Measuring DM Program Outcomes: The Use of Incident and Prevalent Populations

Ying Liu, M.S.
Michael Cousins, Ph.D.
Health Management Corporation

Abstract

The population-based pre- versus post-period design, also known as the adjusted historical control group design, is a popular method to evaluate disease management (DM) programs. This method combines results for members who are newly identified, the "incident population," with results for members who are identified before the evaluation periods, the "prevalent population."

Recently, there has been interest in two issues related to the incident population. First, there is interest in separating the incident members from the prevalent members. The primary rationale for this is to minimize a source of regression to the mean (RTM), and hence to minimize systematic bias. Second, there is interest in creating a cohort of members from the incident population and using them to make an RTM adjustment.

In this white paper, we describe the rationale and need for these approaches and show the results of our investigation into this issue using administrative claims data from two health plans. The conclusions from our analyses are that:

1. Separating the incident and prevalent populations is useful, but

2. The use of a cohort to make an RTM adjustment is unwarranted in the population-based pre- versus post-period design.

1. Definition of the Prevalent and Incident Populations

The incident population represents those members identified *during* an evaluation period. For example, incident members may be identified either during a 12-month pre-program (baseline) or during a 12-month first-year program (year 1) period. These members can be referred to as "Pre-Program (or Baseline) Incident" and "Year 1 Incident," respectively. In general, the incident population represents those members who are newly identified. Specifically, incident members fall into two groups:

1. those who are diagnosed with chronic conditions during the period, or

2. those who joined the health plan during the period with existing chronic conditions prior to enrollment.

The members in the first group are truly newly identified and therefore meet the classical epidemiological definition of incidence. The members in the second group, although newly identified by the operational definitions of the disease by the DM program (e.g., member has two ICD-9 codes for diabetes), meet the classical epidemiological definition of *prevalence*. However, the absence of information about the member prior to joining the health plan means that each member enters with a "clean slate"—even if the member has had the condition for several years.

In contrast to the incident population, the prevalent population includes those members identified *prior to* an evaluation period. For example, prevalent members may be identified either during a 12-month period before the start of the pre-program year or during a 12-month period before the start of the first year program period. In these cases, these members can be referred to as "Pre-Program (or Baseline) Prevalent" and "Year 1 Prevalent", respectively. Loosely speaking, the prevalent population is not newly identified; these are members who met the criteria for a disease prior to the start of each measurement period.

2. Characteristics of the Incident and Prevalent Populations

To better understand the effect of a DM program on the incident and prevalent populations, we examined three incident and three prevalent populations across three years from a large health plan.

We were first interested in understanding how the health care claims costs for the incident and prevalent populations changed over time. Figure 1 shows *trend-adjusted* costs for three incident populations from the same health plan. Moving from *right to left*, the first incident population are those members identified during the time that corresponds to the DM program; moving to the left, the second incident population are those members identified during the time that corresponds to the baseline (labeled as pre-program); moving to the left, the third incident population are those members identified during the time that corresponds to the period *before* the baseline (labeled as pre-pre program).

Figure 1 shows that claims costs for the incident population actually increase before the start of the DM program before decreasing, as expected, during the program year. Therefore, there is no evidence of RTM since claims costs do not decrease until the program year. This same pattern of results was also observed when claims data from a second large health plan were analyzed.

Figure 1: Trended Costs for Incident Populations

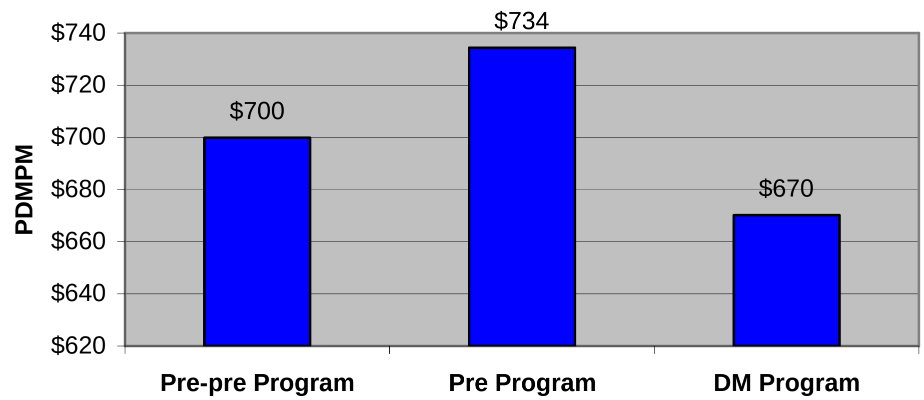
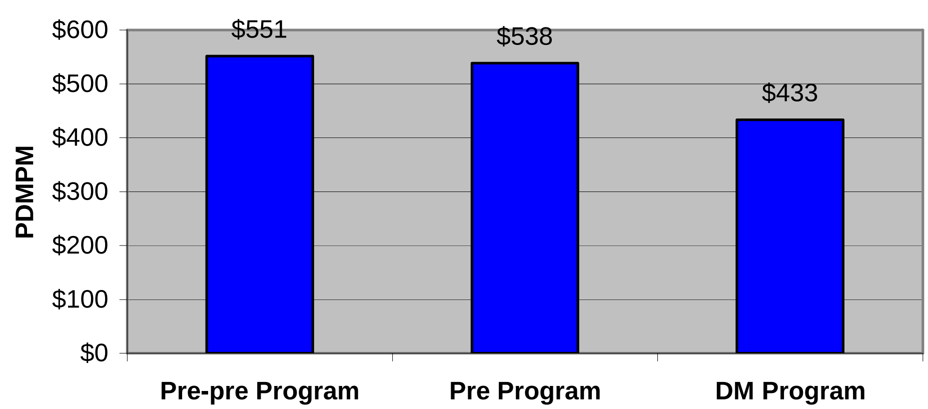


Figure 2 shows *trend-adjusted* costs for three prevalent populations from the same health plan. Moving from right to left, the first prevalent population are those members identified during the 12 months *prior to the start* of the DM program; to the left, the second prevalent population are those members identified prior to the start of the baseline (labeled as pre-program); to the left, the third prevalent population are those members identified prior to the start of the period that corresponds to the period *before* the baseline (labeled as pre-pre program).

Similar to what we observed with the incident populations, Figure 2 shows that claims costs for the prevalent population remain stable before the start of the DM program. Therefore, there is no strong evidence of RTM since claims costs do not significantly decrease until the program year. This same pattern of results was also observed when claims data from a second large health plan were analyzed.

Figure 2: Trended Costs for Prevalent Populations

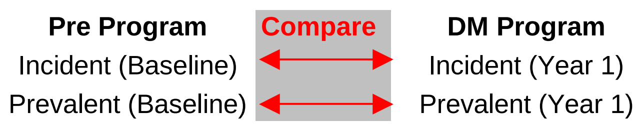


3. Issues with the Incident and Prevalent Populations

3.1 Regression to the Mean

As described above, the incident and prevalent populations represent different groups of members; therefore it is important that population-based pre- versus post-period comparisons are performed *separately* for each of these groups. For example, when comparing costs between two periods (e.g., a pre-program period and program period), it is important to compare the two incident populations separately from comparisons between the two prevalent populations. This is shown graphically below.

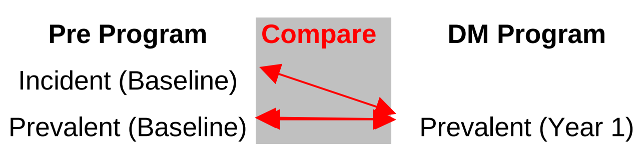
Figure 3
Population Comparisons of Population-Based Pre- v. Post-Period Design



The population-based analysis compares Pre-Program (Baseline) Incident members with the Year 1 Incident members. Similarly, the population-based analysis compares Pre-Program (Baseline) Prevalent members with Year 1 Prevalent members. These parallel comparisons are critical to minimize systematic bias that is inherent to comparisons between cohorts.

Comparisons of costs over time for a cohort have been shown to be associated with RTM. For example, this can occur if one identifies incident members during the baseline (Pre-Program Incident members) and tracks the costs for this cohort group over time (see Figure 4). Cost comparisons between the pre-program and program periods for this incident cohort group will be greatly affected by RTM simply because the Pre-Program Incident members move to the Year 1 Prevalent.

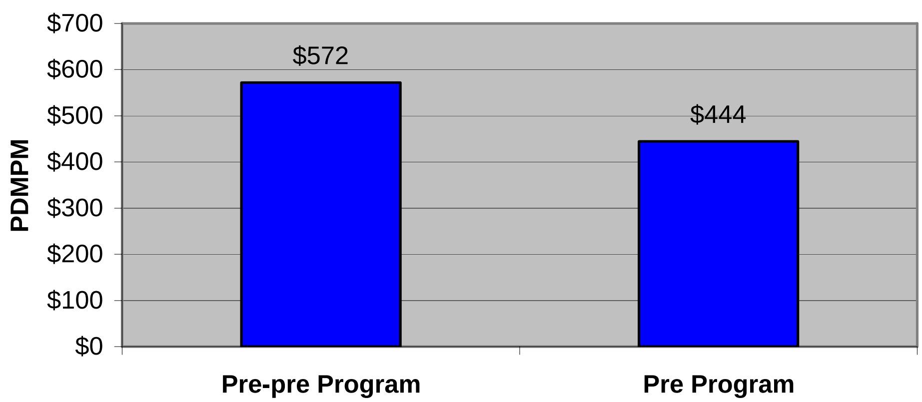
Figure 4
Population Comparisons of Cohort-Based Pre- v. Post-Period Design



Data supports the view that tracking costs for an incident population result in substantial cost reductions that are due in part to RTM. Figure 5 shows the costs of Pre-pre Program Incident members (i.e., these members were identified during the Pre-pre program) during the Pre-pre Program and Pre-Program periods. Note that costs for this cohort decrease over time. This same pattern was found by analyzing claims data from a second large health plan.

The comparison of costs for the incident cohort (Figure 5) is in marked contrast to what is observed when costs for two *separate* incident populations are compared (i.e., Figure 1). There is no evidence for RTM when comparing two separate incident populations.

Figure 5: Evidence for RTM with Incident Cohort



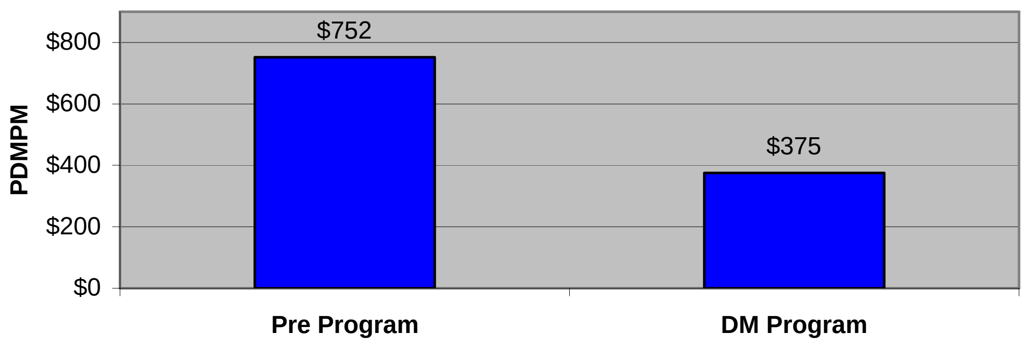
Cost comparisons between cohorts should be avoided if the objective is to measure the effects of a DM program. This is because the cohort, in the example above, the "incident cohort," will pollute the overall results and make interpretation of program impact impossible. This is shown in Figure 6, where "savings" for the Pre-program Incident cohort between the baseline and program year is 50 percent (($752 - $375)/$752). This same pattern was found by analyzing claims data from a second large health plan.

In contrast, savings between the two separate incident populations, such as the Pre-Program Incident and Year 1 Incident populations, is much less (9 percent, ($734-$670)/$734, as in Figure 1)*. This is expected because the two incident populations are

* Note that the costs for the Pre-Program Incident population in Figure 1 are not identical to the costs for the Pre-Program Incident population in Figure 6. This is because the two populations require different criteria; the latter population is a cohort which requires membership eligibility in both the baseline and program period.

mutually exclusive and comparable. If these groups were not mutually exclusive, an RTM adjustment factor might be necessary.

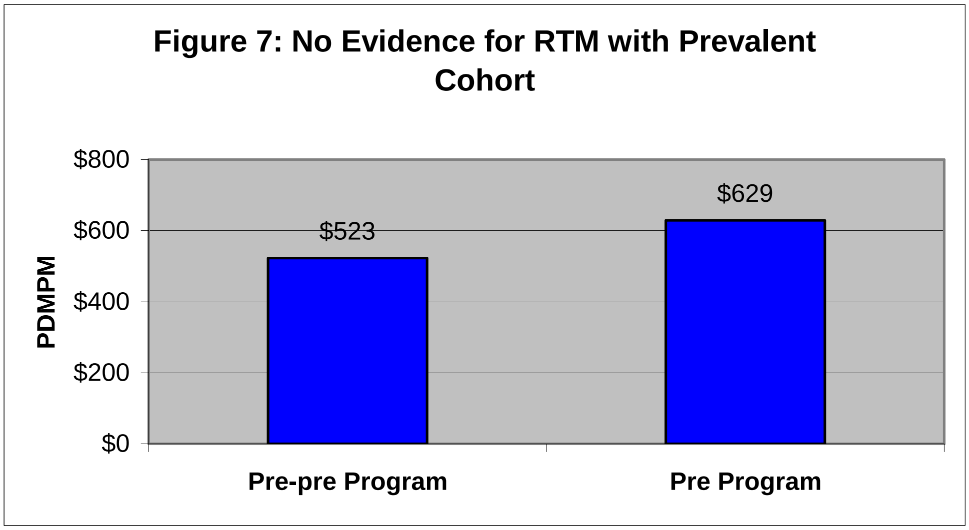
Figure 6: Combined RTM and Program Effect with Incident Cohort



There is no evidence of RTM with prevalent populations, even though they are not completely mutually exclusive (see Figure 2). Although speculative, the reason RTM is not evident at the population level may be because these members have already experienced their sentinel, and often expensive, events. Claims after this incident period are later classified as prevalent, and since the sentinel events are not included, the "prevalent claims" that occur later are more stable and less volatile.

An example of the distinction between sentinel, incident and prevalent claims dollars can be easily observed when considering a coronary artery disease (CAD) program that has heart attacks as identifying events. If a member is entered into a CAD program on the basis of an (expensive) heart attack, by definition the claims related to the treatment of the heart attack are in the calculations for the incident population, not the prevalent population.

Another feature of analysis of the prevalent members is that there does *not* appear to be evidence for RTM even with a prevalent *cohort* (see Figure 7). This may be surprising at first, but it becomes less surprising when considering that this population does not include members with high-cost sentinel events.



3.2 Different Proportions of Incident and Prevalent Members

There are still issues to consider with the incident population in the population-based pre- versus post-period design. Each year the proportion of incident and prevalent members varies, sometimes just slightly. This change can result in the appearance of program savings, even in the absence of a program. For example, if the prevalent population has lower costs than the incident population (this is often the case) and the proportion of the prevalent population increases year after year, the overall PDMPM costs will be lower even in the absence of an intervention.

Fortunately, this systematic bias can easily be overcome. Overcoming this obstacle simply requires calculating the overall savings separately for the incident and prevalent populations and using member months to weight the contribution of each group. These calculations are shown in the next section.

4. Recommendation for Evaluating Program Savings

Our recommendation for calculating program savings is shown below. Based on the above discussion, it is clear that comparisons of cohorts should be avoided and that overall savings should be calculated by examining the independent results from the incident and prevalent populations.

This objective is accomplished by calculating program savings using the weighted average (weighted by member months, MM) of the savings for the incident and prevalent populations to reflect the ratio of the incident population. The formula for this is shown below:

$$\text{Weighted Gross Saving PDMPM} =$$

$$MM_{Incident} / (MM_{Incident} + MM_{Prevalent}) * \text{Gross Savings PDMPM}_{Incident} + MM_{Prevalent} / (MM_{Incident} + MM_{Prevalent}) * \text{Gross Savings PDMPM}_{Prevalent}$$

Note that the PDMPM savings for the incident and prevalent populations are calculated separately (i.e., see Figure 3) and then weighted by member months. This approach eliminates RTM biases that arise from using a cohort (e.g., there is no incident cohort). In addition, this approach eliminates biases due to changes in the proportion of incident and prevalent members. This latter point is illustrated by the calculations shown in Table 1 where the proportion of *less expensive* prevalent members *increases* over time. This scenario does not artificially produce savings; when using the formula shown below, the savings are *zero*.

TABLE 1
Member Month Weighting Adjusts Changes in Incident and Prevalent Proportions

	Pre Program Year	Program Year	Gross Saving PDMPM	
			Formula	Value
Incident				
PDMPM	$100	$100		0=100-100
Enrolled Months	1,000	800	**Gross Saving PDMPM$_{Incident}$ = PDMPM$_{Pre\ program}$ - PDMPM$_{Program}$**	
Weight$_{Incident}$		40%		
Prevalent				
PDMPM	$75	$75		0=75-75
Enrolled Months	1,000	1,200	**Gross Saving PDMPM$_{Prevalent}$ = PDMPM$_{Pre\ program}$ - PDMPM$_{Program}$**	
Weight$_{Prevalent}$		60%		
Weighted Gross Saving PDMPM			***= Weight$_{Incident}$ * Gross Saving PDMPM$_{Incident}$ + Weight$_{Prevalent}$ * Gross Saving PDMPM$_{Prevalent}$***	***0=800/(800+1200)*0+ 1200/(800+1200)*0***

Finally, there are important methodological details that should be considered but are outside the scope of this paper. For example, we recommend that the evaluation only include members enrolled with the health plan for at least 6 months (Cousins, 2003). The interested reader is referred to the presentation by Cousins (2003) and the recent publication by Fetterolf et al. (2004).

5. Conclusions and Recommendations

First, the separate examination of the costs for the incident and prevalent populations is critical to clarify our understanding of the impact of the DM program. Second, when calculating overall program savings, systematic biases related to RTM and different proportions of incident and prevalent members can be eliminated by: a) calculating financial outcomes for the incident and prevalent populations separately; and b) weighting those results by member months. This can be accomplished using the recommended approach and formula shown in this paper, which precludes the need for any adjustment related to RTM.

Finally, it is important to keep in mind that the dissection of incident and prevalent populations described above is only one facet related to evaluating outcomes. Although the methodological design has tremendous influence on the results, there are other important methodological details that need to be considered. For example, minimum enrollment criteria have a large impact on results. In addition, because *any and all* pre- versus post-period comparisons are by definition *non-experimental*, when interpreting and making decisions about financial outcomes it is absolutely critical to measure and consider other outcomes—and for contractual guarantees to extend beyond just financial results. For example, it is important to critically examine the *relationship between* operational, health quality and status outcomes and financial results. Although these relationships will not demonstrate causality—that is, that the DM program is causing the desired improvements—examination of these relationships will allow decision makers to know if the financial results are even *plausible*

References

Cousins, M. 2003. The devil is in the details: Exploring evaluation methodologies for disease management. Presentation at DMAA's Fifth Annual DM Leadership Forum. October.

Duncan, I. An actuarial method for evaluating disease management outcomes. Draft.

Fetterolf, D., Wennberg, D. and Devries, A. 2004. Estimating the return on investment in disease management programs using a pre-post analysis. *Disease Management* 7(1): 5-23.

Fitzner, K. 2005. Principles for robust disease management evaluation. *Disease Management* 8(2): 135-140.

Johnson, A. 2003. Measuring DM's net effect is harder than you might think. *Managed Care* 6:28-32.

Kelleher, M. 2003. Real world ROI and clinical outcomes for diabetes, CHF, CAD and COPD. Presented at Disease Management Summit, May 13, 2003, accessed at http://www.ehcca.com/presentations/dmconference1/5_06.pdf on June 8, 2004.

Lee, D. Evaluating the impact of disease management programs. Available at www.caremark.com.

Lewis, A. 2004. Measuring ROI: This is our final answer. Presentation at 2004 DM Employer Summit. May.

Wilson, T.W. 2003. Evaluating ROI in state disease management programs. *State Coverage Initiatives* Nov. 4(5): 1-6.