Session 114PD
Update on Risk Adjustment Methodologies

Moderator: NEELA K. RANADE
Panelists: RICHARD AVERILL†
            JOHN MULDOON‡
            NEELA K. RANADE
            ALAN Y. WEINER
Recorder: ALAN Y. WEINER

Summary: The search for the perfect risk adjustment method has intensified as the U.S. health care system increasingly moves to the use of capitation arrangements. This panel provides an overview of prevalent risk adjustment methods and focuses on a new method being developed with the help of a grant from the National Institute of Standards and Technology. Differences and similarities between the new method and methods in current use are reviewed.

Ms. Neela K. Ranade: Using risk adjustment, our managed care systems can be compared more fairly for the quality they are providing. What I want to talk to you about is the historical risk adjustment methods that have been used by both government organizations and private sector organizations. Let's start first with the government organizations. The Medicare Adjusting Average Per Capita Cost (AAPCC) method is perhaps the best known. That has been, traditionally, adjusted for age, gender, Medicaid eligibility status, institutional status, and geographic location. I will say more about this later because it has not been a perfect method and has resulted in imbalances. Medicare also carves out those enrollees who are in end-stage renal dialysis and provides special higher payment for them.

---

†Mr. Averill, not a member of the sponsoring organizations, is Research Manager, Lincoln Management Institute, of Lincoln Financial Group in Fort Wayne, IN.

‡Mr. Muldoon, not a member of the sponsoring organizations, is Vice President-Classified Research at the National Association of Childrens' Hospitals & Related Institutions in Alexandria, VA.

Note: The charts referred to in the text can be found at the end of the manuscript.

In addition to Medicare, Medicaid on a state-by-state basis uses different risk-adjusted capitation methods. They have been based on age, gender, aged status, disabled status, and sometimes on activities of daily living such as bathing, eating, transferring, dressing, using the toilet, and walking that are taken into account in nursing home care or long-term care payments. Medicaid also has carve-outs for supplementary security income or special need populations. States also use risk adjustment by having commercial risk pools. If two or more insurers have turned down an enrollee, the enrollee can get health insurance through the commercial risk pool.

Examples of negative risk adjustment methods would be the limits imposed on underwriting, rating factors, etc., by certain regulatory organizations, especially for small group commercial markets.

When we come to private sector payouts, providers and insurance organizations use and have historically used other risk adjustment methods. Typically, we have rating factors such as age, gender, geographic area, income level, occupation, and type of business. The SOA did a very interesting study that produced a risk adjustment method monograph in 1996. They looked at some of the existing and evolving methods. Alan's going to talk more about them. What they found is that the current methods are good predictors for groups greater than 1,000 lives. But there's a 10% prediction error for groups of 500 lives. The prediction accuracy gets lower as the group gets smaller. Of course, the risk can be adjusted by reinsuring a block of business or by underwriting practices where enrollees may be categorized into preferred, standard, or substandard risk.

The whole area has gotten a large impetus because of a recent Health Care Financing Administration (HCFA) mandate. HCFA announced in the Balanced Budget Act of 1997 that starting on January 1, 2000, they will change the current AAPCC method to a diagnosis-related, risk-adjusted AAPCC for HMO capitation. The current AAPCC method hasn't worked very well. It has caused geographic imbalances, overpaid some HMOs and underpaid others, and affected the distribution of HMOs by geographic area.

Besides HCFA implementing this system, several states are in various stages of implementing risk adjustment methods for Medicaid capitation purposes. Various commercial organizations are closely looking at this subject. If you're an actuary who is involved in capitations, whether you're on the payer side or whether you're on the provider side, it's going to be very important to have knowledge of the risk adjustment methods that are available and the nuances of those methods in order to control profit margins.

Alan is going to talk about some of these risk adjustment methods. Then we are going to ask Rich to talk about the 3M Health Information Systems Project.

Mr. Alan Y. Weiner: When the SOA did their study, they compared several risk adjustment methods. Since the study has come out, there have been some losers and some winners in the field. Although not available at the time of the study, the 3M Health Care Product is probably the state-of-the-art item. Rich will introduce that later.

What has happened in the ongoing risk adjustment debate? One of the criterion for a good risk adjustment method is the predictive accuracy. Few risk adjusters are very good predictors. The level of R-squared is used as the basis of determining predictability. The procedure to determine the R-squared is to take one year's claims and run it through the risk adjuster. It's almost like an age-factor-type calculation where you determine the average cost per risk adjustment category. If you take a look at next year's claims for those individuals, you will find that the risk adjusters can predict 6–8% of actual claims for any one individual—maybe the top 10% if they start cutting out different parts of the pie. For example, extremely large claims are cut from the pie. Almost all risk adjusters have various methods of adjusting their database so that the risk adjuster comes out with a little higher R-squared.

The predictive accuracy is not great. However, as Neela pointed out, the traditional risk adjusters (i.e., age, sex, gender, industry, etc.), disability, and nondisability status predict maybe 1%. Although these risk adjusters are not the best, they are better than no risk adjustment at all. What's interesting is that the newer risk adjustment methods are getting better and better.

As far as any clinical credibility and usefulness of risk adjusters, I'm not sure we're at that point yet because of the predictive accuracy. But it is a way of grouping individuals into risk classifications and, ultimately, into disease classifications so that more accurate clinical studies can be made.

The risk adjuster should be objective. This implies that you can't game the system. That means if you add an extra diagnosis code or ten extra procedure codes to that patient when you're treating them, you'll get a higher risk factor with very little additional cost, thereby getting a larger capitation payment the following year. All risk adjusters are diagnostic-coding related or based. Some are based on diagnostic coding and cost. Some are based on diagnostic coding, procedure codes, and cost.

The risk adjuster should be easy to use. Medicare will have all the health care plans send in their data. They will calculate the risk adjuster and send back the

differential in payment. Of course, plans can get the risk adjuster and put it on their own system and do their own runs just to test it. The Medicare risk adjuster, which is called the Personal Injury Protection/Diagnostic Cost Group, will be available from HCFA at no cost.

A risk adjuster has to use existing data. HCFA has to have the data to calculate the risk adjuster. Initially, federal legislation has said, "We're only going to use inpatient data." You may have been reading the press and all the arguments stating that if HCFA is only using inpatient data, HMOs should place more patients into the hospital. That way the HMO will get a much higher capitation. Hospitalizing patients is not easy or inexpensive and may be subject to fraud and/or abuse statutes, so it's not going to be that easy to game HCFA, but the HCFA risk adjuster will work relatively well given the low R-squared.

The 1996 SOA monograph *A Comparative Analysis of Methods of Health Risk Assessment* examined specific risk adjusters. The monograph looks retrospectively and prospectively at eight models. It also looks at some state models and other risk adjustment models. We did not choose the state models because we only have a limited amount of time. But we did take the major national models for which we have updated information.

The problems that the Society found were as follows. High utilizers in year one were low utilizers in year two, and vice versa. Some of this can be adjusted. For example, if you have a maternity case this year, obviously the chance of somebody having a maternity case next year is not as significant, so you can adjust for that individual. Of course, the opposite isn't true. If you're a healthy individual this year, it's almost impossible to predict that you will have a maternity claim next year. They did find that if you go into the hospital from cancer this year, you're going to have claims next year, but not as many claims.

The next problem mentioned in the monograph was that no method worked adequately for patients with very high cost (such as the cancer patient), multiple hospitalizations, or multiple diagnoses. Some of the new systems adjust for this; some don't. They found that 1% of the patients accounted for 25% of the costs, but trying to figure out who was that 1% was not always very accurate.

They stated that an ideal R-squared target was 20%. If that level could be achieved, that model could be of some use. They recommended that if you could achieve that level, then risk adjustment would be viable. The object, again, is that providers should be paid for the patients they assume. Providers will not cover and HMOs have not been sufficiently covering the high utilizers because they're not really being paid for them.

Let's review some current risk adjustment methods. All-patient refined diagnostic-related groups (APR-DRGs). This is just a refinement on DRGs. Medicare pays hospitals on the basis of DRGs with some modifications. Another national system is Mediqual. Adjusted clinical groups (ACGs) is another method. ACGs used to be called adjusted costs groups. All of these are evolving as we sit here.

The DCGs can either be hierarchical condition categories-diagnostic cost groups (HCCs-DCGs) or HCFA principal inpatient diagnostic cost groups (PIP-DCGs), the risk adjuster chosen by Medicare.

The monograph also studied health self-appraisals. There has been very little advancement in self-appraisals. Last, the newest risk adjusters are from Mediqual and 3M. The products from 3M are APR-DRGs and prospective risk adjustment classes (PRACs), which will be discussed by Rich.

APR-DRGs is an inpatient-only retrospective (not prospective) model. This method categorizes DRGs into four basic severity levels: none/minor, moderate, major, or extreme. It also calculates mortality risk. This could be very useful if you were trying to do a mortality table and answering questions such as, "If a patient had cancer of moderate severity, what would be the probability that he or she would die next year?" A few thousand of these APR-DRG systems have been sold to hospitals. Since this is an inpatient risk adjuster, that's the primary market for this. The severity relates to the extent of physiologic decompensation or organ system loss of function experienced by the patient. They'll actually look at what procedures were done, their severity, and whether surgery was done or not done. Each APR-DRG has its own descriptors that decide what level of severity you're in.

The Mediqual system is another retrospective model. By retrospective, I mean that it's really designed as a good predictive (prospective) model. The design doesn't try to estimate next year's cost given a condition or illness this year. What retrospective says is that if you had cancer this year and we say the average cancer costs $14,000 this year, then we go back and run the same database again for the same year comparing the $14,000 to each individual's claims to calculate an R-squared. Given that you're using this year's individual claims and this year's averages, you're going to get a pretty high R-squared.

The Mediqual risk adjuster uses 115 severity clusters, each with 10 levels of severity. For all risk adjusters, you're taking an illness and you're trying to either combine it with other illnesses or split that illness into severity levels. You can go one of two directions. Of course, the more refined it gets, the larger the system, the more complexities, the longer it takes to run. Mediqual uses 69 classifications of International Classification of Diseases-9th Revised (ICD-9s) and 48 selected

procedures to determine the patient's cluster payment. It's a comparative tool and there are retrospective R-squared of 53.5% for total charges, 56.8% for ancillary charges, and 45.9% for length of stay.. An R-squared of 100% or 1.00 implies a perfect relationship. If you minus one, it's the complete opposite relationship. And zero means there's no correlation at all. Mediqual uses total charges, ancillary charges, length of stay, etc. But, again, this retrospective model would not be a model you would want to use to calculate a prospective capitation payment for the next year.

Johns Hopkins developed ambulatory care groups (ACGs). They use something called adjusted diagnostic groups (ADGs), which map into ACGs. Depending on the number of ADGs you have, an individual would be placed into 1 of 106 mutually exclusive ACGs. For example, ADG 32 is a malignancy and maps into ACG 28, which is an acute, likely to recur major illness. It's kind of a broad classification. Now, ACGs only use the diagnostic code. They don't use any cost information. They don't use any procedure codes. This one may be easier to upcode or to add more ADGs or diagnoses to the patient and, thereby getting a higher payment. A single ACG is assigned to a person based on his or her age, gender, and constellation of diagnosis code. They also do some adjustment for birth status and weight if they are recorded on inpatient claims.

More than 150 installations of the Hopkins risk adjuster have now occurred, including HMOs. The state of Maryland is using it for their Medicaid waiver program. Sixty academic organizations in five countries have used it. Depending on what claims you exclude in your pool, you'll get a prospective R-squared of about 6–8%. The monograph excluded claims above $50,000 when they calculated their R-squared.

HCC-DCGs had an R-squared of 9%; PIP-DCGs about 6%. Boston University and Brandeis University devised DCGs under a grant from HCFA. Medicare has taken the PIP-DCGs in coordination with an AAPCC base payment using the 24-age gender groups that Medicare currently uses for capitation payments. Some of the AAPCC eligibility classifications were eliminated, e.g., institutional versus noninstitutional. To the base payment, Medicare adds a payment for an individual's PIP-DCG. Next, Medicare adds a payment on the individual's original Medicare eligibility. Then Medicare adds for Medicaid eligibility, whether the patient is Medicaid-Medicare aged or Medicaid-Medicare disabled. At present, Medicare has used only ten PIP-DCG codes, but this may change in the final regulations.

The national average cost for a Medicare patient is $5,300. The Medicare base payment group includes new enrollees, which includes those who weren't in an inpatient stay during the previous year or those who had an inpatient stay but the

stay was for a day or less. Seventy-five of the inpatient diagnoses were included in the base group because they represented only a minor or transitory disease or a disorder that's not clinically likely to recur, will not result in significant future medical costs, is rarely the main cause of an inpatient stay, or is vague and ambiguous. Medicare risk-adjusts only about two-thirds of the inpatient stays.

The new Medicare status groups do not include the institutional status anymore—those patients who are confined in an institution for long-term care. Medicare Parts A and B will now be combined. In the old AAPCC capitation method, they were separate. The inpatient diagnostic payment uses the highest cost single PIP-DCG. Since hospitalized patients are usually sicker, HMOs will receive greater payment for these individuals under the new system. The new AAPCC will prevent a little of the current positive selection.

HCFA also allows a special adjustment for chemotherapy because diagnostic coding doesn't really give enough information to categorize some of these patients. As a general example, an HMO will get an additional payment of $7,200 for PIP-DCG number 7, which includes bacterial pneumonia, dementia, or kidney infection. If you have all three you're still in PIP-DCG-7. It's probably pretty hard to have all three.

For a complete example let us take a male aged 82. The base payment is $5,617. His original Medicare eligibility was because of a disability. The HMO would get an additional (these are all additive) $2,381. He is not Medicaid eligible, so there would be no payment for that. Again, you could be Medicaid aged or Medicaid disabled and you would get a payment. The beneficiary was hospitalized for asthma and lung cancer. The highest PIP-DCG is 18 for lung cancer, and the HMO gets $12,883 additional. When you add them all together, the HMO gets $20,881 for treating this patient with lung cancer. Now you may think that's not enough. Probably some doctors may not think that's enough. But it's a lot better than receiving the $5,300 average under the old system.

HCFA says we're going to go up with these because this is better than nothing. We have all the data. We've been collecting it. Providing HCFA doesn't postpone this like it postpones a lot of things, it will go up January 1, 2000.

What is the relative risk score for that cancer patient? Considering the average is $5,300, they rate that patient as a 3.9. It's the $20,881 divided by $5,300. A 3.9 patient will have about four times the cost of a normal patient. Finally, HCFA adjusts for the 'gypsy' factors, more properly geographic practice cost Indices (GPCI). There you get your capitation. That's what they're going to do. It's going

to be hard for HMOs to project this. I'm sure this is another actuarial full employment act.

The last type of risk adjuster the monograph examined was health self-appraisals. Health self-appraisals are no longer thought of because they're too subjective and they really don't predict very much of the claims because people can't estimate what they're going to do next year.

The future of risk adjustment is refinements, which include more sophisticated risk adjusters, the addition of family histories, disease progression and transition matrices, complexity of care, treatment goals, alternative treatment protocols, increased use of global capitation for the special needs population, and disease management. Ultimately, it will be based on outcomes and, perhaps, episodes of care.

Mr. Richard Averill: In health care, what's been going on over the past couple of decades? Well, over the years we've been aggregating our payments together in bigger and bigger units. If you go back 15 or 20 years ago in Medicare, they paid cost. Whatever the health care provider spent was what was paid. Then we said we'd pay for items and set a price for that instead of paying cost. Then we said, "Let's aggregate further. Let's pay per diems. We'll just pay for so much per day for hospital care." Then we talked about doing per case. That's the whole DRG system. Then we started talking about paying for episodes. And then capitation rates. Now some plans are cutting deals with providers to just pay percent of premium.

What this has been all about is nothing more than shifting risk from the payer to the provider. This whole evolution in health care has been simply about who is going to take on the risk. Is it going to be the payer or is it going to be the provider? What we've been doing is shifting risk from payer to provider.

To give you some sense of how dollars are distributed, Chart 1 shows Medicare data from 1992. As you can see, 73.5% of the enrollees consumed only 7.6% of costs. Conversely, 9.8% of enrollees consumed 68.4% of Medicare expenditures. If you're an HMO, and you're going to take on Medicare patients on a capitated basis, clearly what you want is that 73.5% of the patients who only consume 7.6% of the dollars to enroll in your plan. That's how you're going to make money. And you want to avoid, like the plague, the 9.8% of the enrollees who are consuming over two-thirds of the dollars. So, there's this huge concentration of expenditures in relatively few patients.

We took on this project because, as Alan alluded to, current capitated risk adjustment methodologies are relatively ineffective for predicting future health care resource use. They are relatively inadequate in dealing with the multiple comorbid patients. The patients with many diseases at high severity levels are the 10% that consume two-thirds of the dollars. Because of this enormous financial incentive for selective enrollment and disenrollment, unless we had a capitated risk adjustment system capitated, payments will not be a meaningful way of controlling health care resources.

Now, if one takes on a project of how to do a capitated risk adjuster, there are certain significant challenges. Clearly, the biggest one is this enormous concentration of expenditures in relatively few patients. You have to clearly deal with that. The majority of patients have low expenditures, so you clearly have to identify those as well. It also means that you have to have databases, both operationally and from a research perspective, that are longitudinal. You want to be able to look at three to five years' worth of data and have all your inpatient data, outpatient data, hospice data, etc., all linked together so you have a complete picture of the patient.

You also want to do this in a clinically meaningful way. Not only did we want to do a good job at predicting future health care resources, but we also wanted to do that in a clinically meaningful way. So this became a management tool. The entity receiving the payment can talk to their physicians and say, "This is what we're being paid for this kind of patient. How did we and how do we treat this kind of patient?" Then what we would have is a dialogue with the physician.

Now, as Neela alluded to, this is a project that was funded by the Department of Commerce's National Institute of Standards and Technology (NIST). It was a 45-month project. The project ends December 1998. It's a multimillion-dollar project. They encourage the private sector to take on certain high-risk projects that would make the U.S. more competitive in the world market. They feel that our health care costs are too high relative to the many other countries in the world and want projects to be done that would ultimately improve our competitiveness in the world market. The participants were 3M Health Information Systems, Actuarial Science Associates (ASA), the National Association of Children's Health Related Institutions, and a wide range of clinical consultants from across the country.

The system provides PRACs. Each PRAC is clinically meaningful and provides the basis for the prediction of future health care utilization and cost. You take your historical data and each patient is uniquely assigned to a clinically meaningful, mutually exclusive class.

This is a clinical model done by physicians, first and foremost. The clinical panels asked the question, "What affects future health care resource?" But it's also a clinical model that has been tested extensively and verified with historical data. For example, there would be a separate clinical model for a diabetic and a different clinical model for someone with congestive heart failure (CHF) but a different clinical model for someone who has both diabetes and CHF. Each of those base clinical models is subdivided into severity levels. So, when we talk about a diabetic we can have a severity level one diabetic versus a severity level four diabetic.

While we looked at the historical data and fed it back to the physicians, the final decisions we made were always clinical. The PRACs are not based on a traditional kind of regression model. It's a clinically driven model tested with data, so predictions are not dependent on coefficients derived from a development database as you would have in a regression model.

We fed the data results back to the physicians. They looked at the results. We would get a dialog going between the data and the physicians. The physicians modified the clinical model based on what the data results were. It was a highly iterative process that we went through. We only used standard data elements that were routinely collected on claim forms. Diagnosis, procedures, age, sex, date of service, site of service, and the type of provider were the data elements used to assign PRACs.

We had an extensive amount of data that we used to develop this system. We had a 5% sample of Medicare beneficiaries from 1991 to 1994. There were approximately 1.3 million beneficiaries in the final analysis database. We also had data on a working population from 1992 to 1995. That was an employee population with their dependents which had approximately 300,000 enrollees in it. We then had Medicaid data from the state of Washington for fiscal year 1992 through fiscal year 1993. That had approximately 250,000 beneficiaries. Clearly, we had some extensive data available to us.

I'm going to start off talking about the analysis of the Medicare data. We have four years, 1991–94, of enrollee data: inpatient, outpatient, skilled nursing facility data, hospice data, home health care data, and physician and supplier data. All of these were linked together around each enrollee for four years. We also had the enrollment database available to us from HCFA so we could determine eligibility and understand when the patients became eligible for Medicare benefits and when they ceased to be eligible for Medicare benefits linked across the four years. We started off with 1.8 million enrollees. Then we did some data editing.

We required full eligibility for at least the first two years because that was going to be what we used to develop the initial model. We excluded HMO enrollees and enrollees with other primary payers because we wouldn't have all the expenditures for those patients or they were still employed even though they were over 65 and Medicare may only be paying secondary. We also excluded beneficiaries who were long-term institutionalized. That was because they could run out of their Medicare coverage and not have all the dollars associated with them.   This data editing reduced our database to 1.33 million beneficiaries.

We then used the patient characteristics from year one and two to predict year three expenditures. The one twist that we had to do was if the patient died during year three, we used our estimated cost in year three as whatever that patient expended during the 12 months prior to death. We've reserved year four for final in-process validations.

We begin by assigning each patient to one of seven PRAC statuses, as we call it. The PRAC status is just a general categorization of the patient's clinical condition. The first is catastrophic conditions. That's usually associated with patients who are dependent on some kind of technology such as a dialysis patient who is dependent on the dialysis technology in order to stay alive. In the next step, catastrophic conditions are subdivided into four severity levels. Second, we have dominant or metastatic malignancies with four severity levels. Third, patients with three or more dominant chronic diseases. These are subdivided into six severity levels. Fourth, we have patients with two or more significant chronic diseases. These are also ultimately subdivided into six severity levels. Then we have patients with single organ system chronic disease, like diabetes. Those are subdivided into four severity levels. Then we have patients with no chronic diseases, but some history of some major acute disease. For example, they can have pneumonia twice in the past year but no underlying chronic disease. That's really a different patient than pneumonia with a chronic condition. Finally, we have well patients. The last two categories have no severity levels.

For our Medicare database, the 1.33 million breaks down for the seven categories into respectively 0.6%, 3.4%, 5.1%, 42.3%, 31.3%, 1.5%, and 15.8%. The charges vary from $64,000 for an average patient with any catastrophic condition all the way down to $2,700 per year for well patients. That's a startling level of difference across the seven PRAC statuses.

How many PRACs are within each of the seven statuses? Obviously, there's only one well group. We have 3 moderate acute groups, 174 different single chronic groups, 66 multiple significant chronic groups, 27 with 3 dominant groups, 21

dominant or metastatic malignancies, and 9 catastrophic groups. With severity levels, these total up to about 1,300 separate individual groups.

As I said, most categories are divided into severity levels. This was done to be clinically meaningful. We wanted to describe the extent and progression of the disease. And, of course, we wanted the severity level to differentiate future levels of resources. For instance, if the patient has plural affusion with (CHF), he or she will be a level four CHF. If the patient has been hospitalized for respiratory failure within the past few years, he or she will be a level four. If the patient has been hospitalized within the past six months with valvular disorders, he or she will be a level four. Since it is extremely detailed, a clinician could look at it and say, "Yeah, I understand what makes this patient a level four CHF patient."

There are huge differences in the level of future expenditures by severity level. Let's look at the costs of prostate malignancies. For metastatic patients, that's going to vary across the four severity levels from $7,000 all the way up to $35,000, so there is a fivefold difference across the 4 severity levels in terms of next year's expenditures. Patients with prostrate malignancies without metastatic disease are going to vary from $5,000 up to $14,000. For this diagnosis, we actually had no level fours in our data.

Expenditures also vary significantly when multiple diagnoses are used. Take diabetes, chronic obstructive pulmonary disease (COPD), and CHF. When patients just have diabetes across the four severity levels, the cost for next year's expenditures varies from $5,000 up to nearly $9,000. For COPDs, the cost varies from $6,000 up to $16,000. If you have pairs of those diseases, diabetes and COPD across our 6 severity levels, costs vary from $7,000 up to $24,000. For COPD and CHF, from $6,000 up to $31,000. Now, if you have all 3 of those across the 6 severity levels, it's going to vary from around $12,000 all the way up to $50,000 per year. This gives you a sense of how detailed this becomes.

So far, we've made very limited use of age and sex. Sometimes in the clinical model we use it to differentiate some diseases. A pediatric asthma patient is a totally different patient than an adult asthma patient. It's basically a different disease.

How would one use the model? From the historical data you would assign the PRAC. There would be a set of payment weights associated with each of the PRAC levels, or you can compute your own payment weights. There would be some multiplier to convert the payment weight to dollars. Then, one would pay per month one-twelfth of that capitated rate on an ongoing basis.

What we did was compute an R-squared to give us a measure of how well our prediction compared to the actual expenditures. This R-squared is per patient. It will be lower than what you're used to because we're trying to predict the actual expenditures on an individual patient-by-patient basis, not to predict the expenditures for the group as a whole.

We tried various alternatives. We only used the inpatient data. We put a 6-month lag in the data and only used the first 18 months to predict the third year. We ignored the death adjustment in some of the runs, and we put a cap of $100,000 on the patient's expenditures or $150,000, $300,000, and $500,000. We tried 15 variations in all. In some cases we used charges and some we tried to predict actual payments that Medicare made. The point is that all the R-squareds were typically in the teens, but we're getting much higher R-squareds than any of the other systems that we've seen out there. Indeed, if you're using 2 year's worth of data with a relatively minor, say, $100,000 type of cap, you're up to nearly 20% in terms of your R-squared.

We repeated the same analysis we did for Medicare for the working population data. That was a standard fee-for-service benefits type of database ages 0-64, and we required at least 3 years worth of continuous coverage unless the last year was shortened by death or they were born in the first year. Those patients were kept in.

The difference for the seven statuses wasn't quite as great as Medicare. Costs vary from $48,000 from catastrophic down to $2,300 for the well group. The seven statuses had the same relative progression that we saw in the Medicare database. Our R-squared for the working population was 14.88%, so that gives you a sense of where we're winding up.

Mr. Robert D. Shapiro: How do you account for the normal turnover in the commercial population?

Mr. Averill: We required that the person be eligible for all three years in this database in order to be included in the study. Since you pay on a monthly basis, if someone disenrolls three months into the year, you would have only paid for the first three months. Disenrollment is automatically accounted for when it occurs. What about a situation where someone enrolls and we don't have any historical data? One of the benefits of having a clinical-based system, and one of the things we plan to do, is to develop a questionnaire that can be filled out and run through the model to give someone an initial PRAC, even if they don't have any historical data. This questionnaire can be used to assign the patient to a PRAC.

Mr. Jon D. Harris-Shapiro: How did you account for inflation?

Mr. Averill: In any prediction, how good is your inflation factor adjustment and what do you want that inflation factor to do? We basically made an inflation factor that kept it budget-neutral in the third year. We're trying to look at how we predicted each individual patient on a budget-neutral basis.

From the Floor: Would it be appropriate to factor out the inflation adjustment?

Mr. Averill: All the risk adjusters have an inflation factor. They all try to factor out the things that are not diagnosis-related. Inflation is one of those.

Mr. Michael Jay Sipos: Have you tested the system on any other databases?

Mr. Averill: We've reserved year four data for final validation, and we will be doing that during the course of the next month. We'll start predicting year four in our database since that will allow us to do a six-month lag, a year lag, a year-and-a-half lag, and so on. We'll present this in our final report to NIST, which will probably be available at the end of January or the beginning of February 1999. We will do every conceivable combination of simulation that you can imagine in terms of lags, what data is used, reinsurance caps, and other variations. We'll probably wind up doing several hundred different variations of this. So, people can take a look at that.

From the Floor: I didn't understand the justification for your adjustment upon death.

Mr. Averill: We did that because we were trying to get our dependent variable for year three. In that last month of death, they're going to have extraordinarily high expenditures. If you just multiplied that by 12, you'd get a very artificial number for what the estimated expenditures for year 3 would be, so we thought a more accurate number would just be whatever the expenditures were 12 months prior to death.

In the prediction models, the R-squared models that we looked at, we also tried one model where we just simply said, "We'll pay them the capitated rate even if they die during the course of the year." Doing that didn't have that much of an effect on R-squared. We tried different ways of estimating what the annual expenditures would be for someone who dies in the prospective year.

Mr. John Muldoon: We used the same logic as Rich but focused on the pediatric population and the Medicaid population, including Medicaid adults. There are two reasons for that. One is from a classification perspective. It was always the intent to

integrate the pediatric component into an adult system. A lot of pediatric conditions continue on into adulthood, so you have to identify that progression.

The other reason is a very practical one. Most Medicaid programs are most concerned about the adult population and emphasize, in particular, the disabled population. To the extent we developed a system that classifies those individuals well and helps the Medicaid agency solve for that problem, they're more inclined to use a system that will also solve pediatric issues.

Let me comment on some of the special challenges of pediatrics. First, of course, you have a lot of different conditions than in adulthood. The second may be not quite as obvious. You have a lot of different disease progressions. Some are permanent. Some are progressive. But quite a few are curative. Some of the anomalies require a lot of surgical intervention in the early years, but you get over it. So, the base here may not predict the future year health service utilization level. Then you have other conditions that are variable and those that tend to attenuate. You have to take all of those into account because you're trying to predict the future. It's not slowly progressive. It's every which way.

You have a lot of low-volume conditions. In fact, most pediatric chronic conditions are low volume. You don't have so many multiple conditions. When you have multiple things going on, most of the time there's a primary chronic condition, and others are complications and manifestations.

We had a clinical committee and consultant structure and tapped into all of the pediatric division chiefs at two children's hospitals. We actually had a physician from each hospital 25% of the time working with us.

The model for the development database we tested was a two-year Medicaid database from the state of Washington. This was a zero to 64-year-old population, noninstitutionalized. We include only those who were not Medicaid duly enrolled and not HMO enrollees. That's just so that you have the full claim history.

Two-thirds of the Medicaid population are children. Of those, only a very small percentage, 2.4%, is Social Security Income Supplement (SSI) disabled and that shouldn't surprise you. Those are mostly children with certain neurological, mental-retardation-type conditions that are the child's equivalent of being unable to engage in substantively gainful employment. Of the adult population, about one-third is SSI-disabled, or 22.1% of the child and adult population.

We classified the population into the seven PRAC categories—1.1% catastrophic, 0.1%, 0.1%, 2.3%, 14.2%, 5.5%, and 76.6% for the well. The 23.4% of the nonwell population accounted for 51% of the dollars.

You can break this out separately by kids and adults, by non-SSI and SSI, and that shows you more of the kind of classification system to understand who the enrollees are. The chronic disease prevalence rates will be a lot different for kids. We haven't done this recently, but some previous runs would probably show you about 13–14% chronic disease prevalence versus about 25% for adults. As to the adults, not surprisingly, the older adults and the SSI-disabled will have a higher chronic disease prevalence rate.

There are also many misconceptions about kids being less expensive than adults. On the whole, that's very true. But that's because most kids are healthy. But if you identify those who are chronically ill and those who have multiple conditions, you're able to break out the cost and it's pretty comparable to the adult population.

This is showing that as we've broken out among those considered catastrophic, it's really a subgroup of lifelong defining childhood chronic disease conditions. These are cerebral palsy and cystic fibrosis (CF)-type conditions. The average charge is $12,000 for lifelong childhood diseased compared to $36,000 for the general catastrophic category, which includes more technology-dependent kinds of conditions. That is because in the lifelong conditions you have the whole spectrum in the disease progression.

What are the most prevalent chronic illnesses? Asthma, diabetes, epilepsy, hypertension, and migraines are the top five physical conditions; schizophrenia, major depression, alcohol dependence, depression/dysthymia, and attention deficit are the top five mental health conditions. It may surprise you that in Medicaid there are as many chronic mental health substance-abuse problems as there are chronic physical conditions. These are the prevalences for those classified in single chronic condition categories. When you count those in pairs and triplets, schizophrenia goes up. There are also many dual cases of alcohol dependence and drug dependence.

Let's discuss how well the severity level system predicts the cost of some of these lifelong defining chronic child health conditions. Take CF, a progressive disease that leads to death. They haven't discovered the cure yet for the genetic defect. A CF patient loses 2–4% of lung functionality per year. Early on, they go through what is called a honeymoon stage. But that 2–4% a year adds up. At a certain point, you begin to hit the pulmonary problems. That's what we pick up at level two. You first begin to see the pneumonia, the asthma, and the chronic sinusitis. At

level three, you have multiple hospitalizations, pneumonia, and home infusion therapy. Level four is the end-stage disease when you get into the pulmonary hypertension and cirrhosis of liver. It's a clinically defined disease progression and the levels pretty well predict the kind of expenditures you can expect to see.

Congenital quadriplegia is the more severe form of cerebral palsy. It is an awful condition to have. You might be surprised to see that the costs are not uniform and the severity levels do make a tremendous distinction. When we're talking about level two, we're talking about kids with muscular skeletal problems, speech defects or mild epilepsy, and maybe some sight and hearing problems. That requires some care in addition to their being quadriplegic, but only a moderate level of care. When you get to level three, then you're talking about things like severe scoliosis, severe epilepsy, blindness, a ventricular shunt, and repeated aspiration pneumonia, which is from the neuralgic incapacity. Then you're getting into a much different cost level. With level four, you're getting into intractable epilepsy, gastrostomy status, and panhypopituitarism. Since cerebral palsy is caused by damage to the brain, it can cause just orthopedic problems or, depending upon how widespread that damage is to the brain, all sorts of problems, so there's a tremendous differentiation there.

Sickle-cell anemia is also a progressive condition. But its presentation varies from mild to severe, and the rate at which it progresses is very variable. So, at level two, you're seeing some sickle-cell crises but not multiple hospitalizations. At level three, you're getting into multiple hospitalizations. With level four, in your adult years, you're showing renal diseases, such as nephritis or recurrent pulmonary infarct. Although Washington State had very few sickle-cell anemia patients it would be very different in, say, New York City. The overall R-squared on this Medicaid database was 30.5%, which was a very encouraging statistical level. This is a database with a lot of severely chronically ill, so the ability to predict is going to be greater.

To summarize, the system is clinically derived. It's a categorical model, not a regression model. It's severity-adjusted. It specifically identifies combinations of major illnesses. It's extensively tested with historical data. It facilitates communication with physicians and management. The predictive performance is very promising.

Mr. Sanford B. Herman: I was giving some thought to how you handle the people who die. I think you're creating an understatement bias in terms of those conditions. As I understand it, the capitation really is monthly. Let's say you die in the second month, and we assume that the costs tend to be lumped very heavily near the time of death, the person is only going to get two months of capitation.

Yet, he or she is going to have a very high cost. Now I would think that the appropriate way of doing it would be to count partially your exposure and build that in there as an approximation. Just assume that the deaths occur in the middle of the year and take the last six months of costs and the one-half year exposure on those situations.

Mr. Averill: I think that's a reasonable suggestion. We're certainly open to any suggestion. As I said, we're going to do all kinds of simulations. Certainly, we should attempt to do that type of model and look at the performance of the system.

Ms. Ranade: I wanted to add that the bulk of the work so far has been the clinical categorization and the clinical models. The last phase will involve the actuarial application. Some of the refinements will need to be worked out during that phase.

Mr. Harris: You anticipated my comment that frequently there is a disconnect between the development of the risk factor, which is done from a clinical perspective, and the application when the rubber hits the road. It sounds like we're on the road to keep those from being disconnected in this particular model. My question is in terms of the Medicare application of the PIP-DCGs. Medicare's going to be taking data that is a year to a year-and-a-half old, assigning each of the Medicare enrollees of an HMO to a category and assigning a price to that category. For most of the health plans I work with, you may have the same number of enrollees, but the bodies that you have a year-and-a-half later are frequently not the same individuals who you had before, so you'll have bunches of new enrollees coming in. They may or may not be clinically and demographically similar to the ones who you had a year-and-a-half ago. I was wondering if Medicare was going to make that adjustment.

Mr. Averill: With regard to disconnect, it's one better than that. ASA will be taking the model and the data and looking at it from an actuarial perspective totally independent of us and will give us a critique of the use of the model for actuarial purposes. That's sort of an independent stage of the project that will be going on over the next couple of months.

With regard to what Medicare will do, Medicare will not have a factor for new enrollees. For enrollees who are assigned a PIP-DCG, that designation will follow them even if they switch plans. You're going to have the disadvantage or advantage of how the prior HMO managed the patient.

Mr. Harris: Just like nature abhors a vacuum, I find that the market abhors retrospective adjustments even though scientifically they may make more sense and have more validity.

Mr. Averill: Medicare hasn't proposed any retrospective adjustment. But the basis is going to be what history Medicare has on the patient.

From the Floor: Your history follows you around?

Mr. Averill: Right. It follows you around. Medicare is fortunate to actually have quite a good data system. They have what's called the common working file, an enrollee file that has multiple years worth of data—every claim, every diagnosis, and every procedure. Indeed, that's how we got our 5% sample, which is a sample from that particular database.

CHART 1
1992 DISTRIBUTION OF MEDICARE ENROLLEES AND PROGRAM PAYMENTS
Average Medicare Payment per Enrollee is $3,391