hepatic glucose production, insulin amount, insulin resistance (efficiency of insulin use), fasting plasma glucose, hemoglobin A1c, glucose tolerance, retinopathy (ETDRS levels), urine protein, creatinine, peripheral neuropathy, foot ulcers, amputations, and many more. The use of differential equations preserves the continuous nature of biological variables as well as the interactions between them. The model is continuous in time with the rate of change of each biological variable being continuously dependent on the values of other biological variables, and any event occurring at any time. Diseases and clinical outcomes are defined in terms of the underlying biological variables, which enables the model to incorporate different definitions and changes in definitions. Interventions, both to prevent CVD and diabetes and to manage it when it occurs, are modeled at the level of the underlying biology. Further details about the model and equations are available upon request.

To create simulated populations that meet specific inclusion, exclusion and other criteria, we used person-specific data from surveys to identify individuals who meet the criteria and then create copies or "clones" of them at the highest level of detail permitted by the data set. For the reference population we use the U.S. population as represented by the National Health and Nutrition Evaluation Survey (NHANES) datasets. These are detailed cross-sectional surveys and examinations of representative samples of the U.S. population at various times. The methods for creating clones are described in a technical report available upon request. Other populations could be created based on other datasets.

People develop symptoms as functions of the underlying biological variables and progression of diseases. For example, chest pain and S-T depressions on EKG are functions of myocardial ischemia. Timing of clinical events is continuous; any event can occur at any time, and the timing can be as condensed or drawn out as occurs in reality. Patients' responses to symptoms and behaviors in seeking care vary depending on individual characteristics and random factors. Test results are functions of the underlying variables being measured. For example, CK-MB test results are functions of the enzyme levels, which are functions of myocardial ischemia. Tests can have systematic and random errors. Interventions are modeled through their effects on the underlying biological variables. For example one of the effects of metformin in the model is to reduce basal hepatic glucose production and thereby lower FPG. Other effects include impact on lipid levels (5) and weight.

The model includes care delivery events such as telephone calls, office and ED visits, admissions, tests and treatments, clinical protocols, practitioner performance, logistics and utilization, personnel and facilities. Care processes and protocols are based on guidelines of national organizations such as the American Diabetes Association and

American Heart Association. Where there are no national guidelines, practices followed at Kaiser Permanente are used for the reference case, but can be tailored to other settings.

To calculate the effects of interventions on quality of life the model continuously tracks the time each person spends with various symptoms or other conditions that affect quality of life. We then apply weights to calculate the total amount of “quality adjusted” time a person spends with each condition. Any set of quality weights can be assigned. For the reference case we used the results of a survey of about 40,000 people reported by Sullivan et al. (6). Weights for people with co-morbidities were obtained from the calculator provided by the same authors. (7)

We test the model for internal consistency and bugs by a variety of methods that include face validity, use of inputs with known outputs, independent duplicate programming of parts, and simulation of studies and trials that have empirically known results. Each of the equations in the model was estimated by fitting functions to data. We confirm the fits by comparing the resulting functions to the data from which they were fitted. To prevent over fitting we chose functional forms with the smallest higher order derivatives (the smoothest curves), and we confirmed each fit visually. Over fitting of curves is also prevented by conducting simulations that involved dozens of equations or spanned long durations of time, and by conducting independent validations that involved data points never used to fit any of the equations. We then validate the overall model by simulating epidemiological studies and clinical trials and comparing the model’s results with the results actually observed in the trials (4). Thus far the model has been validated against 19 clinical trials that are particularly pertinent to this application (4, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25). Eighteen of the trials were chosen by an independent committee appointed by the American Diabetes Association, based on their quality and ability to collectively span the full spectrum of the natural history of the disease, its complications, and its treatments. Counting the individual arms and outcomes of the different trials, a total of 74 validation exercises have been published (4). Overall, the correlation between the model's results and the trials’ results was 0.99. Ten of the 18 trials were not used to build the model, and provide independent validations. The correlation of the model’s results and the trials’ results for these independent validations was 0.98.

The validations cover whatever outcomes were observed in the trials, including the incidence of diabetes, myocardial infarctions, strokes, retinopathy and end stage renal disease. The validations also cover whatever follow-up times were observed in the trials, ranging from about 3 years to 15 years. Finally, the validations span several decades in the progression of the disease, from normal, to high risk, to newly diagnosed diabetes (which itself spans more than a decade, from the time a person first meets the biological definition at FPG > 7 mmol/L [126 mg/dl] to the occurrence of symptoms at FPGs of

about 10 mmol/L [180 mg/dl]), to late complications such as myocardial infarctions and end stage renal disease.

Methods -- Specific

We analyzed 13 different HEDIS measures or variations of HEDIS measures for CAD and Diabetes. Specifically, we analyzed

- The proportion of adults age 18 to 85 who were candidates for at least one HEDIS measure
- The proportion of people who were candidates for each particular measure
- The performance levels in 1988-1992, before HEDIS was introduced
- The number of people who are candidates for each measure, scaled to the approximate current size of the U.S. population
- The effect of achieving 100% performance on each measure, taken together, and considered individually. This provides information about the maximum possible benefit achievable by improving performance on any particular measure or on all of the measures together
- The effect on the U.S. population of achieving an increase in performance of 10 percentage points, for each particular measure
- The effect on the U.S. population of achieving the median and 90th percentile levels of performance of the health plans that reported to HEDIS in 2005 through NCQA (26).
- The effect on individuals of improving performance on each measure

The details are listed in Table 1.

For each measure we calculated the 30 year rates of fatal and non fatal myocardial infarctions (MI), strokes, end stage renal disease (ESRD), and eye surgery and blindness.

Table 1 - Description of HEDIS Measures				
HEDIS Measure	HEDIS Population	Treatment Target	Treatment effects assumed for analyses	Method used to identify people who are HEDIS-improvable in NHANES
Controlling Blood Pressure (CBP)	Adults age 46 – 85 with hypertension	SBP<140 and DBP<90	If SBP>140, then treat to SBP=140. If DBP > 90, then treat to DBP=90	People who had SBP>140 or DBP>90 in NNAHES
Beta Blocker Treatment after a Heart Attack (BBH)	Over age 35, discharged following MI	Prescription for beta blockers at discharge	Treated for 6 months	There are no fields in NHANES that identify people who had just been discharged following an MI. Candidates for this measure (i.e. were discharged following an MI) were specified by running the Archimedes simulation and identified people who in the simulation who just had an MI.
Persistence of Beta Blocker Treatment after a Heart Attack (PBH)	Over age 35, discharged following MI	Treated with beta blockers for at least 6 months after discharge	Treated forever	
Cholesterol Management after Acute Cardiovascular Event (CHM)	Age 18 – 75, discharged in previous year for MI	LDL<100 mg/dl	If LDL>100 mg/dl then treat to LDL=100 mg/dl	
Cholesterol Management after Acute Cardiovascular Event (CHM)	Age 18 – 75, discharged in previous year for MI	LDL<130 mg/dl	If LDL>130 mg/dl then treat to LDL=130 mg/dl	
Comprehensive Diabetes Care (CDC)-HbA1c Control < 9%	Age 18 – 75, with type 1 or type 2 diabetes	HbA1c<9%	If HbA1c>9% then treat to HbA1c=9%	The fields used from NHANES included people who were either told by a doctor that they had diabetes or their FPG was >126. If the person was age <25, then they were classified as having type 1 diabetes. If the person was >25, then they were classified as having type 2 diabetes.
Comprehensive Diabetes Care (CDC)-HbA1c Control < 7%	Age 18 – 75, with type 1 or type 2 diabetes	HbA1c<7%	If HbA1c>7% then treat to HbA1c=7%	
Comprehensive Diabetes Care (CDC)-Eye Exam	Age 18 – 75, with type 1 or type 2 diabetes	Eye exam	Annual eye exam for people on insulin or if HbA1c=8%. Eye exam every two years for all others.	
Comprehensive Diabetes Care (CDC)-LDL Control 100	Age 18 – 75, with type 1 or type 2 diabetes	LDL<100 mg/dl	If LDL>100 mg/dl then treat to LDL=100 mg/dl	
Comprehensive Diabetes Care (CDC)-LDL Control 130	Age 18 – 75, with type 1 or type 2 diabetes	LDL<130 mg/dl	If LDL>130 mg/dl then treat to LDL=130 mg/dl	
Comprehensive Diabetes Care (CDC)-Kidney Screening	Age 18 – 75, with type 1 or type 2 diabetes	Tested for micro-albuminuria in last year	Give annual exams	
Medical Assistance with Smoking (MSC)-Advice to Quit	Age 18 or over, smoker	Seen by MCO practitioner and given advice to quit	Give advice to quit annually	
Medical Assistance with Smoking (MSC)-Medications	Age 18 or over, smoker	Seen by MCO practitioner and recommended medications	Give advice to quit and recommend medications annually	The fields used from NHANES included people who smoked at least 100 cigarettes in life and had smoked in the last 30 days, OR smoked cigarettes more than 15 days during the last 30 days

Design. Our objective was to estimate the benefits that would occur if performance were improved on each of the measures. We did this by using the Archimedes model to conduct a series of simulated clinical trials. In all, 17 simulated trials were conducted. Thirteen addressed the HEDIS measures one-by-one. Four addressed the effects of improving performance on all of the measures at the same time under different assumptions about performance, to be described below.

In each simulated trial a group of people (called a “HEDIS-improvable population” for reasons that will be described below) was either given the intervention applicable to the measure, or was not given any intervention applicable to HEDIS measures, followed for 30 years, and observed for the outcomes of interest.

With one exception, the populations for each scenario were created by randomly selecting from the NHANES III database individuals who met the criteria for improvement in performance of the interventions specified by the HEDIS performance measures. For example, to estimate the effects of improving performance on the HEDIS blood pressure measure, which calls for people age 46 – 85 who have a diagnosis of hypertension to have their BPs controlled to <140/90, we created a simulated population of people with hypertension age 46 – 85 whose SBP>140 or DBP>90. Because these people would benefit from better control of their BPs, they are called “HEDIS-improvable”. It is important to understand that the HEDIS-improvable populations as just defined are different from the people who form the denominators of the HEDIS measures. Using the blood pressure measure as an example again, the denominator for that measure is people ages 46 – 85 who have hypertension. We will call these the “HEDIS-denominator” group. Some of these people are already being treated to BP < 140/90, and their care would not be affected by any further improvements in performance on the measure. On the other hand, some of the people with hypertension still have SBP > 140 or DBP > 90. These people would benefit from further improvements in performance. Because we are interested in the benefits that would derive from improving performance, it is the latter group in which we are most interested for this analysis, and it is the latter group – which we are calling the HEDIS-improvable group – that formed the populations for the simulated trials we calculated. Another way to conceptualize the HEDIS-improvable groups is that they are the complements of the numerators for the corresponding HEDIS measures.

With information about the current level of performance on any given measure, we can use the benefit to the HEDIS-improvable group for that measure to estimate the overall benefit to the entire HEDIS-denominator group, or to the adult population as a whole, that would occur if performance on the measure were improved to any designated new level. The equation is

$$\Delta B = ((P_{\text{new}} - P_{\text{old}})/(1 - P_{\text{old}})) * B,$$

where ΔB is the increase in benefit to the HEDIS-denominator group, such as the number of myocardial infarctions (MIs) prevented in a health plan, P_{new} is the new level of performance on the measure, P_{old} is the old or current level of performance, and B is the benefit of treating the HEDIS-improvable people. This approach assumes that

individuals are chosen at random for improved performance from within the HEDIS-improvable populations¹.

Wherever possible, we obtained information on the original performance levels for each measure from the NHANES III database. Where this was not possible, we used the baseline levels of performance provided by HEDIS. The rates from NHANES III and NCQA are shown in Table 2.

We used NHANES III because it was conducted in 1988 – 1994, just before the NCQA HEDIS measures were introduced, and therefore represents the distributions of risk factors, and the proportion of people who might benefit from improved performance, in the “pre-HEDIS” setting. We also used this database to determine the prevalence rates and distributions of characteristics, risk factors, and biological variables in real HEDIS-improvable populations, as will be described below.

Because NHANES III, properly weighted, is a representative sample of the U.S. population this method also provided information about the proportion of the U.S. population that were candidates for improvement in performance of each HEDIS measure, prior to the introduction of the measures by NCQA.

Each HEDIS-improvable population was then subjected to two or more interventions, analogous to the arms of a real controlled trial. One arm was always a "status quo" arm, in which the chosen population was allowed to progress with no changes in any aspects of their care relevant to the HEDIS measures. They would however receive treatment for any symptoms or health outcomes that might develop. For example, if the measure being studied was control of LDL levels to <100 in people with history of myocardial infarctions (MI), then the population for that scenario would be people with a history of MI and LDL > 100, and for the status quo arm the LDL levels would be allowed to progress gradually as occurs with age, without any treatment. If such a person should develop symptoms of coronary artery disease they would be treated appropriately, but they would receive no prevention interventions.

The status quo arm then served as a comparison for the intervention, in which the HEDIS-improvable population would be treated with the intervention called for by the HEDIS measure. In the example just given, for the intervention arm the people would have their LDL levels treated to 100. In this arm, everyone would also be reexamined at annual intervals and retreated as needed to maintain control. For example if their LDL levels rose to > 100, then they would be re-treated to < 100. In this arm we caused everyone in the HEDIS-improvable population to be treated to the goal. This information

¹ If in reality people are more likely to be controlled if they are already close to a HEDIS performance target, then this assumption will cause our estimates of benefit to overestimate the actual degree of benefit. Conversely, if people are more likely to be controlled if they are far above the treatment goal, then this assumption will cause our estimates to underestimate the degree of benefit.

could then be used in the equation given above to calculate the effect of any new level of performance.

Because the purpose of this project was to learn the effects of achieving the goals set for the various HEDIS measures, we constructed artificial treatments that achieved the goals precisely. They can be thought of as magic pills. In reality, patients would be treated according to algorithms often with a progression of drugs and dosages. Because people start at different levels of a risk factor or variable (e.g. different LDL levels) and respond differently to drugs, in order to achieve 100% performance where everyone is driven to the goal, inevitably some people would be “over-treated” to levels beyond the goal. The magic pills created for this project avoided that effect and the biases it would have in overstating the benefits of achieving the stated HEDIS goal.

To create each HEDIS-improvable population we created simulated copies or “clones” of people in the NHANES III population, where each clone matched the corresponding real person with respect to all the important variables. The method for creating clones ensured that the distributions of all the variables, and the correlations between the variables, were the same in the simulated population as the real U.S. population. For example, in the HEDIS-improvable population for the hypertension measure, while everyone has a BP > 140/90, some would be only slightly above that threshold, perhaps SBP = 143, whereas others would be far above that threshold, perhaps SBP = 168. The method for creating the HEDIS-improvable populations ensured that the simulated populations would have the same distributions of BP levels as the actual NHANES III population. Similarly, people with only slight increases in BP might have a different likelihood of being a smoker, or have a different distribution of weight, than people with much higher BPs; the method we used to create the simulated population preserves all the pertinent correlations measured in the real NHANES III population.

The approach just described for creating the populations for the simulated clinical trials had to be modified for the performance measures relating to beta-blockers on discharge after an MI; whether such patients were given a prescription on discharge and whether they were still taking beta blockers at least 6 months after discharge. The number of people in the NHANES III database who were recently discharged following an MI is extremely small. Consequently, to analyze performance on measures and guidelines relating to this measure we identified a group of people who were at an exceedingly high risk (approximately 90%) of developing an MI in the coming year. We then tracked them and noted the occurrence of any MIs. Because of the nature of this measure it is not possible to calculate the number of people affected by it over a long time horizon using NHANES III data. To estimate the number of people alive today who would benefit from improved performance on a beta blocker post-MI measure, we counted the number of people in the NHANES III database who had had an MI in the past.

Perspectives. We calculated the effects of improved performance from two perspectives: the individual and the population. For the first, we looked at each

measure one by one and calculated the effects on individuals who could benefit from better performance on that measure – that is, people who are actually in the HEDIS-improvable populations. An example is HbA1c control in people with diabetes; to calculate the benefit of this measure from the individual’s perspective we identified people who had diabetes and had HbA1c levels > 9%, and then treated them to reduce their HbA1c to 9%. This focuses on people who stand to benefit from improved performance on the measure, because their current level of HbA1c is greater than the target for treatment. People with diabetes whose HbA1c level was already below 9% would not benefit from a measure that lowered HbA1c to 9%. We did this for all 13 measures listed in Table 1.

For the population perspective we performed the same type of calculation for each measure, but kept track of the proportion of the total adult population (18 – 85) that stood to benefit from better performance on each measure. This enabled calculation of the number of people who stand to benefit from a measure from within a general adult population of any size, such as a 50,000 member health plan, the state of California, or the U.S. population as a whole. For example, NHANES III indicates that approximately 1% of adults age 18-85 have diagnosed diabetes and HbA1c levels > 9%. Therefore in a health plan with 500,000 adults 18-85, there will be approximately 5000 such people.

Interventions

The interventions were the treatments called for by the HEDIS measures shown in Table 1. Each of these was analyzed one-by-one in its own scenario. But as already stated, for the population perspective we also looked at several combinations of measures when applied to the entire group of people who stand to benefit from improvement in performance for at least one performance measure or guideline -- the “sum” of all the individual HEDIS-improvable populations. For this population, which we will call the “HEDIS-improvable super-population”, we calculated the effects of the following combinations of measures:

- Achieving 100% performance on every measure, where treatment targets are interpreted aggressively. For example, if smokers can either be given advice to quit, or be given advice to quit and recommendations for medications, the latter option would be considered the more aggressive.
- Achieving 100% performance on every measure, where treatment targets are interpreted conservatively. For example, if smokers can either be given advice to quit alone, or be given advice to quit plus recommendations for medications, the former option would be considered the more conservative.
- Achieving the level of performance of the median health plan.
- Achieving the level of performance of the 90th percentile health plan.

NCQA provided the assumptions to be made about the level of performance for each intervention for the median health plan and the 90th percentile plan. They are in Table 2.

Table 2 - Performance Levels for Each Measure					
Measure	Porportion of Pre-HEDIS population improvable by measure	Performance levels			
		NHANES 3	HEDIS baseline	Median Health Plan in 2005	90 th percentile Health Plan in 2005
Controlling Blood Pressure (CBP)	13.6%	28%		67.51%	75.35%
Beta Blocker Treatment after a Heart Attack (BBH) [1]	0.3%	Not available in NHANES 3	86%	97%	100%
Persistence of Beta Blocker Treatment after a Heart Attack (PBH) [2]	0.3%	Not available in NHANES 3	86%	69.05%	79%
Cholesterol Management after Acute Cardiovascular Event (CHM) (Goal <130)	1.6%	35%		69.38%	78.33%
Cholesterol Management after Acute Cardiovascular Event (CHM) (Goal <100)	1.1%	10%		51.59%	61.25%
Comprehensive Diabetes Care (CDC)-HbA1c Control 9% (% NOT controlled to < 9%)	1.0%	15%		29.93%	20.90%
Comprehensive Diabetes Care (CDC)-HbA1c Control 7% (% NOT controlled to < 7%)	2.4%	38%		not available	not available
Comprehensive Diabetes Care (CDC)-Eye Exam	3.2%	Not available in NHANES 3	50%	50.35%	66.18%
Comprehensive Diabetes Care (CDC)-LDL Control (<130)	3.2%	52%		65.78%	73.24%
Comprehensive Diabetes Care (CDC)-LDL Control (<100)	5.1%	20%		40.40%	47.45%
Comprehensive Diabetes Care (CDC)-Kidney Screening	3.1%	Not available in NHANES 3	50%	51.58%	65.45%
Medical Assistance with Smoking (MSC)-Advice to Quit	30.9%	Not available in NHANES 3	not available	69.50%	77.20%
Medical Assistance with Smoking (MSC)-Medications	30.9%	Not available in NHANES 3	not available	37.80%	46.70%

[1] BBH data is based on 3-month treatment with beta blockers

[2] PBH data is based on 6-month treatment with beta blockers

A general note is that it was not possible to analyze the effects of testing for LDL, HbA1c, eye pathology or kidney pathology without making some assumptions about the care

that would be provided for people who had abnormal test results. Nor was it possible to analyze the effect of controlling a person's LDL or HbA1c without assuming that they had been tested and found to be candidates for control. Thus for the eye and kidney screening measures we assumed that appropriate care would be given to any cases found, and for LDL and HbA1c we combined the "testing" and "control" measures into a single "testing and control" measure.

Outcomes

For each of these interventions we estimated a large number of outcomes, including the following:

- myocardial infarctions, both fatal and non fatal
- survival following a myocardial infarction
- repeat myocardial infarctions
- strokes, both fatal and non fatal
- retinopathy and blindness (diagnosed legal blindness for all calculations except retinopathy exams, for which actual legal blindness was calculated)
- nephropathy and end stage renal disease
- Revascularization (PTCA/CABG) with the understanding that use of these procedures varies widely from setting to setting. We do this for a typical managed care setting, but there is no guarantee that the results will be accurate for all settings. In fact, they won't.
- Hospitalizations from CHF, with the understanding that criteria for deciding who should be hospitalized can vary across settings.
- Incidence of chest pain (if possible) with the understanding that the criteria for reporting chest pain as an outcome in trials varies from setting to setting. We do the best we can with the available data.
- Life years
- Quality adjusted life years

Quality of Life

For the reference case we used quality of life weights shown in Table 3, drawn from Sullivan et al. (6,7).

Table 3 - QALY Weights	
Condition	Weight
Diabetes and	
MI	-0.18
Angina	-0.18
Stroke	-0.167
CHF	-0.2
ESRD	-0.2
Foot amputation	-0.105
Partial foot amputation	-0.105
Foot ulcer	-0.17
Serious eye disease	-0.19
No complications	-0.04
No Diabetes and	
MI	-0.05
Angina	-0.05
Stroke	-0.037
CHF	-0.07

Time horizon

We estimated outcomes over a 30-year period, reporting both clinical and economic effects at annual and five-year intervals.

Statistical methods

The results were calculated by two methods. First, we calculated the cumulative probabilities that a person who is alive at the start of the simulated trial will have various outcomes. We also calculated the sum of the risks of an event during the coming year faced by people who are alive at the beginning of the year (annual hazard rate). This is sometimes called the integrated hazard rate. It is equivalent to a Kaplan Meier calculation where observations are made at annual intervals.

The two numbers are calculated as follows. To calculate the cumulative probability of a health outcome, take the cumulative number of outcomes that have occurred through the years to the population, and divide by the number of people alive *at the beginning of the calculation*. For the integrated hazard rate, first calculate the annual hazard rate for each year; that is the number of outcomes in that year divided by the number of people *alive at the beginning of that year* and at risk of having the event in that year. Then to get the integrated hazard rate, add up all the annual hazard rates. The calculation methods are illustrated in Table 4.

Table 4 - Calculation Methods						
Year	0	1	2	3	4	5
Number of people alive and without the outcome	1000	930	864	786	731	677
Number of people dying of other causes in the year	20	23	24	21	10	23
Number of people getting the outcome in the year	50	43	54	34	44	53
Cumulative number of outcomes	50	93	147	181	225	278
Number of people alive and without the outcome at the end of the year	930	864	786	731	677	601
Cumulative probability of outcome	0.05[1]	0.093[2]	0.147	0.181	0.225	0.278
Annual hazard rate of outcome	0.05	0.0462[3]	0.0625	0.0433	0.0602	0.0783
Integrated hazard rate	0.05	0.0962[4]	0.1587	0.202	0.2622	0.3405

[1] $50/1000 = 0.05$

[2] $43/1000 + 0.05 = 0.093$

[3] $43/930 = 0.0462$

[4] $0.05 + 0.0462 = 0.0962$

Additional notes on methods

For all the calculations we based the starting or pre-HEDIS performance levels on the performance levels seen in the NHANES III population when possible. These rates are shown in Table 2. For smoking we assumed that all smokers were open to responding to advice to quit and information about medications.

There are two beta blocker measures for which we created separate simulations. For beta blocker use after a heart attack, the measure specifies an ambulatory prescription for beta-blockers rendered within seven days after discharge. The measure for persistent beta blocker use specifies a 180-day course of treatment with beta-blockers. Because the beta blocker measures together have the effect of measuring if beta blockers are used for 6 months, we created a simulated trial in which we assumed that the beta blocker would be given for 6 months and then stopped. Because the intention (even if not the actual process being measured) is for beta blockers to be used forever, we also created a simulated trial in which they are given forever.

For the U.S population, we assumed 210,000,000 adults age 18 – 85.

For the HEDIS measures that call for reductions in LDL, we used simulated treatments that had the sole effect of reducing LDL to the desired goal (either 130 mg/dl or 100 mg/dl). We did not assume that statins would be used, because the measure asks only about the LDL level, and not statin use. If statins were used, the effect would be larger than reported here.

For the smoking measure calling for advice to quit and information about medications, we assumed that smokers exposed to this intervention would quit at a rate starting at 10.5% in year 1 and tapering uniformly to 1.05% in year 30. These quit rates were assumed to occur on top of the spontaneous quit rates. For the smoking measure calling for advice to quit, we assumed that smokers exposed to this intervention would quit at a rate that begins at 5.86% in year 1 and tapers uniformly to 0.586% in year 30. These quit rates were assumed to occur on top of the spontaneous quit rates. These quit rates were provided by NCQA.²⁷

Because our methodology was to conduct simulated controlled trials, the results have confidence limits determined by the sample sizes of the simulated trials and the rates of the outcomes, just as occurs in real clinical trials. Because of this, small effects can be due to sampling error. Confidence intervals were calculated.

The results will be shown as cumulative probabilities. Many of the interventions cause people to live longer, extending the time they are at risk of other outcomes. This can cause what appear to be paradoxical results, where an effective treatment actually increases the risk of outcomes. The interpretation of these paradoxical results is further confounded by the fact that the interventions are selective in their effects, causing shifts in the prevalence of various risk factors in the surviving populations. The shifts in risk factors can selectively affect the rates of other outcomes that are not directly affected

by the particular intervention being applied. To some extent these effects can be reduced by looking

Other apparently paradoxical results arise as artifacts of the subject selection process. For example, it appears that more MIs were prevented when testing the complete set of HEDIS guidelines than the sum of the numbers prevented from each of the guidelines alone. But the single-measure trials start with a subject pool entirely composed of people who are already candidates for the given treatment, while the composite intervention includes all these people but additionally will treat anyone else in the "HEDIS super-population" who develops a need for the treatment over the trial's 30-year course.

¹ (http://www.cdc.gov/nchs/products/elec_prods/subject/nhanes3.htm).

²[Schlessinger L, Eddy DM](#). Archimedes: a new model for simulating health care systems--the mathematical formulation. *J Biomed Inform.* 2002 Feb;35(1):37-50.

³ Eddy DM, Schlessinger L. Archimedes: a trial-validated model of diabetes, *Diabetes Care.* 2003;26:3093-3101

⁴ Eddy DM, Schlessinger L, Validation of the Archimedes diabetes model, *Diabetes Care.* 2003;26:3102-3110

⁵ DeFronzo RA, Goodman AM, for the Multicenter Metformin Study Group: Efficacy of Metformin in patients with non-insulin dependent diabetes mellitus. *New Engl J Med.* 1995;333:541-549

⁶ Sullivan PW, Ghushchyan V. Preference-based EQ-5D index scores for chronic conditions in the United States. *Medical Decision Making* 26:410-420, 2006.

⁷ Available online at <http://www.ohsu.edu/epc/mdm/>

⁸ UK Prospective Diabetes Study (UKPDS) Group: Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet* 352:837-852, 1998.

⁹ The Diabetes Control and Complications Trial Research Group: The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *NEngl J Med* 329:977-986, 1993.

¹⁰ Irbesartan in Patients With Type-2 Diabetes and Microalbuminuria Study Group: The effect of irbesartan on the development of diabetic nephropathy in patients with type 2 diabetes. *NEngl J Med* 345:870- 878, 2001.

¹¹ Diabetes Prevention Program Research Group: Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 356:393-402, 2002.

¹² Heart Protection Study Collaborative Group: MRC/BHF Heart Protection Study of antioxidant vitamin supplementation in 20,536 high-risk individuals: a randomized placebo-controlled trial. *Lancet* 360:23-33, 2002.

¹³ The Heart Outcomes Prevention Evaluation Study Investigators: Effects of an angiotension- converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. *N Engl J Med* 342:145- 153, 2000.

¹⁴ The Heart Outcomes Prevention Evaluation Study Investigators: Effects of ramipril on cardiovascular and microvascular outcomes in people with diabetes mellitus: results of the HOPE study and MICROHOPE substudy. *Lancet* 355:253-259, 2000.

¹⁵ Sacks FM, Pfeffer MA, Moyer LA, Rouleau JL, Rutherford JD, Cole TG, Brown L, Warnica JW, Arnold JMO, Wun CC, Davis BR, Braunwald E: The effect of Pravastatin on coronary events after myocardial infarction in patients with average cholesterol levels. *N Engl J Med* 335: 1001-1009, 1996.

¹⁶ Lewis EJ, Hunsicker LG, Clarke WR, Bain, Raymond P, Berl T, Rohde R, Raz I: The effect of angiotensin-converting-enzyme inhibition on diabetic nephropathy. *N Engl J Med* 329:1456-1462, 1993.

¹⁷ Lewis EJ, Hunsicker LG, Clarke WR, Tomas P, Pohl MA, Lewis JB, Ritz E, Atkins RC, Rohde R, Raz I: Renoprotective effect of the angiotensin-receptor antagonist irbesartan in patients with nephropathy due to type 2 diabetes. *N Engl J Med* 345: 851-860, 2001.

¹⁸ Lewis SJ, Moyer LA, Sacks FM, Johnstone DE, Timmis G, Mitchell J, Limacher M, Kell S, Glasser SP, Grant J, Davis Barry R, Pfeffer MA, Braunwald E: Effect of pravastatin on cardiovascular events in older patients with myocardial infarction and cholesterol levels in the average range: results of the cholesterol and recurrent events (CARE) trial. *Ann Intern Med* 129: 681-689, 1998.

-
- ¹⁹ LIPID Study Group: Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. *N Engl J Med* 339: Frick MH, Elo O, Haapa K, Heinonen OP, Heinsalmel P, Helo P, Huttunen JK, Kaitaniemi P, Koskinen P, Manninen X, et al: Helsinki Heart Study: primary-prevention trial with gemfibrozil in middle-aged men with dyslipidemia: safety of treatment, changes in risk factors, and incidence of coronary heart disease. *N Engl J Med* 317:1237–1245, 1987; 1349–1357, 1998.
- ²⁰ SHEP Cooperative Research Group: Prevention of stroke by antihypertensive drug treatment in older persons with isolated systolic hypertension: final results of the Systolic Hypertension in the Elderly Program (SHEP). *JAMA* 265:3255–3264, 1991
- ²¹ The Lipid Research Clinics Coronary Primary Prevention Trial results. I. Reduction in incidence of coronary heart disease. *JAMA* 251:351–364, 1984. Medical Research Council Working Party: MRC trial of treatment of mild hypertension: principal results. *BMJ* 322:97–104, 2001
- ²² Shepherd J, Cobbe SM, Ford I, Isles CG, Lorimer AR, Macfarlane PW, McKillop JH, Packard CJ, the West of Scotland Coronary Prevention Study Group: Prevention of coronary heart disease with Pravastatin in men with hypercholesterolemia. *N Engl J Med* 333:1301–1307, 1995
- ²³ Rubins HB, Robins SJ, Collins D, Fye CL, Anderson JW, Elam MB, Faas FH, Linares E, Schaefer EJ, Schectman G, Wilt TJ, Wittes J: Gemfibrozil for the secondary prevention of coronary heart disease in men with low levels of high-density lipoprotein cholesterol. *N Engl J Med* 341: 410–418, 1999
- ²⁴ Scandinavian Simvastatin Survival Study Group: Randomized trial of cholesterol in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet* 344:1383–1389, 1994
25. Colhoun HM, Betteridge DJ, Durrington PN, Hitman GA, Neil HA, Livingstone SJ, Thomason MJ, Mackness MI, Charlton-Menys V, Fuller JH; CARDS investigators. Primary prevention of cardiovascular disease with atorvastatin in type 2 diabetes in the Collaborative Atorvastatin Diabetes Study (CARDS): multicenter randomized placebo-controlled trial. *Lancet*. 2004;364:685-96.
- ²⁶ Russ Mardon, NCQA, personal communication
- ²⁷ Russ Mardon, NCQA, personal communication