PART 2: ACTUARIAL ISSUES IN CARE MANAGEMENT INTERVENTIONS

Paper 6: An Actuarial Method for Evaluating Disease Management Outcomes

Ian Duncan[1], FSA, FIA, FCIA, MAAA

March 29, 2005

Introduction

The actuary may not always have the chance to design an outcomes measurement study, and will more frequently be called-in to evaluate a vendor's or colleague's results from an existing study. Whether designing a study from scratch or evaluating a published study, the measurement principles discussed in detail in earlier papers apply.

Occasionally, adjustments are made by a researcher to correct for risk factors that differ between reference and intervention populations. Examples of obvious adjustments are adjustment for the effects of trend, catastrophic claims, age and plan design. It is usually possible to identify when adjustments such as these are required and the extent of the adjustment to be made. Where these adjustments are well documented, it should be possible to assess their validity.

Other factors are not as readily identifiable or quantifiable. Examples of the latter include selection bias (the fact that those members who participate in programs are not randomly distributed within a chronic population), regression to the mean or discrete changes in the population being measured (as, for example, when a new sub-population that is not equivalent to the former population is added to or leaves the program).

This paper describes an actuarial methodology for evaluating disease management outcomes. We address several important issues in this paper:

- Control of exposure,
- Identification of measured populations, and
- Ensuring equivalence between baseline and intervention populations.

[1] Lotter Actuarial Partners, Inc. New York, NY

A fourth actuarial issue, the calculation and application of health care cost trend, is addressed in Paper 7.

The principles of an actuarial adjustment methodology may be applied to any study design in which the results of an intervention group are compared with those of a comparison group. Examples of non-random control designs include geographic, temporal, or product-based controls. The actuarial methodology for assessing managed care outcomes discussed in this paper is one frequently used by health plans and disease management (DM) vendors to assess their financial outcomes, and is an example of an adjusted historical control design. An example of the methodology is described in the American Healthways/Johns Hopkins paper.[2] Although the American Healthways paper does not assign a name to the methodology, it is effectively an adjusted historical method. We begin by describing the calculation in more detail.

The Actuarially-Adjusted Historical Control Design

Under this design, objective criteria are used to define members for inclusion in either reference or intervention population. Certain outcome statistics are measured for that population during the historical period (often referred to as a "baseline" period). Examples of statistics measured in the baseline period include admissions per 1000 of the population, per member per month costs, or clinical markers such as the number of patients receiving beta-blockers, etc.

The measurement period may be adjacent to the baseline period, or not; it is one of the strengths of the actuarial-adjustment methodology that the periods need not be continuous. There may be some, but not complete, overlap between the populations (i.e. the same members will be identified in both baseline and measurement periods) identified in the baseline and measurement periods. This methodology is not a cohort study, however, because we are not following a population identified in the baseline period through to the end of the intervention period, but rather two populations in two periods, identified

[2] American Healthways, Inc., and Johns Hopkins University. 2003. Standard Outcomes Metrics and Evaluation Methodology for Disease Management Programs. *Disease Management* 6 (3) 121-138.

according to the same criteria. Equivalence between the baseline and intervention period populations is assumed to result from the symmetrical treatment of members in each period, that is, applying exactly the same rules in each period.

Generally, the intervention program begins before, or simultaneously with, the measurement period. A naïve observer would measure the effect of the intervention program as the difference between the statistic being measured in the baseline and measurement periods. However, it is an empirical fact that most health care utilization statistics change over time, even in stable populations. Health insurance actuaries and underwriters allow for this effect by applying "health care trend" to their projections (after controlling specifically for directly controllable factors). In the historical control design, savings are not directly measurable. Instead, they are derived as the difference between an estimated statistic and the actual statistic as measured in the measurement period. The estimated statistic is the corresponding historical statistic from the baseline period, projected for a period of a few months or years to the intervention period.

Figure 1 shows a simple example of the application of the Historical Control, or Actuarial Methodology, to the estimation of savings in a population. The outcome being measured is the cost of admissions. However, the methodology could be applied to net paid claims, emergency room visits or any relevant measure of utilization. Baseline medical admissions are recorded for a chronic population (numbering 50,000 chronic members, or 600,000 member months of exposure, assuming every member is continuously enrolled for 12 months). The baseline medical admission rate is projected one year to the first measurement period, applying an annual trend of 5.3 percent. In this particular example, the applicable trend is derived from the comparable, non-chronic member experience of the same health plan (externally derived and not shown in the example). We refer to this population below as the "Index" population, because it is used to create an index to be applied to the chronic population utilization. Whether the Index population experience is in fact comparable to that of the chronic population is a matter for study and debate, and we return to this topic in Paper 8. Any external source of trend experience may be used, provided the experience on which it is based is not affected by the intervention that the methodology is attempting to calculate.

Figure 1. Simple example of the Actuarially-Adjusted Historical Control Methodology

Basic data used in the calculation

	Baseline Period	**Measurement Period**
Period	1/1/2001- 12/31/2001	1/1/2002 – 12/31/2002
Average total population[3]	150,000	150,000
Average chronic population	50,000	50,000
Chronic Member months	600,000	600,000
Chronic population Inpatient Admissions	30,000	28,800
Chronic population Inpatient admissions/1000/year	600.0	576.0
Cost/admission	$7,500	$8,000
Utilization (admission) trend (Derived from an external source, e.g. the "Index" population)	-	5.3%

Example of a Savings Calculation:

In the example below, we apply the data assembled in the table above. The avoided admissions (equal to measured period admissions less baseline period admissions) are multiplied by an average cost per admission to generate overall dollar savings. The average cost per admission may be observed directly from the Index population in the measurement period, or may be estimated by trending forward an average cost per admission from the baseline period, using a suitable admission unit cost trend.

[3] This is an example of a Medicare population. The chronic prevalence (33.3 percent) and number of admissions/1000/year (600.0) are typical of chronic Medicare populations. Both of these statistics will be lower in commercial populations, although the same principles illustrated here will apply.

Estimated Savings due to Averted Admissions =		
Baseline Admissions/1000 * Utilization Trend	600.0*1.053 =	631.8
Minus: Actual Admissions/1000/yr		576.0
Equals: Reduced Admissions/1000/yr		55.8
Multiplied by: Actual member years in Measurement Period/1000		50.0
Total reduced admissions		2,790.0
Multiplied by: Trended unit cost/admission		$8,000
Equals: Estimated Savings due to Averted Admissions		$22,320,000

Once the calculation has been completed, we recommend validating and reconciling the savings to the underlying cost. As a test of reasonability of the result, the underlying cost of a Medicare population ranges between $6,000 to $8,000 per member per year, or (for 150,000 members) a total cost of $900,000,000 to $1.2 billion. While estimated savings of $22.3 million from a program in the chronic population may seem high in absolute terms, relative to the total cost of the Medicare population the savings represent 1.9 percent to 2.5 percent, which is consistent with results from other studies of this type.[4]

Practical Application of the "Actuarially-Adjusted Historical Control" Methodology

The key component of the actuarial methodology is the application of the trend factor that adjusts historical experience to an estimate of current period experience, absent intervention. "Health care trend" is the term applied to the empirical observation that most health care measures (utilization, unit cost, per member per month costs, etc.) tend to change over time. Generally, but not always, trend results in increases in health care measures. The choice of an appropriate health care trend assumption to apply to the baseline experience for calculating savings is discussed in Paper 7. Paper 8 explores the

[4] For an example of pmpm savings from a population study, see: M. Cousins and Y. Liu. 2003. Cost Savings for a PPO Population with Multi-Condition Disease Management: evaluating program impact using predictive modeling with a control group. *Disease Management* 6(4) 207-217.

practical issues of the application of this methodology in a population subject to disease management interventions.

The historical control methodology is an "open group" method, in which a comparable population is selected according to the same criteria in each period. This methodology contrasts with a closed group or cohort methodology (such as the Johns Hopkins/ American Healthways methodology) in which, to be included in the measured population, a member must have been continuously eligible for at least 24 months of the baseline and intervention periods. With regular enrollment, termination and identification of new members that are found in most groups, it is reasonable to expect that the open group method will produce a stable population, year-to-year (at least with respect to common risk-factors such as age, gender and disease prevalence). A closed-group methodology, by contrast, will produce a group that is subject to the effects of aging and disease progression. The theory and practice of trend calculation for use in projecting historical costs is covered in more detail in Paper 7. In summary, when projecting experience from a baseline to an intervention period, it is important to separate the trend measure used between factors that may be allowed for directly (such as the effect of aging and disease progression, or the effect of benefit design features and their changes) from other trend factors such as increase in intensity, changes in medical practice or changes in provider contracts. For the purposes of this paper, we assume that it is possible to find an unbiased estimate of chronic population trend, without the effect of the intervention.

Exposure

If outcomes of interventions are to be rigorously measured, it is critical that members and their associated claims be tracked, allocated, associated and summarized appropriately. Actuaries know this issue as the topic of "Exposure to Risk."[5] Actuaries familiar with underwriting and pricing will recognize that establishing appropriate baseline and intervention period membership populations is similar to the problem of identifying

[5] In disease management exposure has two meanings: (1) A patient is "exposed" to an intervention by being a member of a group selected for intervention or a program. (2) For measurement or actuarial calculations, "exposure" has a meaning synonymous with "denominator," and refers to the entire group eligible for an intervention, or included in a study. The risk-unit is often the member month, and the total "exposure to risk" is the total number of member months measured between the start and end-dates of the study.

enrolled populations for underwriting and pricing. In this section we assume that a valid scheme exists for defining who is in which category. We will return later to the definitions of each category.

Managed Versus Measured Populations

The population to be measured need not be the same population being managed. This may not seem obvious, but a few examples will point out the differences. A DM program may be offered to all chronic members of a health plan. Some of these members may not be good candidates for management (for example, if the member is institutionalized, or suffers from a terminal disease). Conversely, the program may be offered to members who self-identify with a chronic disease, even when they do not have a claims history that would objectively identify them as having the disease. All of these members represent a potential for confounding the DM company's results. In the first example there is potential for confounding because the member represents a chronic individual who will not contribute to savings; while in the second example, the self-identified member will have no counterpart in the baseline period (because members are only self-identifying in the intervention period, destroying the necessary objectivity and symmetry of the identification process). Although the DM company will be managing the care of these members, the DM company and health plan may agree to ignore these members in the actual evaluation. The treatment of non-measured members is independent of the particular methodology chosen to measure results, and may apply, for example, to members in a randomized controlled study.

Member Months

The basic unit of measurement for any evaluation is the member month. In any month, a member is uniquely classified into a single category (defined in more detail below). Members can move between categories from one month to the next, although movements between some categories may not be possible. The number and types of categories used depends on the type of evaluation, the level of detail sought in the study and the types of risk that the study is monitoring.

In the examples that follow, we list a number of different categories that we have used to classify members in studies of DM measurement. However, subsets of membership classes may be combined. The application of the classification rules is in most cases hierarchical.

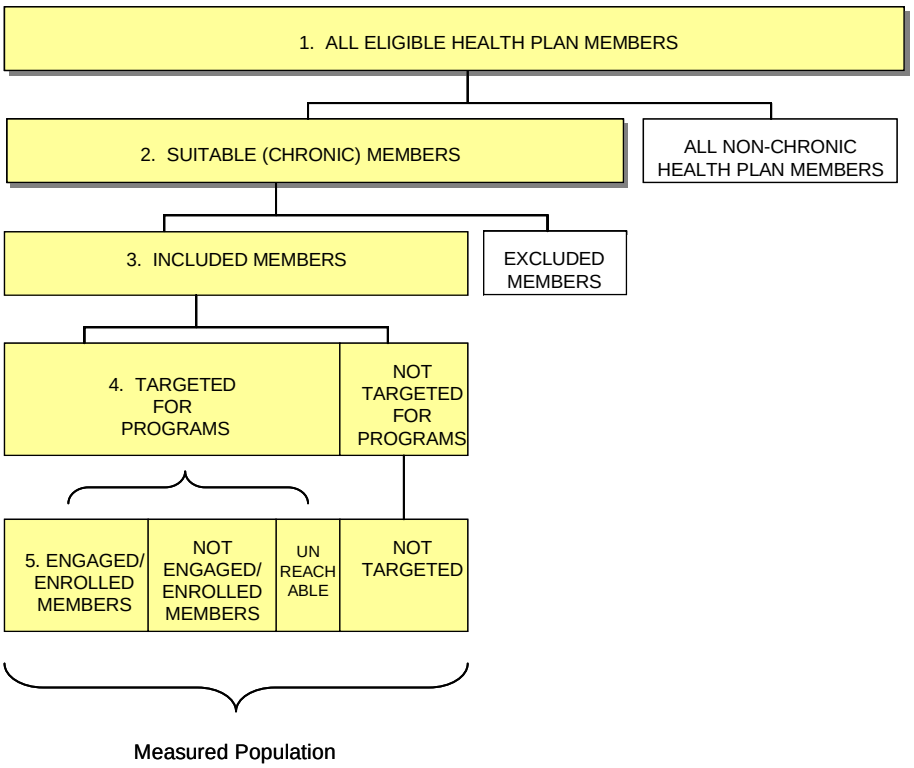
Eligible Members

For measurement purposes, we first determine eligibility for health plan membership, then eligibility for DM services (for example the program may not be available to self-insured groups). While this step may seem simple and obvious, anyone familiar with health plan data will know that determining unambiguous eligibility is not a simple task, and is a task often subject to multiple revisions on a monthly basis. In the figure below, we have assumed that all eligible health plan members are also potentially included in the DM program. Figure 2 illustrates the member classification schematically.

Chronic and Non-Chronic ("Index") Members

Within the eligible membership population, we assign members according to their chronic status. Some companies refer to these members as "suitable," meaning suitable for the intervention program. The assignment of chronic status is determined continuously (i.e. monthly). Any set of definitions needs to be objective and applied consistently within both the baseline and measurement period. Those members who do not qualify as Chronic are, by definition, Non-chronic, a group we refer to as Index (because we will use their experience as the source of our trend estimate).

Figure 2. Initial Member Classification*



* The boxes in this graphic are not drawn to scale.

Excluded Members

In determining the population for measurement, some members will be excluded. We assign an exclusion status to those members who, while eligible for health plan membership, may not be eligible for inclusion in the program population or the measurement population. The question of which members to exclude, and when, is a significant issue in any program measurement. An issue that appears to cause confusion in both customers of DM programs and those who measure outcomes is the difference between the managed and measured populations. This will be particularly true of health plans that use multiple vendors to manage different conditions. A managed population may be whatever the DM company and the customer agree should be managed. For example, it is not necessary that members who are excluded from the measurement be excluded from the management services. These members may be eligible for some or all of the DM

services and still be excluded from measurement. The measured population, however, will be identified by objective criteria and its outcomes will be tracked and measured in order to assess the effectiveness of the program. Exclusions are generally made from the measured population for one of the following reasons:

- The member class is not receptive to disease management. Examples of this category include those members who are residents of long-term care or other institutions and are often under the care of resident clinical personnel.

- The member is a candidate for a program, but the program is administered by another vendor, such as mental health, maternity or psych/substance abuse.

- The pattern of claims that the member exhibits is subject to sharp discontinuity, and can thus distort a trend calculation. This issue is addressed in greater detail in Appendix 1.

- The member's claims are significant, relative to other claimants in the class, and the experience of this particular group is likely to dominate the group, or introduce "noise" to the calculation.

When a member is excluded, it is important to consider the effect of the member exclusion on trend calculations (either in the intervention or Index population). Trend calculations can be affected by the prospective elimination of an excluded member at the point of identification.

More detail about member claim patterns and their potential impact on trend may be found in Appendix 1.

Measured and Non-measured Members

At the next level, we separate measured from non-measured members. Tests for inclusion in the Measurement population may include:

- *Continuous Coverage Test:* In order for a member to be eligible for inclusion in the measurement population (either Chronic or Index) the member must satisfy a

continuous coverage condition. A continuous coverage test is applied to exclude those health plan members who have less than 12 months of continuous coverage in the plan in any year, either because they are new members during the year, or because they terminated during the year. Because members are identified through administrative claims data, the identification of newly-chronic members itself takes several months due to claims processing and operational lags. In addition, newly identified or new health plan members require a start-up period to be contacted, enrolled and begin the program. For all these reasons, a six-month continuous eligibility criterion is usually applied to all members. Different periods of continuous coverage are possible; in a later paper we will examine empirically the effect that varying the continuous coverage requirement has on measurement results.

- *Claim-Free Period:* Another test, particularly important for the newly-identified members, is the claim-free period test. This test is applied as a way of addressing the issue of regression to the mean[6] in the newly-identified chronic population. Including "incident", or newly-identified chronic members from the month of identification, when the identifying event is a hospital admission potentially builds-in regression to the mean (generally a reduction in claims) because the hospital admission usually represents the highest point of utilization for that year. Failure to eliminate the reduction in claims due to the natural course of recovery, risks assigning causality and savings to the program that result from the natural course of the event. Defining a comparable population that includes newly-identified members, and separately tracking prevalent and incident members can address this issue. However, since failure to satisfactorily address regression to the mean in claims data is probably the single most significant source of criticism by health plans, the conservative approach of completely eliminating the identifying event and its reversal is recommended here.

[6] This topic is covered in greater detail in Paper 2.

There is another reason for excluding newly-identified members in the first four months since their identifying event: inclusion of the event itself makes the DM vendor effectively responsible for that event. While there may be ways of identifying members at risk of events who have no prior history of the condition, we doubt that most vendors would agree to be held accountable for reducing events in a population with no prior history of the event. Consistent with the principle that we should measure what we have agreed to manage (and for which we have agreed to be held accountable) we recommend excluding these events unless the DM vendor explicitly agrees to manage and be held accountable for them.

Engaged/Enrolled, Targeted and Reachable Members

In the final line of Figure 2, we illustrate several different outcome states in the "Measured" member category: enrolled, not enrolled and unreachable members. Because this is a population methodology, we measure the outcomes of all members, whether enrolled, not enrolled, reachable or unreachable and targeted or not targeted for intervention.

The issue of "reachability" is an important one for program management and comparison of outcomes. In our experience the typical PPO health plan lacks accurate, up to date contact information on 30 percent to 40 percent of its membership. Restricting measurement only to those members with valid contact information potentially introduces bias into the measurement, and we do not recommend excluding unreachable members. We do, however, recommend reporting data on contact information as part of the reporting of outcomes. Limiting measurement to enrolled members only similarly introduces bias to the results. For comparison purposes the enrollment rate (as a percentage of reachable members) may be reported as part of the reporting of outcomes. Finally, we recognize that not all members may be "targeted" for a program. Some may not have a sufficiently serious condition (based on the identification criteria used for the program) to warrant management by clinical resources. As with other states, however, all members (whether targeted or not) who meet the identification criteria should be measured.

Time Periods

The adjusted historical methodology incorporates the following time periods:

- *The Lookback Period:* Baseline members are identified through claims identified in the "lookback period" or claims identification prior to the beginning of the baseline period (usually, but not always, twelve months).

- *Baseline Period:* The period prior to the start of a program in which the reference population is identified. This period also forms the lookback period for identification of members in the first measurement period.

- *Measurement or Intervention Periods:* The periods during which the program outcomes will be measured. Measurement periods need not be sequential, or sequential to the baseline period, although they are usually close.

Sometimes, a ramp-up period is also imposed, during which measurement does not take place, allowing the program to become established and enrollments to be performed.

Chronic Members and Chronic Prevalence

There is no unique way of identifying who has a chronic disease. In order to be useful for measurement, however, an identification algorithm needs to be objective, stable over time, and cheap to apply. These criteria rule-out many of the methods that involve clinical resources and chart review, and result in administrative claims-based criteria being used in most population studies. Many different claims-based definitions of chronic condition exist.[7] Many health plans use risk-adjuster methods to classify and rank members by risk class, and many actuaries are familiar with these.[8]

---

[7] See for example: HEDIS 2003 Technical Specifications published by the National Council on Quality Assessment. 2004; or Ian Duncan, ed: 2004. *Dictionary of Disease Management Terminology*. DMAA.

[8] See, for example: Cumming, R. B., D. Knutson, B. A. Cameron, and B. A. Derrick. 2002. *A Comparative Analysis of Claims-Based Methods of Health Risk Assessment for Commercial Populations.* Society of Actuaries; or Dove, H., I. G. Duncan, and A. S. Robb. 2003. *A Prediction Model for Targeting Low-Cost, High-Risk Members of Managed Care Organizations*. American Journal of Managed Care. *9(5)* 381-389.

Data: Available Sources

Chronically ill members are often identified from claims data, so data is a central issue. Unfortunately, there is no ideal source of data. Each source has its advantages and drawbacks, which must be weighed against each other. Five types of data commonly available to the health care analyst are medical charts, survey data, medical claims, pharmacy claims and laboratory values.

Generally, we favor identification using integrated medical and pharmacy claims, although care needs to be taken with PPO and other commercial plans where employers often carve out pharmacy benefits. Claims data are not as rich or accurate as survey or medical chart data, but are always available and are generally of sufficient quality to drive risk management programs.

When using claims data for chronic identification, the actuary should consider the problem of false negative and false positive identification.

The Problem of False Negatives

False negatives are chronic members who are "missed" by an identification algorithm. These members are more of a problem for program management than for program measurement. To the extent that a member has a condition that is untreated, claims data will be unavailable and the member will be unidentified. A more difficult false negative problem occurs when the member's provider is not part of the data-submission system (for example, when a member obtains drugs from the Veterans Administration system, or buys them in Canada). Eventually, even these members will have claims for a service that is included on the data-reporting system, and thus be identified. But until this happens, the member will be classified as non-chronic for the purpose of measurement. Different results are obtained, depending on the specific definitions used for identifying chronic members. The following Table illustrates the identification of chronic prevalence using three different sets of criteria applied to the same set of data.

Figure 3. Chronic prevalence* according to different identification criteria

Prevalence of 5 Chronic conditions

	Narrow	Broad	Rx
Medicare	24.4%	32.8%	30.8%
Commercial	4.7%	6.3%	6.6%

Definitions:

Narrow: Hospital Inpatient claims, using the primary diagnosis on the claim or face-to-face office visits only (excluding tests and other services that are not face-to-face).

Broad: Hospital Inpatient claims, using any recorded diagnosis on the claim, plus any professional services, including tests.

Rx: Narrow plus Outpatient prescription drug claims.

Medicare represents a Medicare Risk population with drug benefits.

* Duplicates (i.e. incidence of members with more than one disease) have been removed.

The Problem of False Positives

False positives are members who are falsely identified as having a chronic condition, when they do not have that condition. There are two types of false positives: clinical and statistical. Clinical false positives, as the name implies, are those members who are identified with the condition and later found not to have it. Statistical false positives, on the other hand, arise because the administrative claims used for identification will never be complete, unambiguous or correctly coded. When identification of chronic conditions takes place from administrative claims data, there is a chance of statistical false positives (which may be different than clinical false positive identification). We define statistical false positives as those members who meet a chronic definition in Year 1, but who do not re-qualify according to the same set of definitional criteria in Year 2. This issue is important for disease management outcomes evaluation because false positives, who do not have the condition according to the claims data, are less likely to have high costs. Therefore their continued inclusion in the chronic population, although they no longer meet chronic definition criteria, will likely reduce the average cost (and therefore the trend) in the chronic population, resulting in apparent reduction in cost due to the program.

We should also note that a set of criteria appropriate for identifying members for one purpose may not be the most appropriate for another. For example, one use of identification criteria may be to find members for a management program, and another is to identify members for measurement. In the first instance, specificity is not as important as sensitivity (we need to identify as many members as possible with the condition to implement a successful program). For measurement or other examples involving financial objectives, such as reimbursement of providers, we need to be reasonably certain that the identified population actually have the condition.

Figure 4 illustrates the results of three different sets of identification criteria used in Figure 3, applied to populations in two years.

Figure 4. Prevalence of statistical false positives in a chronic population

		Year 1			
		Narrow	+ Broad	+ Rx	TOTAL
Year 2	Narrow	75.9%			
	+ Broad		85.5%		
	+ Rx			92.6%	
	Not Identified	24.1%	14.5%	7.4%	
	TOTAL	100.0%	100.0%	100.0%	100.0%

In this example, members who are not identified in Year 2 are those members who do not meet the identifying criteria through claims in Year 2. We identify them as "statistical false positives". By definition, since these members meet neither exclusionary nor chronic definitions, they have lower average costs than the chronic group. Thus, keeping them in the chronic group will tend to cause the chronic group's claims to be lower, introducing the potential for bias, and an over-statement of savings that have otherwise occurred.

Included and Excluded Claims

A DM program aims to intervene with members of a health plan who are at risk of medical events (emergency room visits, specialist visits and admissions) for their condition. However, in any system, these members could also consume resources for conditions that

are not subject to management by the DM program, for example trauma, accident, psychiatric, substance abuse or maternity conditions. It is customary in DM evaluations to exclude those conditions that are specifically outside the program, because they are subject to random fluctuation into an evaluation. If the random fluctuation is large enough it could dominate the true effect being measured. An example of such a list of excluded claims, based on ICD-9 codes, is included in Appendix 2.

Results and Use

Once members are appropriately assigned to categories, monthly numbers may be aggregated into measurement years, and calculations may be performed with the resulting totals. Below, we illustrate an actual application of the actuarially-adjusted historical control methodology.

We first summarize data according to the chosen categories for the analysis. In each month of observation, we record the number of Total, Chronic and Index measured members (Chronic and Index non-measured members are not shown). The corresponding cost per member per month for each group is also summarized, allowing us to calculate the index trend used in the savings calculation. In this example, data are accumulated over two periods: the program baseline period, which begins in August, 2000, and the first program measurement period, which begins in the month of October, 2001. Note that for this analysis, a two-month measurement-free period (to allow for program start-up) has been applied in the months of August and September, 2001.

Figure 5. Data summary for use in savings calculation

LINE OF BUSINESS **ILLUSTRATIVE**

		January	February	March	April	May	June	July	August	September	October	November	December	BASE LINE	INTER VENTION
Members	**2000**								100,000	100,000	100,000	100,000	100,000		
	2001	100,000	100,000	100,000	100,000	100,000	100,000	100,000			100,000	100,000	100,000	1,200,000	
	2002	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000	100,000					1,200,000
Chronic	**2000**								20,000	20,000	20,000	20,000	20,000		
Measured	**2001**	20,000	20,000	20,000	20,000	20,000	20,000	20,000			20,000	20,000	20,000	240,000	
Members	**2002**	20,000	20,000	20,000	20,000	20,000	20,000	20,000	20,000	20,000					240,000
Index	**2000**								60,000	60,000	60,000	60,000	60,000		
Measured	**2001**	60,000	60,000	60,000	60,000	60,000	60,000	60,000			60,000	60,000	60,000	720,000	
Members	**2002**	60,000	60,000	60,000	60,000	60,000	60,000	60,000	60,000	60,000					720,000
TOTAL	**2000**								$ 10,000,000	$ 10,000,000	$ 10,000,000	$ 10,000,000	$ 10,000,000		
COST	**2001**	$ 10,000,000	$ 10,000,000	$ 10,000,000	$ 10,000,000	$ 10,000,000	$ 10,000,000	$ 10,000,000			$ 10,500,000	$ 10,500,000	$ 10,500,000	$ 120,000,000	
Chronic															
Measured	**2002**	$ 10,500,000	$ 10,500,000	$ 10,500,000	$ 10,500,000	$ 10,500,000	$ 10,500,000	$ 10,500,000	$ 10,500,000	$ 10,500,000					$ 126,000,000
TOTAL	**2000**								$ 12,500,000	$ 12,500,000	$ 12,500,000	$ 12,500,000	$ 12,500,000		
COST	**2001**	$ 12,500,000	$ 12,500,000	$ 12,500,000	$ 12,500,000	$ 12,500,000	$ 12,500,000	$ 12,500,000			$ 14,000,000	$ 14,000,000	$ 14,000,000	$ 150,000,000	
Index Measured	**2002**	$ 14,000,000	$ 14,000,000	$ 14,000,000	$ 14,000,000	$ 14,000,000	$ 14,000,000	$ 14,000,000	$ 14,000,000	$ 14,000,000					$ 168,000,000
COST	**2000**								$ 500	$ 500	$ 500	$ 500	$ 500		
PMPM	**2001**	$ 500	$ 500	$ 500	$ 500	$ 500	$ 500	$ 500			$ 525	$ 525	$ 525	$ 6,000	
Chronic															
Measured	**2002**	$ 525	$ 525	$ 525	$ 525	$ 525	$ 525	$ 525	$ 525	$ 525					$ 6,300
COST	**2000**								$ 208	$ 208	$ 208	$ 208	$ 208		
PMPM	**2001**	$ 208	$ 208	$ 208	$ 208	$ 208	$ 208	$ 208			$ 233	$ 233	$ 233	$ 2,500	
Index Measured	**2002**	$ 233	$ 233	$ 233	$ 233	$ 233	$ 233	$ 233	$ 233	$ 233					$ 2,800
INDEX TREND															12%

Using the data from Figure 5 in the table below, the savings calculation is as follows:

Estimated Savings due to reduced pmpy =	
Baseline Cost pmpy * Cost Trend	$6,000 * 1.12 = $6,720
Minus: Actual Cost pmpy	$6,300
Equals: Reduced Cost pmpy	$420
Multiplied by: Actual member years in Measurement Period	20,000
Equals: Estimated Savings	$8,400,000

Conclusion

In this paper we have described the actuarially-adjusted historical control methodology for performing DM outcomes evaluations. This methodology provides practical solutions to many of the actuarial issues raised in earlier papers. Previously, we have highlighted the issues of ensuring equivalence within a control group, regression to the mean, reconciliation to the source data, and effect of different member types on the trend and savings calculations. The techniques outlined in this paper are designed to control for many of these problems. In any evaluation there will be issues that have not previously been encountered. We believe that our work provides a framework to address most issues that the actuary is likely to encounter in practice.

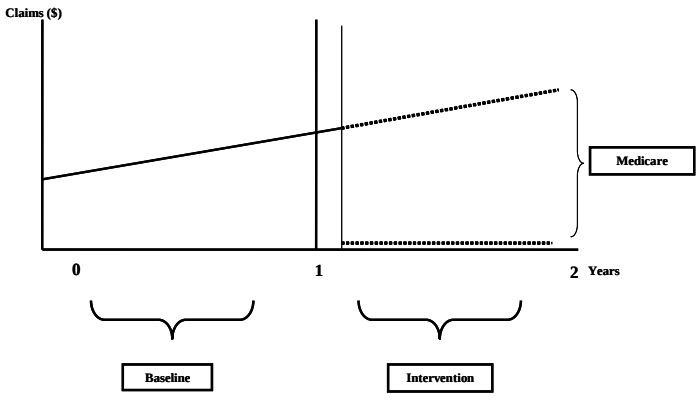
APPENDIX 1: Members Excluded from Measurement

The principle that we apply when determining whether to include or exclude a member is: will the member inclusion or exclusion contribute to a significant discontinuity in claims? For example, if a member were to be excluded only after the member's claims amounted to $100,000 in a year, the member's claim of $100,000 would be part of the baseline experience, while the member would contribute zero to the intervention period, thus potentially affecting the measured trend between the two periods. Below, we show some examples of individual claims patterns. The reader may extend the principles, however, to any other claim pattern.

The following are examples of specific exclusionary conditions, and how they may be handled:

- *End-Stage Renal Disease (ESRD):* The course of ESRD is progressive over time, and management of the condition, while it may delay cost, cannot ultimately reduce or postpone those costs. Claims tend to follow the example below.

  Figure 6. ESRD Claims



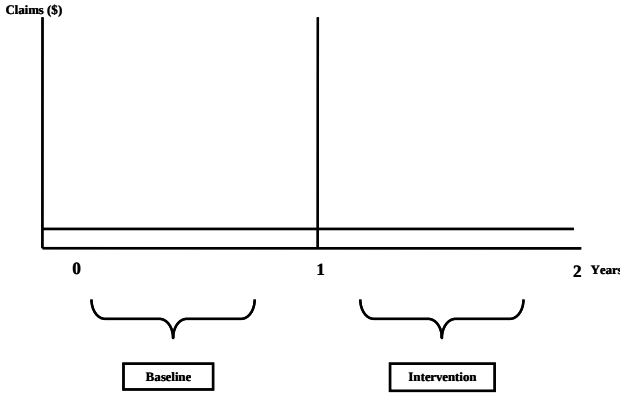
  However, a secular upward trend in claims is insufficient reason to exclude these members. A more compelling reason (for Commercial members) is the discontinuity that occurs at 33 months after first dialysis treatment when Medicare

accepts payment for these members as part of the Medicare End-Stage Renal Disease program. This pattern is illustrated above, as the health plan's responsibility falls to near-zero at the point that the member is eligible for the Medicare program. Failure to recognize this discontinuity may distort the comparison of experience and trend over time. In our work, we exclude these members permanently and retroactively from measurement because the condition is permanent. For Medicare members, where the discontinuity does not exist, the member may be left in the group or retroactively excluded. Retroactive exclusion obviously reduces any potential distortion.

- *Transplants:* Members who have a transplant often experience high and rising claims up to a period shortly after the transplant, at which point the claims are reduced and stabilized. The claims pattern is similar to that of the ESRD member above, although for different reasons. A member who undergoes a transplant should probably be excluded. We recommend retroactive exclusion in order to avoid potential distortion.

- *HIV/AIDS, Mental Health and other Conditions:* for which privacy issues make it difficult or impossible for a vendor to receive complete data feeds, or manage the member. Claims for these members may follow a reasonably regular pattern, and are likely to be lower than the claims of a member with End-Stage Renal Disease. The claim pattern of the member below is an example. For these members, exclusion can occur either prospectively from the point of first identification or retroactively to the beginning of the baseline period.

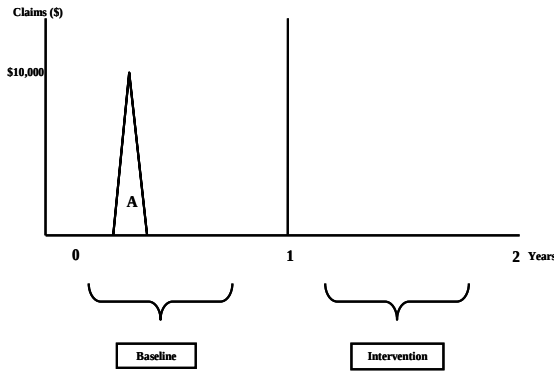
Figure 7. HIV/AIDS Claims



- *Members who are institutionalized:* or who have a history of institutionalization (mental health, hospice, or nursing home) are examples of members who are not reachable or who may not be able to benefit from disease management interventions. These members are also excluded permanently from measurement because their condition or status is more likely to make them permanently unsuitable for the program. These members often have high costs prior to identification with the exclusionary condition, so we recommend exclusion permanently and retroactively.

- *Members with catastrophic claims:* represent members who are not manageable by the DM program, and who are often subject to management by another program (for example, catastrophic case management). The member below is an example of a member with a random, catastrophic claim in the baseline period that is not repeated in the intervention year. These claims tend to be excluded above a stop-loss point, or through the exclusion of the entire member experience for the year. Because these events tend to be acute, traumatic or accidental in nature, members who are excluded in one period are eligible for inclusion in measurement in a subsequent period if they recover and continue to be eligible members. In some instances, a health plan may purchase specific stop-loss coverage. In this case, the specific stop-loss attachment point may be an appropriate level at which to impose a cut-off for evaluation purposes.

Figure 8. Catastrophic Claims



There is a common belief in the industry that a chronic population must be at significant risk of catastrophic claims, because average member cost is high. It is true that chronic patients are more likely to experience high costs (for example, costs above $100,000). However, the incidence of these claimants in a chronic population is still relatively rare. As an example, some data on the distribution of claims by amount within a commercial HMO is provided below. Because this data has been taken from a disease management analysis, the excluded category consists of both members excluded for condition (End-Stage Renal Disease, HIV/AIDS, transplant or institutionalization) as well as members excluded because they are in the population for less than 6 months. This accounts for the relatively large percentage of members who have very small claims. Once the small claims amounts of the short-term members are excluded, most of the claims are skewed towards higher amounts.

Distribution of Members within each sub-population						
Group	< $1,000	$1,000 - $9,999	$10,000 - $49,999	$50,000 - $99,999	$100,000 +	TOTAL
Chronic	52.04%	39.07%	8.42%	0.41%	0.06%	100.00%
Non-Chronic	82.33%	16.62%	1.01%	0.03%	0.00%	100.00%
Excluded	83.91%	13.33%	2.20%	0.32%	0.24%	100.00%
TOTAL	81.00%	17.35%	1.52%	0.09%	0.04%	100.00%

Distribution of Costs within each sub-population cost						
Group	< $1,000	$1,000 - $9,999	$10,000 - $49,999	$50,000 - $99,999	$100,000 +	TOTAL
Chronic	6.61%	37.94%	46.21%	6.77%	2.47%	100.00%
Non-Chronic	25.37%	52.43%	19.30%	2.12%	0.78%	100.00%
Excluded	9.44%	24.06%	29.98%	12.70%	23.82%	100.00%
TOTAL	19.46%	44.81%	25.46%	4.85%	5.43%	100.00%

We do not have a breakdown of the excluded members by chronic/non-chronic status, but the fact that the member has an excluded condition implies that the member's cost is largely driven by the exclusionary condition. The percentage of chronic member costs that arise from members with costs in excess of $100,000 is 2.47 percent of chronic costs, and the portion of non-chronic costs in excess of $100,000 is even smaller. Thus, a health plan with specific stop-loss above $100,000 is not likely to be much affected by the catastrophic claims exclusion. Even an employer with a lower stop-loss limit will not find much of his dollars excluded. For example, at the $50,000 level, 9.24 percent of claims are for chronic claimants with total claims in excess of $50,000, and 2.90 percent are for non-chronic claimants with total claims in excess of $50,000.

- *Members who are eligible for other management programs:* such as members who are participating in case management, or eligible for another disease management program (not part of the measurement program). These members should be excluded based on objective criteria, prospectively from the point of identification as eligible for the program. Members in case management represent a particularly difficult issue for measurement. Because of the selection that exists in the enrollment of patients in such programs, the exclusion of enrolled case management patients would result in bias. Ideally, a set of objective, claims based criteria would exist that would allow the identification of a case management eligible population that may then be excluded as a class from the measurement population. Well-defined, objective criteria exist for many targets for case

management (many of the categories discussed here are candidates). However, many candidates are referred by providers rather than identified through claims, making objective identification difficult. A compromise solution may be to include all members in the measurement, irrespective of the program and who is performing the management. The outcomes so measured would be those for the combined programs. The overall savings may then be split into those from different programs on a reasonable basis agreed between the parties.

Figure 9. Member Exclusions: Summary

Type of Condition	Why Exclude?	Whether to Exclude	When to Exclude?
End-Stage Renal Disease	High and increasing claims period to period. Claim discontinuity when CMS becomes responsible for claims	Medicare: No Commercial: Yes	Medicare: may be included, if the ESRD population is large. Commercial: exclude retrospectively.
HIV/AIDS	Claim discontinuity (increase) post-diagnosis.	Yes	Retrospectively or Prospectively
Transplants	Claim discontinuity (reduction) post-transplant.	Yes	Retrospectively or Prospectively
Institutionalized	Data not always available (psych) and population difficult to manage. Population may already be under full-time management (hospice; long-term care)	Yes	Prospectively
Members with catastrophic claims (e.g. > $100,000)	Significant utilization relative to other members; creates noise and potentially distorts comparison	Yes	Retrospectively for current period only; alternatively exclude claims above a stop-loss point.
Members eligible for other programs	Often the responsibility of another program or vendor.	Yes	Retrospectively (if pre-identification claims are significant) otherwise prospectively; alternatively, include in population and measure overall effect of multiple programs.

APPENDIX 2: Example of Claims Exclusion Criteria

Caveat

The identification of claims for exclusion should take into account the availability of detailed information in the claim system. For example, the level of diagnosis information that is retained within the claim system is a carrier decision. Some carriers will retain multiple diagnoses; others may retain only the primary diagnosis.

In the event that it is demonstrated that an exclusionary condition is related to the chronic condition (for example, if an accident or trauma event were to result from a hypoglycemic episode in a diabetic), the actuary may choose to include a claim that would otherwise be excluded.

Claims with a primary or secondary diagnosis within the following ranges are excluded from measurement.

1. Trauma and Accident

   Typical trauma exclusions include bone fractures, injuries and burns. These claims cover the range of 8xx.xx and 9xx.xx ICD9 series.

Condition	Codes
Fractures	800 – 829
Dislocations	830 – 839
Sprains & Strains	840 – 849
Injuries & Open Wounds Traumatic Complications	850 – 904, 910 – 939, 950 – 959
Late Effects of Injuries, Poisonings, Toxic Effects and Other External Causes	905 – 909
Burns	940 – 949
Poisoning by Drugs, Medicinal and Biological Substances	960 – 979
Toxic Effects of Substances Chiefly Nonmedicinal as to Source	980 – 989
Other and Unspecified Effects of External Cause	990 – 994
Complications of Surgical and Medical Care NEC	995 – 999

2. Psych/Substance Abuse

As we discussed above, members who have a psych/substance abuse diagnosis are not good candidates for a DM program. Often, a health plan carves-out these services and places them with a specialty vendor. It is sometimes difficult to obtain the full history of psychiatric or substance abuse claims in this instance. Members with a history of institutionalization may be under full-time care of a provider, or may not be at the point in recovery where self-care is an option. Nevertheless, this exclusion is likely to be controversial, particularly when the customer and the vendor explicitly agree that the DM program should cover these members.

3. Malignant Neoplasms

Excluded claims are those with diagnosis codes in the range greater than or equal to 140 and strictly less than 210. In addition, claims in the range V10.x are excluded. Cancer is another condition that DM programs are not generally able to manage, and which is often subject to management by a specialty case management program. We do not, however, argue for complete exclusion of members with a cancer diagnosis. Depending on the specific criteria used to identify patients, this could represent a large subset of the chronic population, particularly if the criteria include members who have a prior history of cancer but who are now in remission. These members often represent appropriate candidates for chronic disease management and their measurement is appropriate.

4. Maternity and Childbirth Claims

Unless the DM program targets maternity, maternity should be excluded because a standard chronic program will not cover these conditions. Maternity exclusion criteria are based on primary diagnosis codes within the standard maternity-related ranges identified as normal delivery and "Complications of Pregnancy, Childbirth and the Puerperium" (Diagnosis codes 630 – 679). These codes include:

- Ectopic and Molar Pregnancy
- Other Pregnancy with Abortive Outcome

- Complications Mainly Related to Pregnancy
- Normal Delivery and other Indications for Care in Pregnancy Labor and Delivery
- Complications Occurring Mainly in the Course of Labor and Delivery
- Complications of the Puerperium

In addition, maternity exclusion criteria include appropriate "V" codes associated with pregnancy management. These codes include:

- V22 Normal Pregnancy
- V23 Supervision of High Risk Pregnancy
- V24 Postpartum Care and Examination
- V26 Procreative Management
- V27 Outcome of Delivery
- V28 Antenatal Screening

5. Pharmaceutical Drugs

The exclusion of outpatient pharmaceutical drug claims (retail and mail-order) is probably the most controversial category of potential exclusion. Pharmaceutical drug claims may be a candidate for exclusion, particularly in a large employer or self-insured environment, because this coverage is highly volatile. For example, this coverage is subject to change in benefits design, provider, etc. on a more frequent basis than hospital or physician coverage.

APPENDIX 3: An example of member classification over time

Figure 10. Example of application of member classification over time

This member has no exclusionary conditions and experiences the first chronic identification event (an Emergency Room visit for Chronic Obstructive Pulmonary Disease) on 3/15/2002. The member is continuously eligible 2001 through 2005.

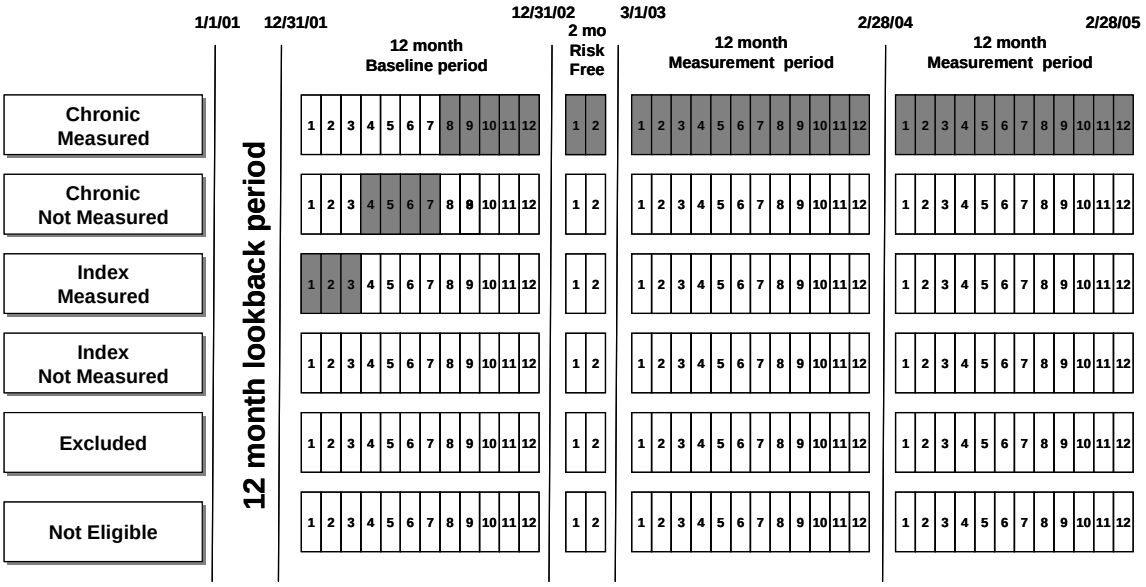


Figure 10 is an example of the application of the exposure classification to a particular member. In the grid, the member's progression over time between classifications may be clearly seen. In this example the member was enrolled in the plan prior to the baseline period (which begins 1/1/02). The member was initially non-chronic and had more than six months of prior eligibility, so was classified as Index Measured for the first three months of the baseline period. (The member does not have any evidence of exclusionary conditions that would result in the member being included in the excluded category.) The member was then identified as Chronic due to a claim in the third month of the baseline period. The member is therefore Incident (newly-identified) Chronic Non-Measured for the four months required for the member to reach the claim-free status that allows the member to be classified as Chronic Measured. The member will be Chronic Measured for the

balance of the measurement period, and (assuming no requirement to re-qualify under the chronic definition and no change in eligibility) will continue in this segment until eligibility ceases.

Several optional choices for a health plan are illustrated by this example: Is the definition of chronic condition “once chronic always chronic” or is some form of regular re-qualification required? Will members who terminate eligibility and then rejoin the health plan receive credit for prior membership (to qualify as measured) or claims/chronic status (to qualify as chronic)?