Paper 8: Testing Actuarial Methods for Evaluating Disease Management Savings Outcomes[1]

Ian Duncan[2], FSA, FIA, FCIA, MAAA, Rebecca Owen, FSA, FCA, MAAA, Henry Dove, Ph.D.

May 21, 2006

Abstract

This paper applies the principles and methodologies discussed in earlier papers in this series to measure cost savings from a two-year disease management (DM) program for the 200,000 Medicare members of a regional managed care plan. We then study how savings estimates change when we allow for four variations in key assumptions: methods for identifying or excluding plan members, requirements for eligibility in the health plan, methods for calculating health care trend and study design. Finally, we also examine the sensitivity of these results to product type by comparing the Medicare results to those of the Commercial HMO/POS product.

Background

Highmark, Inc. is a large regional managed care plan headquartered in Pennsylvania, with more than 4 million members, most of whom reside in and around Pittsburgh. During the period of this study, approximately 2-1/2 million members were in products eligible for the DM program. Eligible members were distributed in two main product groupings:

- Medicare Plus Choice (now Medicare Advantage): 200,000. Throughout this paper we refer to these members as "Medicare" because of the change of designation.
- HMO and Point of Service: 1.1 million

Highmark has been a leader in applying innovative managed care approaches to improving health and reducing cost of its coverage. For several years, Highmark operated a number of medical management and DM programs, both internally staffed and using outside vendors. In 2002, Highmark implemented new DM programs operated by Health Dialog, Inc. a Boston-based care management company. The DM programs were first offered to the Medicare and HMO/POS populations.

[1] Paper 8 of a series of papers entitled: "Evaluating the Results of Care Management Interventions: Comparative Analysis of Different Outcomes Measures" sponsored by the Society of Actuaries' Health Section and the Committee on Knowledge Extension Research.

[2] Solucia, Inc., Hartford, CT (Duncan, Owen). Yale School of Public Health (Dove).

Highmark and Health Dialog wished to estimate cost savings from the DM programs. After considering several different methods, they ultimately decided to use the actuarial methodology described in Paper 6 in this series, which resulted in the savings results that we term the "base case savings."

Highmark's management is aware of the complexity associated with calculation methods used to determine the cost effectiveness of DM programs. As a national leader in the evolution and understanding of effective delivery of health care to members, Highmark's management supported the Society of Actuaries study aimed at quantifying the sensitivity of estimated savings results to variations in the specific components and key assumptions of the actuarial methodology.

Methods

We compare the effect on estimated program savings of changes in different assumptions. All variations are compared with the "base case savings." "Base case savings" are calculated by applying the actuarial methodology described in Paper 6. We report here on the impact on cost savings of four variations of the base-case results from the actuarial methodology, applying each variation to the same underlying dataset (with the exception of the Commercial HMO/POS results, which are obtained from the HMO/POS dataset, applying the base-case assumptions).

Base Case Assumptions

The base case consisted of a number of assumptions, as follows:

1. The base case is a "population" study, in which a "baseline" population measure is compared with the same population measure, calculated in the "intervention period." As discussed elsewhere in this series of papers, the population methodology is used to overcome many of the potential objections that arise from "regression to the mean" that is observed in a cohort of selected, high-risk members.
2. "Chronic" patients identified from diagnoses from medical claims or prescriptions. More details of the specific identification criteria used to classify chronic members are found in Appendix 1. Not all claims were included in the study. Specifically those claims that are subject to volatility (catastrophic claims for example) or are not manageable by the program (for example, maternity claims) are excluded. Details of all categories of excluded claims are found in Appendix 2.
3. A minimum of six months of continuous health plan eligibility is required to be included in the study.
4. No "re-qualification" (as a chronic member) required: once a patient is identified as a "chronic" member, he is always considered to be a chronic member (this definition is consistent with the clinical view of chronic disease).

5. The reference cost is estimated by applying a medical trend factor to the baseline cost of the chronic population.
6. The medical trend that is applied to the chronic population is that derived from the non-chronic population. The assumption underlying the use of the non-chronic trend as an adjustment is that the factors that cause increases in chronic population costs (for example, changes in provider contracts and practice patterns, introduction of new technology, etc.) apply equally in the non-chronic population, and can therefore be estimated from them. Trend is defined as the increase in per member per month (PMPM) cost in a "measured" population. All exclusions that apply to the chronic population (for example, for catastrophic cases, or of members who do not have six months of continuous eligibility) are applied to the non-chronic population when calculating this trend.
7. Cost levels used in the study are allowed charges. While plan design may have some impact on results through its effect on utilization, for the most part the potential for confounding of the effect of changes in plan design is neutralized by the use of allowed charges.[3]

Limitations

Our study, to our knowledge, is the first published instance in which the effect of changes in various methodological approaches and assumptions is tested to gauge the sensitivity of DM cost savings estimates. Further research is needed in this area. For example, in our analysis a series of single alternative assumptions is considered and compared to the base-case cost savings. A multivariate approach might be useful in which the impact of different assumption alternatives would be combined to replicate a particular study and assess the effect on base-case savings.

Due to limited resources, we restricted the analysis of the impact of the different alternatives to Highmark's Medicare plan. We did have access to the HMO/POS plan data and determined base-case savings for these plans to compare to the Medicare results. We have reported the base-case savings for both Medicare and Commercial products, which (on a PMPM basis) are similar. The similarity between Medicare and Commercial baseline savings suggests that the sensitivity analysis for Commercial populations may be similar to that of the Medicare population.

[3] Highmark has recently seen increasing sales of consumer-directed health care plans in which the effect of plan design on utilization is expected to be more significant than in traditional plans. However, during the period of this study, consumer-directed health plans were not a significant component of Highmark's business.

Our methodology includes certain exclusion criteria, both for members and claims. We have listed certain diseases or patient profiles that are commonly excluded in Appendix 2. We have studied different ways of identifying chronic patients. We have not studied the risk profile of patients who leave managed care organizations, patients who refuse to participate in DM programs or those who terminate from them. It was beyond the scope of the present study and the available data to track such measures as the number of enrolled members, the length of time in the program for each member and the rate of termination from the program.

Another issue of interest beyond our study is a disease-specific focus. In this DM program, Health Dialog intervened in a variety of chronic diseases. Health Dialog's program serves multiple conditions, as do those programs of an increasing number of DM vendors. Indeed, Health Dialog refers to their program as a "whole person" model and makes little or no distinction in terms of services provided to chronic members based on conditions and co-morbidities. Purchasers of DM programs and researchers are often interested in the cost savings associated with each disease. It is reasonable to expect that savings from DM would vary by and within disease type. However, this level of analysis is also beyond our scope at this time.

Finally, this paper estimates the sensitivity of calculated claims savings under a number of reasonable assumptions. Purchasers of DM programs usually require an estimate of a program's return on investment (ROI). Estimates of ROI would require information on program costs and costs of administration which were not available to us, and which therefore make the calculation of an ROI impossible.

Alternative Scenario Assumptions

In testing the sensitivity of results to changes in assumptions, all base-case assumptions are held constant and only the specific changes noted below are made. Specific variations analyzed are as follows:

1. Cohort analysis—Savings are calculated based on a before-and-after analysis of a cohort of plan members in a DM program. This methodology (also called "pre-post") was once prevalent in the industry but following publication of methodology analyses such as those by Johns Hopkins (2002), Fetterolf, Wennberg and DeVries (2003), Fitzner (2004) and others, is now less frequently encountered in the literature. We wished to test the difference in results that is observed by applying both population (base-case) and cohort (Alternative 1 below) methods to the same set of data.

2. Definitions/techniques to identify "chronic" patients:
    a. Based on medical claims but *not* prescription drug claims;
    b. Based on primary diagnosis only, as identified from claims;
    c. Based only on claims submitted by hospitals (for inpatient or outpatient services);

3. Trend. We evaluate the use of a non-chronic trend calculated using a retrospective chronic identification algorithm.
    In the base-case, "non-chronic" member trend is used as a proxy for the (unmeasurable) chronic member trend, absent intervention. In this base-case, non-chronic member trend is calculated using a prospective classification of chronic condition (that is, members are assigned to the chronic population from the month that they first meet the chronic identification criteria forward). Alternative approaches to calculating trend are discussed in a related paper (Paper 7 in this series): *"A comparative analysis of chronic and non-chronic insured health plan member cost trends*" (Bachler, R., Duncan, I., and Juster, I.), submitted to the Society of Actuaries Chronic Care Financing Seminar, June 2005. In this paper, the authors identify a bias (called "migration bias") due to the migration of members between groups. One way to mitigate this bias is the use of retrospective identification; that is, members identified as chronic are classified in the chronic group throughout the study beginning with the earlier of the date that the study begins or the date the member joins the plan (including the baseline period). While this technique avoids migration bias, it introduces other potential distortions (we measure the results for members who are not part of the managed population, for example). The extent and nature of this potential distortion remains to be analyzed.

4. Continuous enrollment/eligibility definitions
    To be included as a chronic or a non-chronic member in this analysis the member must have been continuously enrolled for a minimum of six months. Members with fewer than six months of continuous enrollment are included at the beginning of the month after completion of six months of enrollment. For this alternative, we studied the effect on the results of imposing no continuous enrollment requirement.

The base analysis is conducted using the Medicare product line. Finally, we apply the same assumptions to calculations made for the Commercial HMO/POS lines.

Enrollees and claims were tracked for a baseline year (pre-program initiation) and two subsequent intervention years. The alternative methods described above were then applied to recalculate the cost savings.

RESULTS

Base Case Results

In Table 1 we apply the measurement methodology as described in Paper 6 to a baseline and two intervention periods in the Highmark Medicare population. The trend that is applied to the chronic measured population is calculated from the experience of the non-chronic, or index, measured population. Note that we apply a two-month non-measured period at the end of the baseline period to allow for the program to start and for enrollment to take place. Thus, the first period for which trend is calculated is actually 14 months, not 12. The initial chronic measured prevalence is 21.3 percent, indicating that slightly more than one in five members has at least one chronic condition and meets all the other requirements for inclusion in the study. The chronic measured prevalence grows by 28 percent over the three-year study period, to 27.2 percent. It is worth noting that these chronic measurements span a time period during which changes in CMS payments caused plans to begin to be more thorough in both the completeness and the coding accuracy of the diagnosis on claims submissions, which could have resulted in a small increase in the identification of chronic members in the later periods.

As one might expect, the basic cost PMPM of the chronic population is considerably higher than that of the non-chronic population ($448.26 vs. $170.84). Estimated savings in intervention year 1 are 8.4 percent of the estimated chronic cost PMPM, or 2.0 percent of all claims. Intervention year 2 savings are estimated at 12.1 percent (3.0 percent of all claims). As several commentators have noted, estimated savings from the trend-adjusted method, used here, increase considerably from the first to the second period. Non-chronic trends are 9.7 percent and 9.9 percent, respectively, in intervention years 1 and 2. Chronic trend is lower than the trend of the non-chronic population. (For a more complete discussion of factors that influence chronic trend measurement, see Paper 7.)

Table 1: Highmark Medicare Base Case Savings Calculation

Alternative 1: Cohort (Pre-Post) Analysis

Measure	Baseline 8/00 – 7/01	Intervention Year 1 10/01 – 9/02	Intervention Year 2 10/02 – 9/03
Ave. no. Members	158,177	180,290	186,246
Ave. no. Chronic Measured Members	33,628	44,251	50,739
Chronic Measured Prevalence	21.3%	24.5%	27.2%
Trend (PMPM, Allowed Cost)			
Chronic Measured Population		0.5%	5.5%
Index Measured Population		9.7%	9.9%
Claims PMPM, Index Measured Population	$170.84	$187.46	$206.01
Claims PMPM, Chronic Measured Population			
Projected	$448.26	$491.88	$540.55
Actual	$448.26	$450.34	$475.27
Total Cost Savings, PMPM		$41.54	$65.28
Total Savings ($ millions)		$22.1	$39.7
Savings as % of total claims for the line-of-business		2.0%	3.0%

Many earlier studies of DM savings outcomes were performed comparing the historical and current experience of a cohort ("pre-post"). This methodology has come under increasing criticism in the literature and from consultants, who have pointed out the potential for "regression to the mean" to be observed (and counted as savings) when the cohort is identified at the point of claim (that is, at the peak of the members' resource utilization cycle). As a result of this criticism, the pre-post cohort method has tended to be replaced in industry studies by a population methodology in which newly identified members who meet the conditions for chronic identification are allowed to enter the measured population. This is the base-case method used for this study. We wished to test, however, how the results of a cohort study would have differed for Highmark

In Table 2, the identification of "chronic" patients is exactly the same as the base case in the baseline period. However, beyond the baseline year, no new entrants to the cohort are allowed. By intervention year 2, the number of chronic members falls by 43 percent relative to the base case, because new

entrants are not allowed. Thus, we follow a closed cohort forward in year 1 and year 2. We use the same trend assumptions as in the base case, because the index population excludes members who are chronic (and not part of the closed cohort), thus representing a truly non-chronic population.

Table 2: Alternative 1

Measure	Baseline 8/00 – 7/01	Intervention Year 1 10/01 – 9/02	Intervention Year 2 10/02 – 9/03
Ave. no. Chronic Measured Members (Base Case)	33,628	44,251	50,739
Ave. no. Chronic Measured Members	33,628	34,957	29,252
Chronic Measured Prevalence	21.3%	19.4%	15.7%
Trend (PMPM, Allowed Cost)			
Chronic Measured Population		0.9%	6.7%
Index Measured Population		9.7%	9.9%
Claims PMPM, Chronic Measured Population			
Projected	$448.26	$491.88	$540.55
Actual	$448.26	$452.29	$482.62
Total Cost Savings, PMPM		$39.59	$57.93
Total Cost Savings, PMPM (Base Case)		$41.54	$65.28
Total Savings ($ millions)		$16.6	$20.3
Savings as % of total claims for the line-of-business		1.5%	1.5%

Note that the average number of chronic measured members in the cohort is lower in intervention years 1 and 2 than that in the base case. The average number of members in intervention year 1 is slightly higher than that in the baseline year, which appears to be counter-intuitive in a cohort. However, it should be remembered that members included in the baseline year may not contribute a full year to that year's exposure (because they become eligible for measurement during the year) whereas they will contribute a full year in intervention year 1 (unless the member terminates).

Our hypothesis for this alternative was that the cohort methodology would produce higher savings than the base case. However, the savings in this alternative are lower, both on a PMPM and total basis (by 4.7 percent in intervention year 1, and 11.3 percent in year 2). This outcome could be the result of two possible effects: our identification algorithm for the chronic population, particularly the three month claims-free requirement for newly identified members, is effective at minimizing regression to the mean, while the effect of including newly identified members in the chronic measured population effectively

creates some bias because these members tend to be lower-cost than the rest of the cohort. We did not have an opportunity to explore these effects in this study, but recommend this as part of any follow-up work.

Alternative 2: Chronic Identification Criteria

a. Based on medical claims but *not* prescription drug claims;
b. Based on primary diagnosis only, as identified from claims;
c. Based only on claims submitted by hospitals (for inpatient or outpatient services).

Table 3: Alternative 2a—Chronic Patients Identified Using Only Medical Claims

In Table 3, only medical claims (no pharmacy claims) are used to identify chronic patients. In this alternative, the number of identified chronic members is lower in the baseline year by 6.1 percent. However, the number of identified members becomes closer to the original base-case number over time, as other identification criteria identify chronic members.

Measure	Baseline 8/00 – 7/01	Intervention Year 1 10/01 – 9/02	Intervention Year 2 10/02 – 9/03
Ave. no. Chronic Measured Members (Base Case)	33,628	44,251	50,739
Ave. no. Chronic Measured Members	31,586	42,724	49,968
Chronic Measured Prevalence	20.0%	23.7%	26.8%
Trend (PMPM, Allowed Cost)			
Chronic Measured Population		(1.6%)	4.0%
Index Measured Population		9.0%	8.9%
Index Measured Population	$176.92	$192.92	$210.03
Claims PMPM, Chronic Measured Population			
Projected	$471.19	$513.80	$559.36
Actual	$471.19	$463.84	$482.20
Total Cost Savings, PMPM		$49.96	$77.16
Total Cost Savings, PMPM (Base Case)		$41.54	$65.28
Total Savings ($ millions)		$25.6	$46.3
Savings as % of total claims for the line-of-business		2.3%	3.5%

Several things are noteworthy in this alternative: the average cost per chronic measured member is higher than in the base case, which is to be expected, given the use of hospital and medical claims for identification. The index measured population cost is higher in the baseline year (reflecting the fact that some patients with chronic claims, who would qualify as chronic under a different set of rules, remain in the non-chronic population. The non-chronic population trends at a slightly lower rate than in the base case. The higher chronic measured population cost, however, results in higher savings per chronic measured member and higher overall savings (the higher PMPM savings more than offsets the lower number of chronic members).

Table 4: Alternative 2b—The Effect of Using Only the Primary Diagnosis on Medical Claims to Identify Chronic Members

Using only primary diagnosis finds fewer, but sicker chronic members. The number of chronic measured members is lower but their cost PMPM is higher than in other scenarios. The number of identified chronic members in Table 4 is lower than either the base case or alternative 2a (13 percent fewer than the number of chronic members in the baseline period in the base case). Once again, the number of identified chronic members becomes closer to the base case over time.

Measure	Baseline 8/00 – 7/01	Intervention Year 1 10/01 – 9/02	Intervention Year 2 10/02 – 9/03
Ave. no. Chronic Measured Members (Base Case)	33,628	44,251	50,739
Ave. no. Chronic Measured Members	29,190	39,526	49,344
Chronic Measured Prevalence	18.5%	21.9%	26.5%
Trend (PMPM, Allowed Cost)			
Chronic Measured Population		(1.7%)	1.4%
Index Measured Population		9.1%	7.6%
Claims PMPM, Index Measured Population	$179.81	$196.17	$211.00
Claims PMPM, Chronic Measured Population			
Projected	$484.58	$528.66	$568.63
Actual	$484.58	$476.44	$483.31
Total Cost Savings, PMPM		$52.22	$85.32
Total Cost Savings, PMPM (Base Case)		$41.54	$65.28
Total Savings ($ millions)		$24.8	$50.5
Savings as % of total claims for the Line-of-business		2.2%	3.8%

Savings per chronic member per month are 26 percent higher in year 1 and 31 percent higher in year 2, compared to the base-case savings. Despite the smaller number of chronic members, the higher PMPM savings result in total savings that are higher than the base case, by 12.2 percent and 27.2 percent in year 1 and year 2, respectively.

Table 5: Alternative 2c—Identifying Chronic Members Using Hospital Claims Only

If enrollees are identified as "chronic" from diagnoses obtained only from hospital-based claims, the number of chronic members is lower than in any other scenario (15 percent lower than in the base case). Again, the number of identified chronic members converges to those in the base case over time.

Measure	Baseline 8/00 – 7/01	Intervention Year 1 10/01 – 9/02	Intervention Year 2 10/02 – 9/03
Ave. no. Chronic Measured Members (Base Case)	33,628	44,251	50,739
Ave. no. Chronic Measured Members	28,710	40,902	49,402
Chronic Measured Prevalence	18.2%	22.7%	26.5%
Trend (PMPM, Allowed Cost)			
Chronic Measured Population		(2.0%)	6.3%
Index Measured Population		7.6%	6.8%
Claims PMPM, Index Measured Population	$186.75	$200.85	$214.59
Claims PMPM, Chronic Measured Population			
Projected	$463.32	$498.32	$532.39
Actual	$463.32	$454.18	$474.72
Total Cost Savings, PMPM		$44.14	$57.67
Total Cost Savings, PMPM (Base Case)		$41.54	$65.28
Total Savings ($ millions)		$21.7	$34.2
Savings as % of total claims for the line-of-business		2.0%	2.6%

The inclusion of members in the non-chronic population who otherwise would qualify as chronic has the effect of raising the initial PMPM claims of the non-chronic group, while at the same time reducing this population's trend. Because of the reduced non-chronic trend, the savings per chronic member per month are 6 percent higher in year 1, but 12 percent lower in year 2, relative to the base case. Total savings in intervention year 1 are close to those of the base case, but year 2 savings are considerably

lower (-14 percent) due to the combination of lower PMPM savings and lower measured member months. The point at which a chronic member is identified and re-classified from the index to the chronic population appears to affect the measured trend for both the index and chronic groups, and thereby to affect the savings calculation. This analysis suggests that a commercial purchaser of DM services comparing results of different programs needs to know not only how the chronic population is identified, but also what has been done with the claims of the "suspect" chronic population—those that would qualify as chronic under different criteria.

Table 6: Alternative 3—Retrospective Chronic Identification

In the actuarial method, as used in the base case, a single trend factor is calculated from the index population for year 1 and year 2 and applied to estimate base-case cost savings. Alternative 3 (Table 6) applies a different chronic identification algorithm, in which members are classified as chronic at the beginning of the baseline period, irrespective of when the chronic member first meets the identification criteria. Similarly, the index population is identified as the complement of the chronic population. This approach to identification results in a non-chronic population trend that is not affected by the migration that potentially affects the base-case trend calculation. This approach results in average claims PMPM that are lower for both the chronic and the non-chronic populations. Consistent with the lack of migration from non-chronic to chronic population over time, non-chronic trend is higher than in the base case. Chronic trend is also higher because newly identified chronic members are not added to the population over time.

Measure	Baseline 8/00 – 7/01	Intervention Year 1 10/01 – 9/02	Intervention Year 2 10/02 – 9/03
Ave. no. Chronic Measured Members (Base Case)	33,628	44,251	50,739
Ave. no. Chronic Measured Members	50,699	54,278	54,575
Chronic measured prevalence	32.1%	30.1%	29.3%
Trend (PMPM, Allowed Cost)			
Chronic Measured Population		11.9%	11.7%
Index Measured Population		11.7%	12.5%
Claims PMPM, Index Measured Population	$158.58	$177.17	$199.37
Claims PMPM, Chronic Measured Population			
Projected	$375.92	$420.01	$472.63
Actual	$375.92	$420.48	$469.62
Total Cost Savings, PMPM		($0.47)	$3.01
Total Cost Savings, PMPM (Base Case)		$41.54	$65.28
Total Savings ($ millions)		($0.3)	$2.0

In the case of the retrospective chronic identification method, we first note that the number of chronic measured members is relatively constant over the three years of the study, rising slightly from the baseline period to year 1, and remaining relatively flat in year 2, as members who are identified as chronic in subsequent periods are assigned to chronic status in earlier periods. The number of chronic measured members in the final year (intervention year 2) is higher than in the corresponding base-case scenario because members who are identified in later periods are not required to meet the same timing requirements for measurement in this methodology, and are thus counted as chronic for the entire period for which they are eligible. The chronic measured population numbers, which are higher than in the base case in the last year, may appear anomalous. In the base case, newly identified members are counted as chronic only after they satisfy all claims-based criteria, and only after a three-month waiting period. On average, therefore, assuming continuous identification, newly identified members contribute slightly less than one-half year to total chronic member years in the year of first qualification. Under the retroactive identification method, these members contribute immediately on identification, and contribute 12 months for each measurement year for all years in which they were a plan member.

The average PMPM cost for the chronic members is lower than that of chronic measured members in the base case (by 16 percent), which is consistent with our expectations given the number of non-chronic members assigned to this group in the early years. The index population trend is higher than the base case, (11.7 percent vs. 9.7 percent in year 1 and 12.5 percent vs. 9.9 percent in year 2). In the base case, the chronic measured group experienced very little trend (0.5 percent in year 1 and 5.5 percent in year 2). In the retrospective chronic identification method, the chronic measured population experiences a very similar trend to the index population. The lack of significant differential in the chronic and non-chronic trends results in small savings with this method (in year 1, savings are actually negative, although this result is not statistically significant).

Table 7: Alternative 4—Continuous Eligibility Criteria

In the actuarial method, as used in the base-case, members were included in the study at the later of their attainment of six months of continuous eligibility or the beginning of the baseline period. In order to test the sensitivity of the results to the continuous eligibility criterion, we analyzed the results with no continuous eligibility requirement. In order to make this alternative operational, we also removed the requirement that members be at least three months post-chronic identification before they are included in the measurement population. Results were as follows:

Measure	Baseline 8/00 – 7/01	Intervention Year 1 10/01 – 9/02	Intervention Year 2 10/02 – 9/03
Ave. no. Chronic Measured Members (Base Case)	33,628	44,251	50,739
Ave. no. Chronic Measured Members	39,811	50,394	56,063
Chronic Measured Prevalence	25.2%	28.0%	30.1%
Trend (PMPM, Allowed Cost)			
Chronic Measured Population			
Index Measured Population		9.8%	10.7%
Claims PMPM, Index Measured Population	$169.99	$186.66	$206.77
Claims PMPM, Chronic Measured Population			
Projected	$548.59	$603.05	$667.33
Actual	$548.59	$544.67	$561.50
Total Cost Savings, PMPM		$58.36	$105.83
Total Cost Savings, PMPM (Base Case)		$41.54	$65.28
Total Savings ($ millions)		$35.3	$71.2
Savings as % of total claims for the line-of-business		3.2%	5.5%

In this alternative, a number of members who are clearly high cost (primarily those initially identified through a hospital claim, for example) are now included in the measurement population for the study. This raises the average PMPM cost from $448.26 to $548.59 (a 22 percent increase). This increase in the base cost, coupled with a slightly higher trend in the index population, increases the estimated savings PMPM by 41 percent (first year) and 62 percent (second year). The change also increases the number of measured members, and thus the total savings. One of the objections often raised by purchasers to the base-case methodology is that the use of claims truncation increases the savings. Our analysis suggests that the opposite may be true. The inclusion of large claims increases the base PMPM and therefore the overall savings.

Table 8: Alternative 5. Results in Commercial Products

This study has focused on Highmark's Medicare population. However, the program was also implemented within the Commercial (HMO/POS and PPO) populations. We also performed an assessment of the DM program in the Commercial HMO/POS population; we did not pursue the PPO because the product was in transition and the program was judged too immature to permit the DM program results to be analyzed.

Commercial HMO/POS Savings Calculation

The initial chronic measured prevalence is 3.7 percent, indicating that initially relatively few members were identified with one or more chronic conditions, and met all the other requirements for inclusion in the study. The chronic prevalence grew by 51 percent over the three-year study period, to 5.6 percent.

Measure	Baseline 8/00 – 7/01	Intervention Year 1 10/01 – 9/02	Intervention Year 2 10/02 – 9/03
Ave no. Members	1,030,204	1,107,120	1,027,539
Ave. no. Chronic Measured Members	38,126	53,799	57,444
Chronic Measured Prevalence	3.7%	4.9%	5.6%
Trend (PMPM, Allowed Cost)			
Chronic Measured Population		5.5%	8.4%
Index Measured Population		20.3%	12.5%
Claims PMPM, Index Measured Population	$56.89	$68.43	$77.00
Claims PMPM, Chronic Measured Population			
Projected	$237.94	$286.19	$322.03
Actual	$237.94	$251.07	$272.15
Total Cost Savings, PMPM		$35.12	$49.88
Total Savings ($ millions)		$22.7	$34.4
Savings as % of total claims for the line-of-business		1.1%	1.6%

Estimated savings for the HMO/POS population are similar in magnitude to those for the Medicare population, both on a PMPM and total basis. However, the mechanism by which the savings are derived differs significantly: the Medicare population consists of far more chronic members with a cost PMPM that is approximately twice that of the HMO/POS population. Conversely, the trend assumption used to project the baseline cost to the intervention period is higher in the HMO/POS population than the

Medicare population.[4] At the line-of-business level, the lower prevalence and costs result in savings that are lower than in the Medicare population (although still significant).

Discussion

At present, there is no consensus among actuaries, health economists, accountants or health services researchers on how DM program cost savings estimates *should* be made. Our results demonstrate that, even within the same general methodology, the estimates of cost savings can vary substantially, depending on how chronic patients are identified, how patients are included based on continuous enrollment criteria, the treatment of non-chronic members who are "suspect" chronic patients and who may later be added to the chronic population and how trend is handled. The following table summarizes all comparative PMPM savings results:

Savings PMPM under Different Scenarios

Table Number	Scenario Number	Scenario	Intervention Year 1 10/01 – 9/02	% change compared with base case	Intervention Year 2 10/02 – 9/03	% change compared with base case
1.	0.	Base-case	$41.54	-	$65.28	-
2.	1.	Cohort	$39.59	(4.7%)	$57.93	(11.3%)
3.	2a.	Medical claims only identification	$49.96	20.3%	$77.16	18.2%
4.	2b.	Primary diagnosis only identification	$52.22	25.7%	$85.32	30.7%
5.	2c.	Hospital claims only identification	$44.14	6.3%	$57.67	(11.7%)
6.	3.	Retrospective identification	($0.47)	(100.0%)	$3.01	(95.4%)
7.	4.	Continuous eligibility criteria	$58.36	40.5%	$105.83	62.1%
8.	5.	Commercial HMO product	$35.12	n/a	$49.88	n/a

[4] The trend assumptions used in the baseline calculations, although they appear high in the first intervention year, have been reconciled with Highmark's overall population trends for the relevant periods.

In summary, how and when chronic members are identified can significantly impact the savings. Many of the methodology choices and assumptions are subtle and difficult to identify in published studies and vendor reporting. It is thus critical for DM companies and purchasers to increase their disclosure in this area. Our results show that in a large population study, whether results are estimated using a cohort or population methodology has relatively little effect on savings. This result appears to contradict some of the current thinking in the industry, which has largely moved away from cohort studies. However, our result may be influenced by the fact that we used the initial chronic cohort, which was identified from a mix of hospital, drug and medical claims. Our identification criteria find members in all stages of disease, while cohort studies that identify high-risk, hospitalized (or recently hospitalized) members are more likely to be subject to the effect of regression to the mean at the individual member level. The retrospective identification method generates trends that are similar in the chronic and non-chronic populations, and therefore appears to generate little or no savings. This method, while it has a certain intuitive and statistical appeal, runs counter to clinical perceptions of disease (members are classified as chronic, even though they have no chronic condition and are not managed as such). As Table 6 shows, the PMPM costs of both the chronic and the non-chronic populations are lower, initially, than in any other methodology, reflecting the mix of chronic and "future chronic" members assigned to the chronic population. Whether this is a true claim cost PMPM or whether some adjustment should be made to it requires more analysis. Much more work needs to be done in the area of understanding and applying appropriate trends at the population level before these results can be accepted into the mainstream.

Finally, the Commercial HMO/POS results are presented for completeness, even though they are not truly comparable with the Medicare results. Purchasers of DM programs will be encouraged that the Commercial results are as close to those of the Medicare population as they are. Given the similarity of the underlying results to the Medicare base case, we expect that similar sensitivity will be found in the Commercial population to that observed in the Medicare population.

Conclusions

The calculation of cost savings of DM programs by actuaries is still an inexact science. This study demonstrates that estimates of cost savings can vary substantially, depending on the methods used to identify and exclude patients. Our research in these papers, however, supports the rigorous, actuarial approach to the analysis, for example:

- Rigorous identification of members who meet defined criteria and their inclusion in the study population at appropriate times;
- Control of study eligibility over time, linkage of members and their claims; and

- Careful development of an appropriate adjustment trend assumption.

Purchasers of programs should take equal care about the assumptions and data decisions that are taken "behind the scenes" in the calculation of program savings. Purchasers will increasingly look to their trusted health care financial advisors, actuaries, for guidance on assumptions, methodology, calculations and to benchmark results. As this paper shows, the results of a particular program can vary widely, based on these assumptions. Purchasers may wish to become involved more directly with the setting of assumptions in studies, and be willing to accept a range of outcomes, rather than a 'point' estimate of savings.

At the same time, actuaries should avoid the simplistic conclusion of dismissing DM savings calculations as unreliable because of "regression to the mean." As our analysis throughout this series of papers has shown, regression to the mean, which may occur at the individual member level, may not be observed in a large population with offsetting changes in member utilization. There are a number of potential biases in a population study of DM outcomes; regression to the mean may be one of them, depending on how the population is selected, but it may not be the greatest source of distortion.
For the actuary who is approaching a new study or the review of a vendor's results, we provide a check-list of issues/assumptions that should be considered in Appendix 3.

Our review of the literature of DM financial outcomes studies shows that there are very few published, peer-reviewed studies of populations. Much more data analysis and publication is needed for us to understand and to begin to develop the necessary tools (for example, risk-adjustment; durational adjustment, etc.) and data that may assist us to correct for known biases.

More research of evaluation methods is also needed, which will require funding from DM companies, managed care organizations and the federal government. Clinical researchers and epidemiologists have striven to improve methods to conduct clinical trials and have considered and critiqued alternative research designs to evaluate new drugs and medical and surgical procedures. Health care actuaries have an opportunity to make a substantial contribution to an exciting, evolving field of academic and practical importance. Further efforts in this field are warranted to validate financial models proposed by managed care organizations or DM companies. Purchasers and actuaries who "validate savings" should be aware of the implications of different approaches that are commonly used.

REFERENCES

American Healthways, Inc., and Johns Hopkins University. 2003. Standard outcomes metrics and evaluation methodology for disease management programs. *Disease Management* 6 (3): 121-138.

Fetterolf, D., Wennberg, D., and DeVries, A. 2004. Estimating return on investment in disease management programs using the pre-post analysis. *Disease Management* 7 (1), Winter.

Fitzner, K., et al. 2004 Principles for assessing disease management outcomes. *Disease Management* 7(3), October.

————. 2004. *Disease Management Program Evaluation Guide.* Washington, D.C.: Disease Management Association of America.

Linden, A., Adams, J.L., and Roberts, N. 2003. An assessment of the total population approach for evaluating disease management program effectiveness. *Disease Management* 6 (2), Summer.

APPENDIX 1: Chronic identification criteria

For the DM program, chronic members are identified by Highmark using the criteria below. In addition, some members may be referred by physicians or health plan staff. These members are included in the management program, but excluded from the evaluation (both the intervention and reference (index) groups). Newly identified members were included in the chronic measurement population after a three-month waiting period (to allow for any effect of regression to the mean). During this period, these members were excluded from both the chronic and index measured populations.

Diabetes Mellitus

At least one admission or ER visit with a primary diagnosis of diabetes (ICD-9 codes 250, 357.2, 362.0, 366.41, 648.0);
OR two professional visits in a 12-month period with a primary or secondary diagnosis of diabetes; (in range of 99 series E & M codes) and (92 series for eye visits);
OR one or more dispensed insulin, hypoglycemic or anti-hyperglycemic (therapeutic class 172, 173, or 174).
EXCLUDE 648.8x gestational diabetes.

Chronic Obstructive Pulmonary Disease (COPD)

At least one admission or ER visit with a primary diagnosis of COPD (491.xx, 492.xx, 494.xx, 496.xx) or primary diagnosis =466.xx with secondary diagnosis = (491.xx or 492.xx or 496.xx);

OR at least four primary diagnoses or secondary diagnoses (four encounters in range of 99 series E & M codes);

OR two primary diagnoses or secondary diagnoses (two encounters) for COPD (in range of 99 series E & M codes) AND 2 medication dispensing events for beta agonists, cortico-steroids, atrovent, serevent, theophylline or O_2.

Congestive Heart Failure (CHF)

At least one admission with primary diagnosis of CHF (402.01, 402.11, 402.91, 404.01, 404.11, 404.91, 428.xx);
OR three or more physician encounters (in range of 99 series E & M codes) with a dx1 or dx2 for CHF (in 12-month time frame).

Coronary Artery Disease (CAD)

At least one admission with ICD-9 procedure code for CABG (36.1x, 36.2x); or ICD-9 procedure code for PTCA (36.01, 36.02, 36.05, 36.09) or any primary diagnosis of acute coronary ischemia (410.xx-414.xx).
OR four or more physician encounters (in range of 99 series E & M codes) with Dx1 for acute coronary ischemia.

Asthma

At least one ER visit primary diagnosis with ICD-9 code =493.x, OR at least one inpatient discharge with code 493.x;

OR at least four outpatient visits primary or secondary diagnosis ICD-9 = 493.x (in range of 99 series E & M codes);

OR two outpatient visits primary or secondary diagnosis ICD-9 = 493.x (in range of 99 series E & M codes) AND at least two asthma medication dispensing events.

APPENDIX 2: Claims Exclusion Criteria

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 ICD-9 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 & Other External Causes	905 – 909
Burns	940 – 949
Poisoning by Drugs, Medicinal & 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 & Medical Care NEC	995 – 999

2. Psych/Substance Abuse

   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 pick up members who have a prior history of cancer but who are now in remission. These members often represent appropriate candidates for chronic DM 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: subject to change in benefits design, provider, etc. on a more frequent basis than hospital or physician coverage.

APPENDIX 3: Checklist of Issues/Assumptions for Conducting or Evaluating a Study

No.	Issue	Discussion	Comments
	STUDY DESIGN		
1.	Study design	What is the specific methodology used for the study? (For evaluation of study designs see Paper 5.)	
2	Study design—Population study	Does the study use a population methodology? (Population studies are preferred to cohort studies in order to minimize regression to the mean.) It may not always be possible to conduct a valid population study—for example for smaller employer groups. How is this addressed?	
3.	Study design—Reference Population	A reference population is defined to compare to the intervention population (e.g., historical; geographical, product, etc.)	
4.	Study design—method for overcoming small populations	For example, is there some form of credibility weighting employed in adjusting the results?	
5.	Study design—intervention time period	What is the study time-period (preferably 12 months post-beginning of intervention program). Where program is not in place for 12 months, how is this to be handled? Where launch is "staggered" or members are continuously invited, how is this handled?	
6.	Study design—baseline time period	What is the baseline time period, when the study is a historical adjusted methodology?	
	POPULATION DEFINITIONS		
7.	Population definitions—chronic population	Objective claims criteria used for identifying eligible chronic population. Is methodology hospital claims only; hospital + medical; hospital + medical + prescription drugs? How will issue of false positive be handled as wider criteria are employed?	
8.	Population definitions—excluded	Objective claims criteria used for identifying excluded population (e.g., AIDS, transplants, etc.)	

	population		
No.	**Issue**	**Discussion**	**Comments**
9.	Population definitions—excluded population	Are there any excluded members who are excluded based on subjective criteria (for example members in case management)? How will these members be evaluated?	
10.	Identification criteria—multiple years	Does the study require that members be re-qualified through claims each year in the study or do members accumulate in their groups as long as they are benefits-eligible?	
11.	Population definitions—newly-identified chronic and excluded population	How are the newly-identified members of the population handled? Are they included in the study immediately on identification? Are they included at the next anniversary? Is the treatment symmetrical in all years?	
12.	Eligibility criteria	Is there a requirement for continuous eligibility in order to be included in the study?	
13.	Eligibility criteria	How are terminating members (from health plan) handled? How are terminating members (from program) handled? Do they contribute to the study up to month of termination? Are they retroactively removed?	
	DATA		
14.	Data exclusions in dataset provided	Are there any data that will not be provided for analysis (for example certain states have privacy restrictions that pre-empt HIPAA)?	
15.	Data specifications	What dataset will be used for the study?	
16.	Data validation/ Reconciliation	What data validation process has been (will be) performed in order to ensure that data are complete, reliable and balance back to audited financial statements of the client?	
17.	Data exclusions in study	Define any excluded conditions (for example, maternity or cancer claims, where these are not the responsibility of the DM vendor).	

No.	Issue	Discussion	Comments
18.	Claims run-out	The same run-out months are included in each year of data used in the study (i.e., additional run-out does not continue to accumulate on earlier years).	
	POTENTIAL SOURCES OF BIAS		
18.	Prevalence creep	Are members re-qualified over time?	
19.	Trend bias	How is the potential bias in trend due to the effect of migration between non-chronic and chronic groups allowed for?	
20.	Geographic and product controls	Are the starting claims costs adjusted to place all members on the same actuarial basis? Is the control group sufficiently representative that its trend is a valid proxy for the intervention group?	
21.	Selection bias	How were members included in the study? Is the effect of selection bias overcome by performing a population study or is the study limited to enrolled members only?	
	TESTS OF EQUIVALENCE		
22.	Test the intervention and control populations for equivalence	Test the intervention and control populations (baseline and measurement populations, for example) for equivalence with respect to risk factors: - Diagnosis-related group (DRG) distributions; - Provider distributions; - Age/sex distributions; - In-/out-of-network services.	
	TREND		
23.	Method for calculating trend	Where there is a trend adjustment, what trend is used for a proxy? How is this adjuster calculated? How representative of the non-managed chronic group expected trend is the adjuster? How stable is this measure over time? Has the measure been tested *prior* to the beginning of the program in order to evaluate its suitability as a proxy?	

No.	Issue	Discussion	Comments
24.	Validate the calculated trend	Is the vendor's choice of a trend assumption reasonable in relation to the plan's recent trend experience?	
	REPORTING		
25.	Reports are auditable	Reports are provided that support simple checking of the calculation at the aggregate level, while also supporting drill-down audits of selected cells and calculation components.	
	CALCULATIONS		
26.	Audit the calculated savings numbers	Have the savings numbers been calculated (mathematically) correctly? Are the numbers reasonable based on other experience of the actuary? Of the industry? Is there independent support for the savings (e.g., observed trend moderation)?	
27.	Audit the components of the calculated savings numbers	Are the components that make up the calculation reasonable (for example, the PMPM numbers for services categories; utilization numbers; unit cost numbers)?	
28.	Are the calculated savings plausible?	Validate the savings by decomposing into utilization reductions. Do the implied utilization reductions seem reasonable? A reduction of $100 PMPM in a chronic population translates into a reduction of X in inpatient admissions. Is there support for this level of reduction in inpatient admissions?	