RECORD, Volume 31, No. 2*

New Orleans Health/Pension Spring Meeting
June 15–17, 2005

Session 41 Seminar
An Introduction to Care and Disease Management Interventions, Part 2

Track: Health

Moderator: Ian G. Duncan

Panelists: Henry Dove†
Ian G. Duncan

Summary: In 2003, the Health Section of the Society of Actuaries, recognizing the increased need for rigorous calculation of the financial outcomes from care and disease management (DM) programs, sponsored an extensive research project into the actuarial issues of these programs and their financial measurement. The study encompassed both theoretical and practical aspects, including analysis of outcomes from an extensive disease management program that has been in place for a number of years at Highmark, Inc. Part two of the seminar examines some of the key issues and findings from the research, including an actuarial methodology for measuring disease management outcomes and practical applications of DM outcomes measurement within a health plan population.

† Dr. Henry Dove, not a member of the sponsoring organizations, is lecturer at Yale University Health Management Program in Hamden, CT.

Note: The charts referred to in the text can be downloaded at: http://handouts.soa.org/conted/cearchive/neworleans-june05/041_combined.pdf.

MR. IAN G. DUNCAN: I'm going to discuss Papers 6-8 of the Society's research study, "Evaluating the Results of Care Management Interventions: Comparative Analysis of Different Outcomes Measures," plus some conclusions and some pointers to further study. The study is available on the Health Section's Web site: http://www.soa.org/ccm/content/areas-of-practice/health/research/eval-results-care-man-int/. I wish I could say that after two years our work is done, but I'm hopeful that it will be completed by the end of the summer. Paper 6 is entitled "An Actuarial Method for Assessing Disease Management Savings Outcomes." Paper 7 is a comparative analysis of chronic and non--chronic trends. Paper 8 is an application to some actual data.

By way of background and explanation, in Paper 6 I refer to a method that some people call "adjusted historical." I decided for marketing reasons to call this the "actuarial method," because in an industry in which a lot of the work is done by informatics people and clinical people, actuaries don't have much of a toehold yet. It seemed to me that there were good reasons why, if you apply the title "actuarial" to a particular methodology, it might create some opportunities for actuaries. But there are other good reasons why it should be called "actuarial," because it uses adjustments and trend, and if anybody "owns" trend, it's the actuaries. It's clearly an actuarial concept and an opportunity for us to have an opinion on a study, not to mention all the other adjustments like age and sex and geography.

This is the prevalent industry methodology; most large vendors seem to use some variation of this methodology. It is trend-adjusted historical control, and, again, since trend is an actuarial concept, there are opportunities for us. Here's a very simple example of how people apply it in practice. There are many variations on this, but this is the very basic concept. You create a baseline cost per member per year, trend it forward and subtract out the actual cost per member per year. The difference between the trended baseline and the actual cost is called "savings" and attributed to the value of the intervention. Because it's on a per member per year basis, you then multiply this by the number of member years in the population, and you can then calculate or estimate a total estimated savings.

But here's the rub: The whole savings as calculated depends on what number you use as the trend number. If you can get an accurate fix on your trend, you can get an accurate estimate of savings. If you happen to use an aggressive assumption about trend, you might be able to produce some very large savings.

In this paper, we go into some detail about assigning the populations. This is intended not so much for the actuarial audience, but for the general audience to understand the notion that you need to control who is in what population for any period of time, and that when you do these studies, people get assigned to different buckets. For example, if you start with all eligible health plan members, then you have to figure out which of those members is a chronic person. It is difficult to

identify chronic people. There is no universally accepted definition in the industry of a chronic disease, let alone how you identify the people who have that chronic disease. But if you assume for the purposes of your study that you know who is in that chronic disease bucket, then you need to figure out if there are any members that you want to exclude from this study. For example, there might be people suffering from some kind of catastrophic condition who are exceeding their stop-loss limit, so you don't want to have them and their results confounding the results of the study.

The population of chronic people is divided into those who are excluded and those who are included in the study. There will be some people who will be targeted for programs (maybe some high-risk people) and some who will not be targeted for programs, but they all remain in the chronic population. This methodology is a population methodology. Once you've identified your chronic population and taken out the excluded people, the chronics minus the excluded people are the population that will be analyzed for savings purposes.

This contrasts with methodologies that only analyze the targeted people or, out of the targeted people, only analyze those who choose to enroll in the program, both of which could produce quite substantial potential biases. The only way to do this in a completely unbiased way, or even get close to an unbiased way, is to objectively identify a population such as the chronic population and then not allow any kind of selection, either on the part of the patient or on the part of the vendor, to interfere with that population. Finally, you identify the measured population. The people whose experience you're going to measure is the sum of those who enroll, those who don't enroll and those who are unreachable, as well as those who are not targeted. These categories sum up the chronic population, less any excluded.

I'll give you an example of people who are often excluded and, more importantly perhaps, why they're excluded. One group that is frequently excluded is people who have end-stage renal disease. One reason is that end-stage renal disease is a very severe condition. It's expensive. It's often managed by specialty care management companies and not often by typical disease management vendors. Remember, this is a trend-adjusted method. Think about what happens to the average patient who has end-stage renal disease. Their costs go one way. They increase until the point at which they become eligible for Medicare, when they just drop off the cliff and become zero. So first you've got an increase and then a sharp discontinuity downward in claims. Any kind of discontinuity like that has the potential to distort your trend.

HIV/AIDS is another example of a common exclusion. Those members might be excluded because they're being managed by a separate vendor, as they're certainly not easy patients to manage in a typical disease management program. Transplant cases are another example of a type of patient whose costs frequently increase quite a lot up to the point where they have the transplant, then drop precipitously

and become quite low following it. Again, that could potentially distort your trend.

People are often excluded either because they have unusual claim patterns that are not like average cyclical claims, or because, as in the case of the institutionalized, they're members who are difficult to reach as well as difficult to manage. For example, they might be in a psychiatric institution or in a long-term-care institution. These are people that the average disease management program isn't set up to manage and has great difficulty managing, so they're best excluded, provided they can be objectively identified. Type 1 and type 2 diabetes can be identified, but cannot be identified through claims. It's the same thing with heart failure; types 3 and 4 can be differentiated clinically from types 1 and 2, but not through claims, which means you introduce an element of subjectivity. For this kind of population study, you need to be able to assess the population objectively.

We need to track people on a monthly basis. In our study, we were very rigorous about assigning people to different buckets. There was a "chronic measured" bucket and a "chronic not measured" bucket. We called the non-chronic population the "index population" because that's where we got our trend observation. There may also be some people who are simply not eligible. We keep track of them because, as I keep telling my clients all the time, at the end of the day, for each month I want to be able to reconcile those numbers back to some kind of validated, audited and reliable data source. That's something that's not often done by the industry, at least not to a standard that I'd like to see.

Then you can summarize individual member data. If you can organize people by class and by month in this way, then everything else falls into place because you're able to immediately attach claims to the individuals and you can sum up their claims. But if you do your people count separately from your claims and then try to match them up, it is much more difficult. This method allows you an audit trail so that you can go back and audit the numbers.

Paper 7 is entitled "A Comparative Analysis of Chronic and Non-Chronic Member Trends." In the DM industry, the trend number generally comes from one of two sources. It either comes from the experience of the non-chronic member population or from the experience of the entire population. In a commercial population with only 5 or 6 percent chronic people, it probably doesn't make much difference whether you use your non-chronic trend or your overall population trend, because you would not expect the two to be that different.

But what this method assumes is that the non-chronic trend is a valid proxy or measure of what the trend experience would have been in the chronic population, absent the intervention. In our Paper 7 we wanted to explore whether, in fact, this was true. The assumption is generally made in the industry that it's okay to use these numbers. We wanted to see whether that held out.

Because you're all actuaries (with maybe a couple of exceptions), here's a trend question for you. Which of the following is true?

A. Chronic member trend is higher than non-chronic member trend.
B. Chronic member trend is lower than non-chronic member trend.
C. Chronic and non-chronic trends are about the same.

Let's assume that it's a closed group for the sake of argument. Who votes in favor of "A," the chronic trend is higher than non-chronic trend? One brave soul. Who votes for "B," the chronic member trend is lower than non-chronic member trend? Who votes for "C," that they're about the same? It's interesting that when you ask an audience of non-actuaries that question, people vote for "A." They think that because chronic member costs are higher than non-chronic member costs, the chronic trend must also be higher. But I've asked this question of other audiences of actuaries, and some people vote that chronic member trend is lower than non-chronic member trend, and some people vote that they're the same. In fact, the correct answer, or at least my answer to this, is that it's non-disclosed alternative "D," which is "It all depends on how you define these things." But I think the people who voted for "B," which was the majority of the audience here, is probably generally right.

Let's look at why this is so. We looked at a commercial database of more than 1 million member lives over four years, from 1999 to 2002. It's the Reden and Anders database, for those of you who know the data. We excluded groups that didn't have four continuous years of exposure, but not individuals. So if the employer came in or left during the period, they were excluded, but people were allowed to move in or move out from those groups that were included. We were doing this analysis at the allowed charge level, not the net paid.

The total overall trend for the entire population over the four years, not distinguishing between chronic and non-chronic, worked out to about 16 percent. That may be a little higher than you're used to over this period, and this may be due to us restricting the analysis to continuously enrolled member groups. For the non-chronic members, the trend was slightly lower. It worked out to be just under 14 percent, which is a little lower than the average. But, as was said earlier, in a commercial population with relatively few chronic members, you would expect the total trend and the non-chronic trend to be about the same. The really interesting number is the chronic member trend, which is less than half that, at under 6 percent.

As I said earlier, it entirely depends on how you define who's chronic. The key point about these results is that you're identifying chronic people prospectively during the four-year period. At the point that somebody meets your criteria for being included in the chronic population, they're moved from the non-chronic population into the chronic population.

As an example, an individual might be in the non-chronic population for a couple of months and then be found as chronic because he or she meets the criteria. When you meet all the claims tests that are applied, you're moved into the chronic bucket. So what you see is a constant migration of people from one population to another. You may or may not think this is reasonable, but in the industry that's frequently the way it's done. People are identified and then put in the population at either the point of first identification or back to the first of the year in which they first meet those criteria.

FROM THE FLOOR: Would your methodology permit someone to go from chronic to non-chronic?

MR. DUNCAN: We didn't test that in this particular study simply because we didn't have the time and resources to do that. You obviously can move non-chronic to chronic, but there's a lot of debate as to whether you can move from chronic back to non-chronic. The clinical answer to that is you can't, because a chronic disease by definition is permanent; you have it for life, so you don't ever move back. But we're identifying these people through claims data, and when you do that, depending on how sensitive or how specific your criteria are, you potentially identify a lot of false positives.

Moving from one year to another, reapplying the same test at the beginning of each year, if you requalified people, you would sift out the false positives. We did not do that. In year one, chronic prevalence was 4 percent, but because we did not requalify people, by year four it had doubled to 8 percent. That has some very significant implications for the savings calculations later on, because over time you're building up your chronic population that, ultimately, will drive the savings in combination with this trend assumption.

In order to see whether we'd get the same results by applying a slightly different test, we compared the prospective identification with what we called a retrospective classification of people into chronic buckets. If we identified an individual at any point in the four years as a chronic person, we counted that person as chronic from day one. That way, we avoided this problem of migration between populations over time.

The overall trend is still 16 percent, but now the non-chronic trend has actually gone up a little, to just over 17 percent. The non-chronics are now people who were never identified as chronic at any point over the four years. The chronics are those people who are chronic at any point in the four-year period. But we avoided the migration effect and are comparing their experience over the four years. What we see is that, give or take some statistical fluctuation, the numbers are pretty much the same. So one way, at least in theory, to avoid the potential distortion from assuming that chronic trend and non-chronic trend are the same might be to classify your populations differently and retroactively classify everybody into one

class or another from the start. Then you would be able to use your non-chronic trend as an estimator for your chronic trend. In our paper, you'll see the result of testing some of these assumptions on real data.

I want to say one more thing about chronic versus non-chronic. We also looked at whether it might be possible to use risk adjustment as a technique to adjust for the changes in risk over time. We used DxCG risk adjustors, which many actuaries use and are familiar with. I'm not entirely satisfied with the results, but there does seem to be some promise in using risk adjustment to take the effect of progression over time out of the trend. If you adjust for the change in risk, it might be possible to use this prospective method without having to reclassify people from the start, because there are problems when you reclassify people back to the beginning of the period. You avoid the migration problem, but you create some other issues. I'll talk about that a little more when I discuss Paper 8.

Now Henry is going to talk about some of our results of Paper 8.

DR. HENRY DOVE: After I describe the nature of this study, Ian is going to talk about the various alternatives that we analyzed as we calculated the savings from disease management interventions. This involves a very interesting opportunity that we had. Health Dialog was the disease management vendor that provided these services to Highmark. Highmark has 2.5 million members in three populations. They have a Medicare Advantage Program, which has about 200,000 members. They also have a commercial population and another PPO population.

What struck me as interesting about this opportunity was that we had three years of data starting from July 1, 2000, all the way through September 30, 2003. Health Dialog was providing disease management services for the last two years, and so they were interested in assessing the savings that were produced for Highmark during this period. As a non-actuary, I think this is an example of the opportunities that actuaries and other health services researchers are likely to have over the next few years, because the disease management firms and the managed care organizations are investing hundreds of millions of dollars in these vendors. They want to have good estimates about the cost savings because, after all, they're generally for-profit companies.

So how does one do this? Some academic purists would say that the only way is to have a randomized controlled trial. That's the only way that you can make these calculations. I didn't realize until a few years ago that the first randomized controlled trial didn't even start until 1947. That shows you that randomized control trials are a recent innovation. Purists like them, but these randomized trials are very expensive to do. The drug companies will let you know that it costs a billion dollars before they can bring a new product to market. So one of the challenges that clinical epidemiologists face is to think of other ways to evaluate therapy or to evaluate programs like this besides randomized control trials.

Actually, the notion of a control group is something, as Ian mentioned, that should be part of any rigorous evaluation. For historical purposes, to my knowledge, the first time a control group was mentioned in the medical literature was around 1830, so the idea of a control group is a fairly new idea as well. Some of what Ian is going to be addressing when he talks about the evaluation of these interventions over this two-year period is very important. There could be some very good opportunities for health-care actuaries. I think health-care actuaries could be involved with other studies like this. They could become the people who are relied upon to certify the cost savings, because these are not simple calculations. We're talking about literally tens, if not hundreds, of millions of dollars. I think that this is a great potential opportunity.

In the research for Paper 8, we used the base case, or what Ian called the actuarial methodology. We only did this on the Medicare Advantage population, which included about 200,000 lives. In order to be considered in the study, the patient had to have six months of continuous enrollment in the base case. In order to identify who was chronic and who was not chronic, in the base case we looked at pharmacy data, hospital claims and other medical claims. We considered either the primary diagnosis or looked at all the secondary diagnoses in order to decide whether an individual had a particular chronic condition or not.

In this base case, we calculated what the savings were in year one and year two. But the purpose of the paper is to investigate using different methodologies, so we have five different alternatives. The first alternative involved a cohort study, where we simply identified a certain group of chronic patients and then followed them forward and looked at the savings attributable to those patients in that cohort.

The second alternative was to consider three different ways of identifying the chronic patients. First, we only looked at hospital claims, because it's believed that hospital claims are more accurate than a physician's claims or claims from the pharmacy benefit managers. Next, we used all of the claims except the pharmaceutical claims to identify the patients. We only used the hospital claims or the physician claims, but we used primary as well as secondary diagnoses. Then we used all the sources of the claims from hospitals, physicians and pharmaceutical, but we only used the principal or primary diagnosis. We could say that we believed the primary diagnosis, but we didn't necessarily trust all of the secondary diagnoses. So those were the three ways to identify the population under study. The third alternative methodology was something that Ian talked about before, which is the retrospective identification of the chronic members—looking backward, so to speak. The fourth alternative was to examine the requirement involving six months of eligibility. Suppose we didn't require continuous eligibility? What impact would that have on estimating the cost savings? I don't think we have conclusive answers for that. In the fifth alternative, instead of looking at the Medicare population, we looked at the commercial HMO population or at the point-of-service

population. Those were the different alternatives that we considered as we had this incredible database over a three-year period that we could use for these kinds of calculations.

The baseline period ran from August 1, 2000, to June 30, 2001. The actual cost per member per month (PMPM) was $448.26. So that was the starting point. Then we evaluated the total costs and calculated the cost savings per member per month in the first year and in the second year. We also used the trend factors that Ian talked about of 0.5 percent for the chronic population and 9.7 percent for the non-chronic population in the first year. The base-case analysis showed cost savings of $22.1 million in the intervention year one and $39.7 million in year two. The cost savings as a percentage of total claims for the Medicare business was 2.0 percent in year one and 3.0 percent in year two, calculated using the actuarial method.

I'm going to turn it back over to Ian because he's much more familiar with the alternative methods. It is interesting how the savings vary according to the assumptions that you make. As Ian said before, the base savings only represents the method that we proposed to do evaluation of disease management interventions. I think that Ian would agree that the final method has not yet been designed. This is the best that we have come up with thus far, but I think the opportunity for actuaries is to think of other ways to do these kinds of evaluations. Five years from now, when someone is talking about these calculations, hopefully the underlying data will be better. The methods for handling the definitions of the chronic patients will be more precise, and the enrollment or eligibility data will be more precise. But this is a starting point.

MR. DUNCAN: It is important to note that the underlying population has been growing over this period. It started at 158,000 members and grew up to close to 190,000 members. It is hard to control for growth in the population. As I said earlier, we did not requalify people at all, so there has been some increase in chronic prevalence, from 21 percent up to 27 percent. It may not seem all that much; in a Medicare population you don't see the sort of rapid growth in chronic prevalence that you do in a commercial population when you don't requalify people because in Medicare, once they have a chronic condition, they generally do reappear the next year.

The index measured population trend is 9.7 percent in the first year and 9.9 percent in the second year. These numbers do not seem exceptional at all. They seem perfectly middle of the road and are, in fact, maybe a little low for a trend assumption in 2001, 2002 and 2003. That's what was used to do the projection. The projected claims are claims per member per month for the chronic population. If these numbers seem high to you, it's because they're chronic people's claims. They're going from $448 dollars per member per month up to about $540, and the actual claims are lower. There was a fairly significant increase—about a 50 percent increase from year one to year two—in the savings calculated per head. Some of

that is an incremental effect of two years of trend being applied to that base number.

The first alternative was to look at a cohort design. We wanted to test if you just chose a population and followed cohort forward, does that do anything different from the sort of open population where you're identifying people constantly and bringing in new people? We started with 33,000 people, and it slightly increased in the second year. But in the third year, it effectively declined down to 29,000 people because it's a closed cohort. The trends are exactly the same, and the projected numbers are also the same. The actual numbers are slightly different, but they are pretty consistent.

My expectation was that the numbers would be much more different than they, in fact, turned out to be. So this kind of closed cohort gave me results that were a little surprising, and I'm still thinking about why that may be. My initial thought is that it has to do with the fact that we apply some fairly rigorous conditions, such as the six months of continuous eligibility and the three months of exclusion after identification of the first event.

Regression to the mean has always been something that people have been concerned about in these studies. Clearly, if you look at an individual and the identifying event of an individual, an individual can have a high claim because, for example, he or she is in the hospital and that might be the individual's identifying event. Three months later, the individual is back to a stable condition again, so if you just looked at that one individual over time, you might appear to be saving money, but two things go on. First, in this methodology we exclude the first three months of experience after the identifying event. It doesn't get counted to avoid counting the regression.

Second, this is still a population methodology. You've got a large population of people, and some of their claims are going up and some people's claims are going down. What matters is what's happening in the population as a whole. I don't know if it's a fallacy or simply a misperception or perhaps people haven't thought about this enough yet, but there's an extrapolation that happens in people's minds. The behavior of an individual whose claims decrease quite sharply from the point of identification forward is extrapolated to an entire population. People talk about the regression to the mean and the population as though the population behaves like an individual. But, in effect, people's claims are going all over the place at any point in time. If you track populations, there's far less extreme movement than there is in individuals, and that seems to be the case here.

If you remember in Paper 7, we pointed out that migration bias drove the difference between non-chronic and chronic trend. In Paper 8, we wanted to see what would happen to those savings calculations if we go back and classify somebody as chronic from the first moment in time, not from the point of identification forward. The

results were really quite remarkable. What we calculated as savings in the first base-case example disappeared in the first year. In the second year there were savings, but they're much lower. This seems to suggest that if you want to avoid the effect of migration bias and you classify your population as chronic or not chronic from day one, you can correct for the migration bias, but somehow or other your savings disappear.

Now, is this the right answer? By classifying a number of people as chronic who are not chronic on day one, your chronic pool now consists of people who are both chronic and non-chronic. So you're following a population of people who are chronic and non-chronic, some of whom are being managed—obviously, the ones who are not chronic are not being managed—and you're measuring the effect on that entire pool. We feel that there has to be a way to be able to sort out those people and extract the true chronics from that pool. What you're seeing in terms of the savings disappearing is probably an overstatement. It might be that the savings calculated in the base case were overstated. These results seem to be an understatement. Perhaps the true number lies somewhere in the middle. We've got to be able to figure out exactly where it is, but at least the different tests show us what the results might be.

Removing the six months of prior eligibility and this three-month exclusion requirement drives the savings up. In the base case, savings in the first year were $41.54 per member per month. The cohort savings are about 5 percent lower. If you identify people from medical claims only, excluding the drug claims, then savings actually go up on a PMPM basis in the first year. If you use the primary diagnosis, they go up even more. Using only the hospital claims produces a result very close to the base case. The retrospective identification, as I talked about a moment ago, wipes out your savings in the first year, but elimination of the no continuous eligibility or waiting period requirement in the first year increases your savings by about 50 percent. We also did an analysis of the commercial HMO and point-of-service product savings. They are not all that different from the Medicare Advantage—a bit lower, but not much different.

What does this tell us? What can we conclude from this? The first point is that savings results can vary considerably depending on identification, method and assumptions. In other words, these results are extremely sensitive to the assumptions that you make. In everything I've ever read, be it in the peer review literature or for client work that I've done, disclosure is inadequate. You never know all the assumptions that are going into studies. Because assumptions can have such significant impact on the results, it's very difficult to determine, short of redoing an entire study yourself using your own assumptions, whether results are credible.

You might be able to do some digging into results if you have a good audit trail. I recommend a good audit trail and, obviously, complete documentation of assumptions. But caveat emptor: If you look at results that can go from the base

case, they essentially have a 100 percent swing either up or down, depending on some of the assumptions. In order to understand specific savings results, a great deal of information and disclosure is required. We did not perform any tests varying multiple results; all of these scenarios are univariate scenarios. Obviously, if you combine some of the scenarios, you're going to get different results, but we did not do that. We continue to test other assumptions. The one that we would like to complete before we finish this study focuses on what happens if you requalify people on an annual basis.

One of the things frequently asked about, which we did not look at here, is what savings results look like by disease. This is a population study; the program that was put in effect by Highmark was a population management study. Frankly, there are enough problems and difficulties in assessing the results at the population level, let alone trying to deal with how you classify somebody into a particular disease category and how you move them over time as comorbidities develop. This makes the assessment even more difficult. So that's something that we're going to leave to others to look at.

DR. DOVE: I'd like to make a couple of other comments. When people want results by disease, they assume that we can associate a cost with diseases. But we can't. We don't have the tools for doing that at the present time. If you get a diagnostic test on a patient that has congestive heart failure, it's very hard to determine if you should associate that cost with the patient's congestive heart failure or diabetes or any other things he or she might have.

One question that I have about this, Ian, has to do with what proportion of the people were actually managed by the disease management company. Let's say you had 24 percent of the patients that were chronics. But it may be that only 5 percent of those were actually contacted and actively followed by the disease management company. Looking to the future in five years, we should have information on whether a person who is a chronic got an actual intervention, whether he opted out of the program or whether he dropped out of the program. Do you have information on that?

MR. DUNCAN: That information was not available, Henry. It was not captured and so it was not available to us. But there are a number of other tests that one would like to apply, such as how long somebody is managed. It can be very difficult to contact people. In one population you might only be able to contact 60 percent of the people, and in another group you may be able to contact 80 percent of the people. Another is how you compare results from different populations with completely different profiles of risk characteristics at the population level. It's hard enough to do this kind of simple population study, let alone allow for all the myriad complicating factors, but we have to start somewhere.

FROM THE FLOOR: My question is whether you looked at either the applicability of

methodologies or which methodology you would use based on the data source, whether it was commercial or, for example, a Medicare data source. Also, did you look at the applicability if it was a commercial purpose versus a Medicare purpose?

MR. DUNCAN: I'm not sure that I understand the question, but I will say this. Even though it's a Medicare Advantage population, it's a Medicare Advantage Program run by a commercial HMO, so this is not Centers of Medicare and Medicaid Services (CMS) data. This is the health plan's data.

FROM THE FLOOR: Right. You're getting the data from a commercial provider, and you can choose to identify people prospectively or retrospectively. Obviously, you always have a data problem. My question is, if you had received similar data from Medicare rather than getting it from Highmark, do you think you would have been able to have fewer false positives by using a prospective rather than retrospective, or vice versa?

MR. DUNCAN: I don't know. I never thought about that. Ultimately, if you wait long enough, isn't the data going to be the same? Though if I wait for Medicare, I can never get the drug benefits. These people did have drug benefits in the Medicare Advantage, so I'll never get that. But for everything else, if I wait long enough, the data ought to be the same, right?

But you have raised something that I want to discuss. We talked a lot about research earlier, but we actually did not do the work. Highmark did the work. We were very insistent on that because of privacy and other issues, so Highmark did the work under our direction. But even so, one of the things we ran into is a new provision in the Health Insurance Portability and Accountability Act of 1996 (HIPAA) that says that you can't do research on health claims data without obtaining the permission of the patient. This is new as of early 2003. There are certain ways around it. You can de-identify the data, but the rules for de-identifying the data are very rigorous. You can set up an independent review board. The rules are skewed toward academics who, obviously, have these independent review boards in place. So as actuaries who hadn't been down this road before, we found some complicating factors. I hope, as a result of some of the things we've said, that we've inspired the rest of you to go out and start doing research. But I would caution you to read the HIPAA rules first.

MR. ROBERT PARKE: I find your results very interesting, but what they suggest to me is that all of the savings are an artifact of the calculation. We've done some internal research ourselves where we thought it would be interesting to apply those same kinds of methodologies that you used to a client database that didn't have a disease management program in place. We used three years' worth of longitudinal data, doing exactly the same things, and we got similar results to what you're showing in terms of savings. It makes us all wonder about the savings you're putting up there and how valid are they really, when you can actually mirror exactly

those calculations in a population that doesn't have any specific disease management program in place. So it's an interesting sort of observation, but it doesn't answer the question of whether these are real savings.

MR. DUNCAN: That's exactly the point. I agree entirely with that. But I tie that back to my comment about economics. One of the things that we may be losing sight of here, because the disease management programs have a certain size and scale and are applied to these very large populations of people with certain conditions, may be that in the entire population you're not going to see a lot of savings if you were to apply a valid test, and maybe we have to think about whether there's a way to identify a narrower but more manageable population and assess the savings on them. I don't know the answer to that, but I think that one of the things that the industry loses sight of is who the patients are that they can really manage toward the savings.

MS. CINDY MILLER: Ian, I'd love to have a methodology that's the actuarial methodology, but given the wide variation in results, depending on the assumptions and the classifications, has it occurred to you that maybe it's the methodology that's at fault? It's entirely possible that maybe we just haven't found the right way of defining it and that if you tweak it enough and refine it enough, it will be credible. But something between a loss in the first year to $20 million in savings is too wide a variation to me to say that this is the methodology that we can espouse as the right way to do this.

MR. DUNCAN: I think you need to distinguish between methodology and assumptions. A lot of this is driven by the particular assumptions and the particular definitions that you apply. I don't think that we've got the right answer yet. I think that's the point of Rob's comment earlier: we don't have the right way to do it and the right answer yet. But maybe for all of us who do these kinds of things, if we could get more coherence around one set of assumptions, then at least we could benchmark our results against each other. That might be the best that we can hope for for the time being.