# RECORD, Volume 27, No. 1[*]

Dallas Spring Meeting
May 30–June 1, 2001

## Session 96IF
## Data Quality Concerns for Health Actuaries

**Track:**              Health

**Moderator:**      KARL G. VOLKMAR
**Panelists:**        JAMES E. DRENNAN
                       DAVID V. WILLIAMS†
                       J. MICHAEL CROOKS

*Summary: As the volume of data available to health actuaries from both inside and outside their organizations continues to grow, so do concerns regarding the quality of that data. Panelists address topics such as techniques for testing data quality, the actuary's responsibility for verifying the reasonableness of data, and considerations in modifying data sources for specific situations.*

**MR. KARL G. VOLKMAR:** I have been a consulting actuary with United Actuarial Services in Carmel, Indiana for the last two years. Prior to that, I was the second vice president of supplemental health product development at the Conseco Companies and the vice president of supplemental health product development at Pioneer Financial Services. I have more than 15 years of insurance company and consulting experience working with a variety of individual and group health insurance product lines.

Our panel includes James Drennan, principal of Tillinghast-Towers Perrin, who has been with them for 13 years, 10 of them in the St. Louis office and the last three in Atlanta. Prior to consulting, he worked for Blue Cross/Blue Shield of Louisiana as the chief actuary, Philadelphia American Life as group actuary in the Houston office, and the Texas Department of Insurance. He graduated from the University of Texas at Austin.

Mike Crooks works in the Kansas City office of Ernst & Young's Health Actuarial Services Group. He has 16 years of experience in the life and health industry, primarily focusing on pricing, product development, financial management, and

_____

client reporting.

David Williams is a health care consultant for Milliman USA, formerly Milliman & Robertson, in their Hartford, Connecticut office. He was previously with a small Connecticut HMO startup, and before that he worked for Kaiser Permanente in medical economics.

The Actuarial Standard of Practice (ASOP) No. 23 is organized and consistent with other ones that I've seen and reviewed.

**SECTION I**
**PURPOSE: SCOPE, AND EFFECTIVE DATE**
In section I, the purpose is to provide guidance to the actuary in the areas of data selection, the review of data for appropriateness, reasonableness and comprehensiveness, and in making appropriate disclosures. It applies to all areas of practice as of December 31, 1993. Basically, it says if you're an actuary, this applies to you. It applies to everything that has to do with data for every practice area as of December 31, 1993.

**SECTION II**
**DEFINITIONS**
Section II lays out some definitions. One of them is appropriate data, which is data that is suitable for the intended purpose of the analysis—data that is relevant to the system or process being analyzed. Then there is comprehensive data which is data obtained from inventory or sampling methods that contain each data element of record needed for the analysis. The reason I point out these two is that there is a pretty significant difference between them. I think every actuary has worked with data whose variables interrelate. This is very important, because you can have appropriate data for a variable and for a number of individual variables, but you may not have comprehensive data that shows you how the variables interrelate.

**SECTION III:**
**BACKGROUND AND HISTORICAL ISSUES**
Section III really boils down to two comments. First, data is important and therefore, data quality is important. That's obvious to us. Second, data appropriateness and disclosures regarding imperfections are important to users. I think this is well addressed in the standard.

**SECTION IV:**
**CURRENT PRACTICES AND ALTERNATIVES**
Section IV was meant to address the current practices at that time, and yet I think it's very true today, which is why we need this standard. It's important that data is relevant to the system or process being analyzed. Data quality assurance ranges from reliance without checking to a complete and independent verification, which I think is still true today, depending on the circumstances involved and the purpose of

the data. Actuaries provide analysis for a broad range of uses, from limited distribution to public exposure.

Last, in this section there is a range of practices with respect to documentation, disclosure, sources, material biases, imperfect data, adjustments made, and so on. All of these points really show the need for the standard of practice meaning that there is a wide range of things going on, the issues are very important, and hence, we have the standard to follow. I think this will come out more when we go through the case studies at the end.

**SECTION V:**
**ANALYSIS OF ISSUES AND RECOMMENDED PRACTICES**
I think this is many of our careers in a nutshell. Section V basically begins by saying you're not going to have the data that you want to have in a perfect world, and you're going to have to do your best with what you have, document what you don't have, and document the limitations of your work. I think that's very true. The other thing that it reaffirms is that this standard does not recommend an actuary audit data. One of the things that was thought in the fellowship admissions course was that ASOPs are guidelines. In many situations it's hard to know exactly how ASOPs apply, so actuaries must rely on their judgment to make these kind of decisions. I think you'll see this to be true throughout. There are three sections on recommended practices: selection of data, the use of imperfect data, and reliance on data supplied by others.

**Selection of Data:** The intended use will indicate the nature and extent of review needed and the number of alternatives to be reviewed. In a lot of ways it's common sense that if you're looking at a big-ticket item, you're going to spend more time looking at it and trying to understand it. In evaluating any alternatives, the actuary should consider data elements and possible alternatives and then select the data, considering the appropriateness versus the purpose of the analysis— including how current the data is. Actually, the standard very specifically talks about paying attention to internal and external consistency, the reasonableness and comprehensiveness of data, any limitations, modifications or assumptions that you need to use it, and the cost and feasibility of alternatives. It clearly says that one of the costs that you need to consider is the time and effort that it's going to take to get the data that you need.

**Imperfect Data:** This standard really asks two questions. First, will the imperfections produce material biases in the results? Second, which I suppose is an even bigger question, are the data so inadequate that they cannot be used?

**Reliance on Data Supplied by Others:** This is actually the majority of what's in the recommendations. Any reliance should be disclosed. Data accuracy and comprehensiveness are the responsibility of the supplier, which is interesting, because before I started working with the standards in practice, I always thought

that it was the responsibility of whoever had to do the project. That's the way it seemed in real life sometimes. The data should be reviewed for reasonableness and consistency, and again, I think tying in with what we talked about before, the intensity of it should vary the purpose and the possible impact.

When a review is performed, data should be reviewed for questionable values or inconsistent relationships. The standard goes on to say that additional compilations are not required. If material defects emerge, the actuary needs to determine the nature and extent of auditing that may have been performed. The bottom line is, if it justifies additional review, then you need to do it. Last in this section, if it's impossible or impractical to perform a sufficient review, the actuary should disclose that along with any resulting limitations.

As a consultant, I know things often are not quite as clean as you'd like them to be, but I think the standard is very clear about explaining the impacts of this and documenting it.

**SECTION VI:**
**COMMUNICATIONS AND DISCLOSURES**
Section VI states that you need to maintain documentation supporting the specific data used for a time period that will depend on the purpose of the analysis, client needs, regulatory requirements, etc. You need to keep that documentation as long as you need it with the presupposition that you're going to document it. Something that actuaries are not necessarily the best at is documenting things. To us, many times, the work is done when it's done. I know both as a company actuary and as a consultant, it's been interesting at times to re-create things that have already been done.

Actuaries' reports should include the following: the sources of data and the materiality of potential biases. This is sometimes hard to estimate, but it is important. Any adjustments or modifications made because of imperfect data should be included also. This is something that I think is missing probably more than all the others. People tend to understand what the key variables are and how they're going to vary, but it's very difficult to take raw data and then get to the data file that was used for whatever the project is. Those things aren't documented as much. Again, report the extent of reliance on data supplied by others. If data has not been reviewed, any unresolved and potential material concerns about the data must be disclosed. There's a caveat at the end of the standard that says for any procedure that departs materially from this standard, the nature, rationale, and effect of the departure must be disclosed. That's par for the course, as we all know.

Now that I've laid the groundwork, the other panelists are going to talk about what this subject means to them. Then we're going to go through some case studies and have some discussion.

**MR. J. MICHAEL CROOKS:** I requested to discuss a specific topic—what's going on at this point legislatively in data quality and how that may impact both our ability to  access data and how good that data may be.

One thing Karl talked about was the issue of data quality, which probably means different things to most of us. To me, there are only two aspects to it. One is what you think is the normal, which is how reliable or accurate the data is that you're receiving. The other is the issue of access to data. The quality of the data you have may be materially impaired by the fact that you don't have access to the data that you want or need.

When I started looking at what was going on legislatively, there were really three major categories. The first one is the Health Insurance Portability and Accountability Act of 1996 (HIPAA), which we all are fairly familiar with, and it really overlaps the other categories as well. The second is the issue of a right to privacy, which has been becoming more and more prevalent in court cases and other things. From the general discussion, it looks like there will be some sort of right to privacy which will probably be codified through federal government or through some state agencies in the near future. Last is the issue of genetic testing. It's pretty clear where this issue is going to land as far as what access insurers may end up having to test results.

With HIPAA, there are really three aspects that are going to impact data quality. The first two go hand-in-hand. You have privacy standards and security standards, both of which are going to impact your ability to get your hands on data to some extent. The way that things are written right now, there are provisions that allow insurers to have access to it for rating purposes. What's not so clear is if you're doing studies along some other lines that aren't clearly related to rating an account. Whether or not you'll run into privacy standards issues if you try to access information for a particular purpose is one concern perhaps for future data quality, both along the lines of what's coming in HIPAA and some other legislation that's going on that's out there.

The good side, at least from the standpoint of data quality, is the administrator simplification. This has been pushed and pushed, and now I think it's about a two-year horizon until it's going to have to come into effect, unless it gets pushed again. This has the potential of actually improving the quality of data, or at least the access to data that we may have, or may require, for uniformity of data records that are going to be reported. They're going to be reported electronically, so hopefully you'll have access to more data as it's submitted electronically and it will be easier to get your hands on it.

Another thing Karl mentioned was the consideration of the cost of getting the data in the first place. If it's being submitted electronically, you should be able to obtain that data more readily. One concern would be that if you can standardize the data

fields, and if you can require that they be transmitted electronically and privately, you can't require that they have anything worth looking at inside of them. HIPAA is just one of the many initiatives out there right now that has a right to privacy incorporated in it. The Department of Health and Human Services (HHS) has proposed some national legislation. The Gramm-Leach-Bliley Act, which is really a sort of banking act but actually addresses all financial institutions, also has some privacy and security clauses in it that really will affect, and currently do affect insurance companies, probably no more than the HIPAA's going to, so that's not our big issue.

Some states are taking action, and the big thing is that it's risen to a fairly high level in the public's awareness, and because of that, I think you're going to get some sort of legislative response coming in the near future. We may have a little bit of delay with the recent change in the makeup of the Senate. There's that slim possibility that they won't be able to come to an agreement on what some sort of HMO bill of rights would look like, but almost all of them have some sort of security issues in them, and if that doesn't come through, HIPAA almost surely will.

HIPAA originally addressed electronic transmission of data and the security and privacy involved with that. It seems fairly clear, with some of the recent reporting that's come out, that it will probably be extended from a regulatory standpoint to all transmissions, which will have the justification of saying if you don't secure all transmissions of data, then you can't realistically control electronic ones. I think you can expect that there won't be a loophole in that respect, and that something will come down in the next couple years, at the very latest, that's going to somehow restrict your access to data that's available.

The final piece on the legislative front is genetic testing. It's a field that's still in its infancy. They just completed the mapping and they're starting to dig through and try and figure out what they can do with it. For a lot of us who have worked in the group field, that's probably not a big issue. We would waive those sorts of things. We don't do a lot of medical underwriting. In small group, that's always a potential, but that's being regulated out as well. The real potential would have been in underwriting individual products. I'll throw a few out there, but the main thing is it's such a hot and high profile item, that the likelihood that we would ever be able to use genetic testing on anything is probably zero.

If you think back, there are plenty of other issues that we basically have been prohibited from using. You can look at race, sex, and in some instances, we're even restricted with AIDS because it's such a high-profile item, but there are still restrictions there. The likelihood that there's going to be any kind of genetic testing that ends up flowing through and improving our ability to predict is virtually nil.

As a quick recap, I think the quality of data is probably going to be enhanced by the move to electronic submission and standardization. As I mentioned earlier, though,

my big concern is having seen these sorts of things that have come to pass. If the data is not being used specifically for reimbursement, I would be very leery of trusting what's in that field if it's not somehow the employee information which would be linked to determining whether or not there's coverage.

As an example, I would throw out the quality of ICD-9 data before diagnosis related groups (DRGs) came around and your compensation was being driven by what codes were going into those fields. I think you're likely to find the same thing now. They come out with this standard that says here's a field you need to provide. But if some of those fields aren't being used to determine the payment that's going back, I think the likelihood of the quality being good is very low.

The other issue is security. How much it's going to impact us is not clear because, as I said, most of the stuff that's out there, including HIPAA at this point, gives exclusions for insurance companies to use data when it's needed for pricing purposes. My concern would be that, as it evolves, where's that line going to be drawn? If I'm doing an experience study that's generic for some reason, am I going to be able to roll that in and really say that what I'm trying to figure out is how my rates are going to be affected long-term on a group basis? Or are they going to say, "You're not rating a group, you're looking at your block of experience trying to determine some other subject, so you can't have access to this data." That's probably a fairly minimal concern though. From the survey I did, it looks like the legislature is balancing reasonably well business's need to access with the concerns the public has about privacy.

Finally, I will give you my idea of quality, and where I stand. I think standardization is a good idea, but again, I'm really concerned about the quality of particular fields as long as they're not being related to reimbursement. I just don't see that there is any likelihood that those are going to have useful data in them.

**MR. JAMES E. DRENNAN:** I'm going to talk about the valuation issues related to data quality, primarily claim liabilities for medical, dental, and prescription drugs, and types of short-term, short-type products, and how they relate to statutory valuations and your internal valuations. These are things that you, as an actuary, might be working on for your company, or a consultant might be working on, and data problems you may have.

One of the key things that came out of an earlier session was always to restate your prior carrier. That will help you figure out whether your current valuations have some data problems. If you haven't restated your prior carrier, you may think something is going on, and when you do restate, and you find it's not, you'll see the errors coming through.

Another thing that really helps, which is probably obvious, is to get more months of runout. That's almost always the best answer. Get more months of runout, check

the data, and see what's going on. Karl mentioned that there's supposed to be a disclosure of imperfections. I find that very rarely happens to us. We generally wind up having to find them ourselves, often a year later.

Testing is obviously what you should do early on. That's the actuary's responsibility as outlined in the standards. The first one is clearly to compare to paid claims. You can use the ledger or a preliminary financial statement, but you should really look at paid claims compared to something that's fairly consistently accurate and something that you feel you can rely on. You may have to investigate the best source.

Another test is to look at days per thousand. Generally, this is going to help you only on the in-patient side, but it does give you some hint if something is going on. If you look at a history of days per thousand over time, and you see something funny going on, it should make you suspicious. The reason we like this is that it comes from different sources, and that's often good. The days per thousand often come from a different department using different data, and it may show something that you don't see in your claims triangle.

Looking at per member per month (PMPM) is clearly an indicative test of your data. When you look at it over a period of time, obviously you'll find some swings month to month, so typically we look at it on a running 3-, running 6-, and running 12-month basis. If you see something you haven't seen in the past in your most recent valuation period, you should get suspicious and look for an explanation. There could be some obvious ones; such as they added a new product such as Medicare. Then you would probably want to do it separately. You'd want to split it in this case. It could be a new system. That often is the culprit in data problems. It's the most common one I have seen if there's a new system and an old system. I like to get the triangle separately and find out whose planning the runout. Did the old system plan the run, or did it cut off immediately and run on the new system? Look at them separately to see some patterns, see if there's a difference, and then try to make a judgment going forward. If you combine them and there's some overlap, the runout from the old one is overlapping the new one, and you may not see those patterns on the new system. Obviously, you're trying to predict the future patterns, so you want to look at what the most recent clear data indicate.

We frequently see negative payments or refunds. These clearly come about from overpayments. Your system may have overpaid providers, or maybe it's just routine errors. Sometimes it's a system change. I've seen one that was an advance payment to providers, and then they recovered them later. We'll talk about that a little bit more. The main thing we try to do is tie the refund back to the original payment. It sounds easy, but it rarely happens. You may find the original paid date of the claim, but where it was in your claims triangle is almost never available. As a fallback, we do lag factors separately. We'll take the lag factors before these negative payments. I call them gross factors. Because you can't take

out these new payments at exactly the right cells, you basically make the assumption that the overpayments have the same average lag as the net payments. That's just an assumption, and you have to test that a little bit, but you really can't test it well enough to ever know for sure. Then with those completion factors, those lag factors, take your paid and reduce it by your refunds as well as you can and then apply the life factors and get a reduced incurred number. It's not very good for reinsurance, because reinsurance claims typically have a very long table. If you've done a lot of refunds, it doesn't work well, and you may have to use some alternate methods, such as assuming a different lag pattern based on your best guess, but that's a real problem with data.

Some of these are continuing and some of these are one-time payments. Most of the time it's a one-time payment—your claim system overpaid and you're doing a recovery. We often have clients that come in and say, "We think we overpaid by so many millions of dollars; therefore, we should reduce the reserves by that much," and it's not that easy. The refund recovery rate typically is not 100 percent. In fact, it's quite low in my experience, because the money's been paid out and getting it back is very difficult. If it's an ongoing provider that you're going to make payments to in the future, yes, you can take it out of his future payments. If it's to a member, it's very difficult to get it back. So the percentage reduction is a real test of debate between the clients and us, often as to what they think they'll recover— it's never any firm number, it's just your best guess. We usually give some credit when we can clearly see that there have been overpayments. It's a very common problem. We're finding more and more of that in recent years, and I'm not sure why, unless we're just getting more complicated provider payments.

A payment prior to the incurred date sounds implausible. The first time I saw it, I thought it was just a coding error. Well, we got into it, and we saw after a couple of times that it really wasn't a coding error. I talked to another actuary who said he'd seen it on different types of claims too. The most common one was a provider group that was very unhappy and was threatening to pull out, so they were given some funds in advance to be applied as claims came in.

It was coded as paid prior to the incurral date. Clearly, that makes your triangle more like a square.

**FROM THE FLOOR:** Isn't that an accounting issue?

**MR. DRENNAN:** It is an accounting issue. They put it in our claims triangle. They put them in the payments that shouldn't have been, so we pulled them out and said these are amounts that you have prepaid; therefore, that's a recovery.

**FROM THE FLOOR:** Advanced deposit?

**MR. DRENNAN:** Advanced deposit, yes. So, clearly, it should come out. Then you

should calculate your life factor separately without it. Now, if there are future payments, that's an accounting item that you need to consider separately. We had difficulty explaining the concept that it's not part of your future claims liability. It's a prepaid item, so it's a recovery and it should not affect your future amounts.

The biggest single problem that I have seen is with changes in claim systems. It makes sense that people want to have better and newer claim systems, and the first thing we try to find out is what's different about the new system. Is there some unique feature that makes it better, and will that affect the run-out payment pattern? For example, does it have more edits? Will it reduce claims by not paying some that were paid improperly in the past? Will it have electronic transmission?

Back to your comment about the electronic submission which could improve turnaround time. If the claims come in, often we'll see there's an up-front repricing system, then that's sent over to the claims system. If that transfer is done electronically, you could save some time. A new system may be at another site further away, so you will add some mail time if it's done by mail. You need to look at those things and try to evaluate what's different about it.

The methodology we would typically use is to do separate data sets on the new system and the old system to see the different patterns. Typically, there's a slowdown in payments during the transition and a speedup just prior. We find that the claims departments try to get everything cleaned up prior to the transition so you'll have a faster payment pattern than during a transition, which may be one to three months. It will be slower until they get up to speed, then who knows what will happen after that, and you'll have to try to evaluate what's going to happen now. If it's in the middle of your evaluation period, it really is difficult. You hope that they will do this in an off time if you're looking at year-end valuations, such as summer, so by the year-end you may have some pattern to follow.

The incentive for these claims processes on the old system is often contrary to what we, as actuaries, would like to see. I have seen some in which the claims processors were going to be let go and they were outsourcing. The claims processors were given a bonus if they got things caught up before they left. The quality of claims was very poor during that period; almost everything got paid. They had no incentive. They were paid to get things caught up quickly and get everything out the door. We noticed the per month per member claims (PMPMs) just had a real increase, and then they gradually came back down after the transition. I didn't view that as a long-term phenomenon, that was a short-term one. Financially, it was a disaster for the client in that short period of time. I thought it was a mistake on their part. They were looking at the transition more than the overall picture. I've seen it when it wasn't quite so blatant, when there wasn't the big bonus, but it happens quite often. Anytime the claims department is told to get everything caught up, they typically do not do a thorough job of adjudicating. They will tend to get things done quicker. It seems to be very common, as I've seen it

several times. To test that, we basically evaluate the PMPM claims and look for these increases or spikes. You want to be careful not to use that as a long-term trend if you think that's what's happening. You look at it as a short-term phenomenon and try to estimate what you think the long term is.

Backlog counts are very helpful if you have good data. I don't think I have ever had good data on backlog counts. Once, I actually went back and got backlog counts for every quarter for several years. I went retrospectively and looked at what the restated claims were, because I knew the exact number for each point in time and tried to do a test of how the change in backlog related to that. It had almost no relationship. The backlog would go up and the liability would stay flat or go down. They might actually go opposite sometimes, and sometimes they'd go the same way. There was no real relationship. I concluded that the method was probably not very accurate in counting the backlog. If you're going to use it, you ought to test it and see if it makes any sense, and you need to look at the method. If they're just counting envelopes that are sitting on the shelf and an envelope may hold one or 10 claims, that's pretty inexact.

One company actually opened every envelope and had a front-end entry system. It actually entered all claims up-front and then adjudicated them later. Then I had a very good count. I knew the number of claims. I actually knew a submitted dollar amount. That was one of the best systems I've seen. That's pretty rare. In fact, that was so expensive that they finally scrapped it. Lose a few, win very few.

Looking at the timing of check runs will tell you something about your backlog. This is not unreported, this is more of your reported claims, claims that are in-house but unpaid. If your check runs are once a week, and you can look at how that falls within your valuation time period, you might have a little bit more or less.

For instance, if a check run is five days prior to your evaluation date, versus one day prior, it will change your backlog and change your claims liability. You can make some judgments about that, although you have to use some approximations. Also, you can find when the end of the month falls. A shorter month may make a difference. You also have to take holidays into consideration.

We've noticed a significant difference between when the holiday falls relative to the end of the month. For example, Christmas falls on the same date every year relative to the end of the month. But you need to look at what the staffing is. If a client has a very limited staff, and you find that they have a lot of people off during the holidays, you'll see big jumps in backlog, versus the larger claim system in which they have pretty good coverage and people working fairly consistently. So you ought to get a count of staffing by end of month. I didn't do this at the time, but after I was told that was what had happened with a client, and I realized that with a small claims staff, it's very material. You can lose 20-30 percent of your staff just because of holidays or vacation time.

To test backlog we often use utilization data. Again, this may come from different sources, such as a precertification unit. You can get a rough idea of how many new claims are being precertified, and sometimes even what they are being precertified for. A little bit of information will help you to know if you're getting more precertified claims. You may then have a larger claims backlog, claims liability coming in, and if it's gone down, you may have a little smaller backlog.

An interesting story is about a client that had assumed or taken over some different competitors that were in a different geographic location. The client was having the precertifications sent in. Somehow this one unit in some small town forgot to send in a whole box of precertifications and messed up the whole system. So it's very inexact to rely on hard copy things like sending in a box of precertifications. That's something you ought to watch and not trust completely. This client was using this as its primary method and it didn't work.

I mentioned reinsurance before, and it's fairly similar to what we've said about recoveries. The first thing I tend to do with reinsurance is to look at the asset side to see if it's set up as a recoverable. If the accounting department doesn't feel like it's a recoverable, or if it's not sound or certain enough to be recovered, then I don't really want to count it as a reduction in liability.

You want to adjust your prior claims triangles if you can. Again, you probably won't know exactly where, so you may have to use lag factors with an end, and then make an adjustment and reduce your gain incurred for the reinsurance recovered. Again, those are the ones that have been recovered. Those that have been filed and not recovered are a judgment factor, so you want to look at how certain it is and if it's been set up as a recoverable. You may want to use a percentage reduction to assume that the reinsurer may reduce it a little bit and not pay the whole amount. Again, it's very prone to how material it is. In a big block of business, sometimes your reinsurance is very small.

I have seen some instances in which it was quite material, where they were getting a lot of large claims—specific stop loss where they had a low amount deductible. If it's very material, then you really need to treat this carefully, because it can affect your liability quite a bit.

**MR. DAVID V. WILLIAMS:** I'm not a data quality expert either. (Laughter) I'm not an actuary, but I dig around in databases a lot, and I've had two roles working with the actuaries. My first role is to find data problems when I'm digging around in data warehouses and building data warehouses, and I get paid well for that. My second role is to introduce new problems to the actuaries and provide them with the data. The sites that I have are not necessarily comprehensive or in order, but this is not always a problem.

The first site, and we've covered this pretty well, is the claims inventory. A medical director told me the story of an HMO that he kept having payment problems with. He knew that there were claims out there that weren't getting paid in a timely manner, and when they kept going back, the staff kept assuring them that everything was fine. They said, "Look at our time to payment, from the time the claims come in the door to the time they're paid, it's all very reasonable, within industry standards." But we wanted to test that.

So they went in and as they walked through, they peeked out the back door and found piles of claims back there that hadn't come through the door. Those piles of claims sitting back there hadn't been stamped as received yet, because they knew that as soon as they got stamped as received, they would start flying through. The whole issue of unopened claims and the time from receipt to payment different from your claims triangles can be a big problem. Also, claims that have come in the door, the opened claims and the reversals and the recoveries, those are all issues that you can talk about.

These are just various types of things that you might find as a result of a data warehouse. Some of these are technical issues that you might not necessarily think about as actuaries, but they're the way that your data reporting systems are designed. There are a lot of different errors that can be introduced. Everything from unreasonable relationships, where you're trying to relate things to one another that aren't necessarily related, to nonstandard formats, in which you have five digits in one column, four digits in the next one, and they're not quite matching up because they're only matching up on the first four digits, to poor design of your data structures. Just be aware that there are lots of different issues in underlying data structures that can introduce many problems in your data if you're not aware of them.

Table 1
## KEY PERFORMANCE MEASURES

|  | **WORST** | **MEDIAN** | **BEST** |
|---|---|---|---|
| Claims Turnaround | 57 | 18 | 6 |
| % of Claims Reworked | 6% | 2% | 0.5% |
| EDI Claims PTC | 0% | 30% | 85% |
| Concerns / 1000 | 17 | 7 | 2 |
| Claims FTEs / 1000 Members | 0.25 | 0.4 | 0.6 |

This is just a handful out of about 80 or 90 that we typically tried to work with when we built data warehouses and reporting systems. It's good to have some idea of benchmarks for what to expect and what is a reasonable range.

These performance measures then can help you judge what is reasonable and at what point a material bias or a material problem in the data might be occurring, the claims turnaround time, or percentage of claims that are reworked. This goes to the issues that we were talking about—if you have claims payment folks that are trying to get a push through the holidays, especially around Christmas, or a lot of stuff is getting pended and not fully worked.

You can bring that rework percentage down if you've got a higher percentage of electronic data interchange (EDI) claims. If you can get that up above the 50, 60, or 70 percent range, then that's going to clean up your data quite a bit.

You also want to look at the complaint log that comes from customers. If the claims aren't getting paid right, and it's flowing through the doctors, sometimes that's going to reflect back in the customer service areas. If you start looking at the concerns per thousand or the complaints per thousand, particularly if they can categorize that into issues around your claims data, then you'll start to get another indicator of where your data is going.

Also, adjust your claims staffing levels. If the level is adequate, you're going to have fewer reworked claims. If the level is inadequate, the staff is overworked, and if there are a lot of pended claims, then that's going to introduce more errors as they try to work with the skeleton staffing and get through their staffing issues. If an insurance company has performance measures, it can use them to help judge staffing levels.

One of my favorite topics is the notion of membership. Membership is the denominator for your per thousands and your PMPMs, so it's very important. Surprisingly enough, most membership accounting areas struggle at some point with getting this right. It's not necessarily even the insurance company or the HMO's fault. A lot of times it's the nature of the relationship between the employer and the insurance company, and so there are two ways that they send the membership over. One is the negative tape and the other one is the positive tape, one that gives you the and the other one that gives you just the changes. A lot of times the human resources area in a company, particularly if it's a large company, may take months before it gets those life-change events recorded in its system and then passes them on to the insurance companies. For instance, if a new baby is added or if an employee terminates, it may be months before that change actually gets taken off the system and reported to the insurance company. This particularly becomes a problem when you're trying to look at member months in total.

As a quick example, if an insurer is gaining membership, one quick judgment that you can make about the accuracy about its membership data is simply to take the past 12 months and add the 12-month single monthly membership together. Each month you'll also have the running member month total. If the totals of the 12-

month and the final membership member month count are different, that percentage of difference is going to start to give you an indicator of how far off, how much retroactivity is processed at a given time.

Table 2

# Membership Example

| | Gaining Membership | | Declining Membership | |
|---|---|---|---|---|
| | Monthly Membership | Member Months | Monthly Membership | Member Months |
| January | 1,350,000 | 1,350,000 | 1,350,000 | 1,350,000 |
| February | 1,390,502 | 2,709,923 | 1,296,005 | 2,621,136 |
| ... | | | | |
| November | 1,814,330 | 17,260,149 | 897,559 | 12,184,786 |
| December | 1,868,768 | 19,128,918 | 861,665 | 13,046,450 |
| Sum of Montly Memb | 19,159,496 | | 13,071,320 | |
| Difference | 30,579 | | 24,870 | |
| | 1.6% | | 2.9% | |

Typically, in a situation where you're gaining membership, you're going to be putting new members on faster, because as soon as they come on, they're going to want to have a claim paid. When the claim hits, you're going to do a little research, and if the new members don't show up on the eligibility files, you're going to go back, research, find out the members really are eligible, retroactively instate them, and get them on to get the claim paid. On the other hand, if members are terminating, and particularly, if your membership is declining, it's going to be a longer time before you realize that they're off the rolls. You're going to keep paying the claims against that because they're still showing up as eligible even though they've terminated perhaps six to nine months ago. The lag time in a declining membership is typically worse than in a gaining membership.

An HMO that I worked for had significant problems in this area. It was paying claims as long as two or three years after the person terminated because it hadn't reconciled those accounts for so long.

Once those claims have occurred, there's virtually no way the HMO can go back

and recover them. The HMO had no premium coming in when it did the retroactivity, but it sometimes had large claims payments going out.

Capitation is something that we haven't covered yet. I know that capitation is declining somewhat, but it's still out there. One of the things to watch for is that capitation payments typically lead to underreporting by physicians. Although there may be incentives to try to get the claims in that were incurred, you may want to keep a wary eye out for potential underreporting in the capitation system. They don't look like adjudicated, and they're not adjudicated like the fee for service. Again, if they're not tied to the actual payment, then I would be suspect of the quality of the data that you're receiving. Some of the doctors will just check off 99213 for each healthy-person visit. If you look under the capitated claim, particularly in the staff model, you're going to get some doctors who lazily just mark those off to get the paperwork off their desk, because it really doesn't make any difference in how they're paid, and the claims data gets suspicious.

A couple of other items to look at are home-grown codes, as was mentioned before. HIPAA and the simplification rule should help that. Also, in creating data, sometimes you'll group data into various categories.

One potential area of introducing error is in how things are mapped from one system to another and how those transformations are made and then any groupings where you're grouping things together. You may be excluding or including things that may or may not reasonably fall within that group because an analyst someplace along the road either misunderstood or misapplied the logic that was set up in order to get things into those buckets. Some of those may be very subtle and very small items, but potentially they can add up systematically over time and create more significant issues.

There are a few other things to take into consideration, including financial audits. If the health plan reports Health Plan Employer Data and Information Set (HEDIS) data, they're required to go through a data audit. I don't know how many of you have seen this data audit, but it's fairly extensive and it covers a number of the issues that we've talked about today. They report the findings, and if you simply ask for the results of that audit, and if the HEDIS data is available, you can learn a lot about how the data is organized, what kind of problems they were seeing, and what kind of corrective action plans may be in place to correct that data. Those obviously apply to measures that come out of HEDIS, but if they're fully reporting HEDIS, they're fairly extensive and they include most utilization data.

The cost data is a little more iffy because HEDIS doesn't get into the cost issues although there are areas that would impact them.

Also, look at any internal audit corrective action lists that they might have, and provider membership complaints. You might want to look at their Department of

Insurance (DOI) complaints too, to see if they are significant and if any lawsuits are out there.

This case study is fairly simple, but it's designed to get the ball rolling and to start people talking and thinking. This is a case about a small HMO that had been asked to develop a cost model. In developing the cost model, it received three numbers for membership, and the numbers were different enough to cause a little more than $1 PMPM difference in the cost model. The sources of the data included membership from marketing, membership from finance, and membership from accounting. You've got those three numbers and you're not sure which one to pick. Let's open the discussion up about where you would go from there. How would you proceed? What would you do in that report?

**FROM THE FLOOR:** I don't think I'd rely on marketing.

**MR. WILLIAMS:** Material bias, right?

**FROM THE FLOOR:** At our company we have two sets of enrollment reporting on a monthly basis. They're reported and they're restated. Restated will go back six months. For an accurate study, we always use restated, which is delayed or a late cancellation. For example, in small groups, most cancellations are for nonpayment of premium, and you wait for the 30-day minimum grace period. After that, we're already two months behind, so we always go back six months restating. And we don't use marketing. They never report a cancellation, and they always report potential sales. You've got a plus 5 percent, plus another 5 percent. Finance is using a database. We don't even look at it. It's our actual database.

**MS. KRISTIN LYNN JOHNSON:** I'm with USAA Life Insurance Company. We had a similar problem, more on the life side than the health side, and it all stemmed back to definitions—exactly what do you count and when do you count it? Do you count it when you accept the premium, do you count it when the coverage actually starts, or do you end it when the grace period starts? As soon as we could all agree on definitions, we could all agree to have different counts, as long as they were within a tolerance of each other.

**MR. WILLIAMS:** That's good. The existence of a data dictionary that defines things specifically for various areas is an important element of controlling data quality, because that very issue of different definitions applied differently across departments is very common. It occurs virtually every place I've ever been at.

**FROM THE FLOOR:** What we do, and I'm not saying we have been very successful, is define the sales lapse and in-group changes, because that's the only part that we can nail down. You're going to get the people who move from small groups or, say, larger groups when people retire and move to Medicare supplement. Is that really sales, or are they moving from case to case? Are they

new sales, in-group changes, or just lapse and, therefore, for the kids over 19 or 20 who become individual policies, do you count them as sales or lapse?

This can cause a lot of controversy because the business unit usually is agreeing. That's where you have to be very careful.

**MR. WILLIAMS:** Any other responses to this one? I think what's important here is you have to figure out what it is that you need the data for, and as long as you're consistent, and especially if you're doing historical data and you're trying to pull it all up to a current date, you need that consistency. Then you can even do sensitivity testing. If I use the marketing membership, what information is that going to give me versus the accounting? It may not really matter as long as you don't mix accounting source for historical and marketing for current, or at least understand the differences.

**FROM THE FLOOR:** Okay. So for consistency you don't use membership and accounting's count once and then the finance department's for the next one.

**MR. WILLIAMS:** Let me give the conclusion to this case. This concerned the CEO so he pulled in his membership and accounting person, finance person, and salesperson into a meeting and asked them what the right number was. Of course, sales defended their count, finance defended theirs, and membership and accounting defended theirs. In a fit of frustration, the CEO opened the door and pulled the janitor in and asked, "What's the membership?" My job was to go in and figure out what the right membership number was so they could zero in on it for the reporting systems. It turned out that the janitor was the closest. The finance department had made all the adjustments for the retroactivity and nonperforming accounts and things like that. As you might have guessed, the finance department had the most accurate numbers that you could work with, and they've been pretty stable and steady over time.

**MR. VOLKMAR:** Let me take the second item there—backlog unknown due to a merger system—because we kind of hit the claims switch to a third party pretty well. I had an actual case in which, due to a new system, there was actually a separate claims adjudication unit to handle special problems. This new system obviously had a lot of problems so there were a lot of appeals and they had a large backlog that was not disclosed to their own finance department, nor to us. It was very material, and a year later, or six, nine months later, they disclosed it to us. The question is, what do you do the following year when you find out that your prior number was significantly understated, and how do you trust the most recent data?

**FROM THE FLOOR:** A couple of things come to mind. First, you want to take a look retrospectively at what happened, speak to the various people, see if there is still a lack of communication, and understand what they've learned from that. The

second thing is perhaps to the degree that you have comparable data, or you look at what was before and what was after and see what's coming in, see what patterns have emerged and get a sense where things are. Was it a blip that I can't trust the data from a year ago? This looks good. My senses tell me that, after speaking to everybody, everything seems okay and it passed some reasonable tests that you have—some relationships to various indicators, whether it's premiums or members or what not. That's the first thing that comes to mind.

**FROM THE FLOOR:** I think your problem in the very beginning is the testing. If you do your per thousand, the days, and the average PMPM, you can kind of get a feeling of where it is. Also, I find out when you have any doubt, don't talk to the VPs, talk to the supervisors. They know what's going on. They know what those problems are that the VP might not be aware of. At least every other month I talk to the claims supervisors and they'll give me the straight answer.

**MR. VOLKMAR:** That's a distinct advantage an in-house actuary has over a consultant.

**MR. DAVID A. SHEA JR.:** Trigon Blue Cross/Blue Shield in Richmond, Virginia. This looks like my case study.

**MR. VOLKMAR:** It's not yours.

**MR. SHEA:** I know. We had purchased a small insurance company that we have since relieved ourselves of, but the claims over a period of three years were on three different payment systems. It started out on theirs and then the decision was made to move it to ours, because it's better, faster, and smarter. We had more data elements, so you could look at the payment patterns and you knew exactly when it moved to ours because the paid claims for one month were zero. They were going from $4 million to $5 million a month to zero, and then the next month they were $20,000, and then it ran back up again. Then when we decided to relieve ourselves of the business, it moved to another company and we sped up payments and it went to zero again. You would think you could rely on some tried-and-true measures like trends and PMPM work, but in this particular instance it didn't work because the trends were higher than anyone had ever imagined they would be. Sometimes just a word of caution is that every time you look at something and every way you slice it, it looks reasonable, but the answer may be very different.

**MR. VOLKMAR:** The same thing happened to us. We looked at the PMPM and did not catch it. There was enough change going on. It looked reasonable, or at least we convinced ourselves that it was reasonable at the time. So it doesn't always work, especially if there's a change in membership. This particular company was getting out of some lines of business, so there were some other changes going on that kind of masked the missing data.

**FROM THE FLOOR:** One similar observation that I found is that there are times when things seem to be going normally, and then there are times when you get blips in the data. You get changes in patterns for reasons that you really don't know about. If that happens, there's no way you're going to know about it. You'll either be underreserved or overreserved, depending on your base reserve before any margin. You're going to be different because there's no way you could have anticipated that blip.

**FROM THE FLOOR:** I think another related point is a major change in provider contracts. Coming off a bunch of capitation, what do you assume for claims? It gets trickier dealing with different regions, which had different provider contracts and now have different provider contracts, so it's very hard to make any assumptions.

**FROM THE FLOOR:** I heard you say that the one point that seemed to be critical was this idea that there had been a system change in the complaint ratio. They had this special unit set up just to deal with complaints, and I'd say the first thing you have to find out is that the system is now a year older. Typically, the bugs are out of the system, so at that point the first assumption would be that the processing is reasonably normal. At this point I would go back and check to see if the special department is even in existence anymore. Then, in looking at your runout analysis, you know that you made a mistake. But to assume that the mistake is still there, and you have to be underreporting into the next period, probably isn't a good assumption without looking at a lot of other things and making sure that the same level of errors isn't still occurring..

**MR. VOLKMAR:** What we did in this case was a full claims audit. We had our specialty claims people come in and verify that the problem was pretty much caught up, solved, and we didn't expect another one. That gave us some confidence going forward that we had all the data. We did go back and get the missing claims, actually in a separate triangle, and they did have an unusual pattern because there were complaints and they tended to pay the older ones first because they had been sitting there. They had a real long pattern. Then they finally caught up. It would have changed our completion factors quite a bit if we had known about it; however, it probably wouldn't have been the correct long-term completion factor. So we may have had reasonable completion factors, but we had understated claims data. We did include it all and did our most recent year, and I think we're reasonable again. That's what we have to sign to.

One interesting sidelight is the auditors were asked to go back and restate the prior year. Now, as you might imagine, the auditors from the prior year were no longer there. They had been dismissed, and a new audit firm came in that was actually restating the prior year reserve, which is a major deal for an accounting firm. That's in process right now, and we gave them our restated numbers and so forth. It

made me much more cautious. Every time I get burned, I get much more cautious. You never can catch everything. Do you want to do one more?

**MR. CROOKS:** Yes. A company rolled out a new online management information system, which was going to give all the executives of this company the ability to drill down into what was going on in the company, help them manage it better. There was a big kickoff at which all the people who mattered were in attendance. It just so happened that the company in this example was in a city that we had just priced for purposes of PPO, and this was quite a while ago. The numbers that came up were erratically different. We were being presented with things like per-diem rates, mainly calculated off of the business, as well as some utilization statistics. In general, that's what we had to work with. How would you proceed, and what do you think were some of the possible causes of the difference?

**MR. SHEA:** I'm going to presume that there is one per-diem rate, that there's not a mixed bag of variations, like a per diem for newborns.

**MR. CROOKS:** To clarify, there were per diems presented for four different categories, med surg, neonatal, etc. This was in the late 1980s, when per-diem contracting was not all that common, so this was in fact the first time that this area was getting the per-diem contracts.

**FROM THE FLOOR:** There wouldn't be a mix issue. In other words, you would have one per-diem rate for that city and you would expect to see that when you looked at your data, no mix-issue differences. If there is a mix-issue difference, I would look into that. You ask the silly questions first, but sometimes they turn out to have answers. You ask if you pulled from the right data field, perhaps you balanced back to your general ledger to see if it matches something on a page in an incurred basis. That would be the next step. Then check your provider reimbursement system to make sure the correct per diems were loaded.

**MR. CROOKS:** In this case, we had a situation in which one person looking at it came up with one answer, and was pretty confident that was the correct answer. Another person, who would put together the management information system, had a totally different answer that didn't necessarily look like it made sense.

**FROM THE FLOOR:** Okay. Those are the things I would check first, the data source, the match between what was contracted versus what was being paid, and then balance to a system that you can depend on like your general ledger.

**FROM THE FLOOR:** Certainly, either the general ledger or whatever you were using before. I presume that I have something that was in use beforehand. Now they went to a new management system. The question is, "Why is it so different?" The first thing I would start with is looking for a problem in a specific area; I'd look first to see the totals, if the totals match the data that I have.

**MR. CROOKS:** The totals from both sides actually matched in this case. So they're the same data source, and if the totals are worked out—utilization times rates and everything else were the same.

**FROM THE FLOOR:** Okay, fine. So the next step is that I would break down the totals by the various subcategories, to see how one matches versus the other, and presumably there's some switches in there. I would take a look at that and see how that matches up. Is this one area or a couple areas giving a problem, or is everything giving a problem? Depending upon what you find, you may have to go down to the basic sources and start investigating what you have; nevertheless, it's easier if you have a total and, if you can find the sources and go back to the sources, find out what's mismatched. I guess with the degree of what you're finding out, I'd start big and then break it down and see where it starts resolving.

**FROM THE FLOOR:** I was operating under the assumption that there was something wrong with the numerator, but there could be something wrong with the denominator. The utilization count could easily be wrong. I know, for example, if you don't correct bills on hospitals, you'll double the number of days that you count. A lot of times on claims systems they will count the mother and the newborn as two separate days if you're not careful. That would be the other thing I would check—the utilization measure.

**FROM THE FLOOR:** Is there a possibility the medical dates—I'm assuming medical is medical, not including surgical, not maternity—may be understated in total?

**MR. CROOKS:** In fact, we basically stumbled across what the problem was—homegrown codes. In this case, there was a 799 diagnosis code, which is undiagnosed illness if you look it up, and we were amazed at the number of undiagnosed illnesses that were going in. This indigent population was just walking in off the street to get out of the cold, and that's how they coded those people. What was going on was they were getting so many of these 799 codes, especially at low dollars. Again, that was one assumption. It turned out that, at some point in their history, they decided to use that for in cases on the border where they were splitting off the mother and the baby from the code. So you had these $200-a-day babies, super-high frequency, three or four days at that time each, and it was skewing the medical per diem dramatically.

The lesson that I took away was the idea that you really want the person who is rolling your data up to understand what he's looking at. In this case, a systems guy who was doing all the work was putting together this really neat information system, but he had no clue what the data meant or what it should look like when it was rolled out. If he looked at how low that medical day rate was, although it wasn't terribly low, it should have clued him in to the problem. But because he had no background, he just rolled it up and presented it and said, "Here you go, guys,

here's your great system." I don't know if they ever implemented the system; because of the issue, the data was not reliable enough. If you don't have someone who understands what he's rolling up for you, then the answers you get back will be bad.

**MR. DRENNAN:** Our case study is about a company that was looking to develop benchmarks from its experience data. It had a bunch of standard metrics again. The resulting numbers appeared totally unreasonable when compared with industry benchmarks and with the management's. What do you look at?

**MR. ROWEN B. BELL:** I'm with Blue Cross/Blue Shield Association. We had some similar problems. All of the Blue companies report their utilization figures to us. What we've always found is that there are wide and superficially unexplainable variations between the dozens of different Blue companies on the cost per day, admissions per thousands, but when you multiply everything together, it all makes sense and is reasonably consistent. It's just that each plan has a different way of counting things. The correlation is fine. Each plan is consistently reported, it's just that you compare state A to state B and it just doesn't make any sense. People count things differently. Every plan has its own way of doing things and I would think the same thing here. If you multiplied the right things together, you'd come up with something close to your industry benchmark, and it's just a decomposition issue.

**FROM THE FLOOR:** I used to report that stuff to you and you don't want to know how it was done.

**FROM THE FLOOR:** I work for a Blue plan and we know how it's done. We count them correctly. Again, utilization statistics are tough to compile. A claim payment system is not set up to do that, so you have to create it. We create an outpatient visit by doing the same provider, same patient, same day, and that creates visits for us. They look reasonably accurate. I'd start looking at the utilization counts again. Admissions with the interim bills can be double what they normally are. Extended stays for admissions, they count them every month and the person's continuously in the hospital. Again, I think the culprit would be the utilization counts. I'm sure when you multiply everything together, it balances to a general ledger, your income statements, or whatever. Again, look at those numbers.

**MR. VOLKMAR:** We probably have time for one more before we hear the real story.

**FROM THE FLOOR:** It's a possibility it included the Medicare supplemental data in there.

**MR. DRENNAN:** Okay. How about the real answer? It basically was that the utilization was being recorded and counted. They had days fields that the system

carried, but as you went down and started investigating claim by claim, it didn't make any sense. You had outpatient claims with multiple days and you had inpatient claims with no days. So what it came down to was that the system and how it was set up and how it was generating that field were just flawed. It was classic garbage in, garbage out. Someone set up a system to determine what an admission is as opposed to what a day is. You would think a day is a little easier, but they had problems with it, and so when they rolled it all up, the base was wrong and so the numbers looked out of place.