## Session 8PD

## Risk Adjusters: Are They Working?

**Track:** Health
**Key words:** Risk Management

**Moderator:** JOHN M. BERTKO
**Panelists:** JAMES S. MATTHISEN
LESLIE F. PETERS
DAVID C. SKY
**Recorder:** LESLIE F. PETERS

*Summary: Debaters examine what has been done and is currently happening regarding the use of risk adjusters to spread risk among various carriers. What approaches are being used? Are they working? Are they equitable? These approaches include:*
* Understanding the purpose of risk adjusters
* Gaining an appreciation of the difficulty to find a risk adjuster method that is both equitable and administratively feasible
* Gaining knowledge of the risk adjuster approaches being used

**Mr. John M. Bertko:** I've been working in risk adjustment for six or seven years now, and I've been talking to the guy who recruited me to do the session and to people at the Society. I took the moderator's prerogative and changed the title of the section. It was to be a debate; that is, is risk adjustment working? but judging from the panel we have it really is, how is risk adjustment working? This session will cover some of the real-life circumstances that are happening today. I've had some experience in a number of different states—California and looking over James Matthisen's shoulder in Washington state, with the Alpha Center who's doing an incredible amount of work with the health care researchers on this and with what's going on in Medicare Risk.

---

Note: The chart referenced in the text can be found at the end of the manuscript.

So, there are working models of risk adjustment in the commercial sector. There are working models certainly in the Medicaid sector. And soon there should be working models in Medicare. Let me just say that if you haven't paid much attention to this yet, Medicare, of course, both on a personal basis and also have some effect on your personal life and your business, and, as many of you know, the Balanced Budget Act in January 1, 2000 says there shall be health-based risk adjustment going beyond the demographic adjusters and the adjusting average per capita cost. One of the people I've heard give their speech on risk adjustment from the Health Care Financing Administration says the train has left the station. Now, that doesn't mean there might not be a piece of track that derails at some time between now and 2000, but someone's got to go over there and do that. If things continue to go along, we will have an operational risk adjustment system in the year 2000.

Let me take just a moment to introduce our speakers panel. First we have James Matthisen who is with William Mercer. James is a consulting actuary with the Government Health Care Practice there and has worked with managed medical care and health care reform for many years and would describe himself, or at least he let's me describe him, as the principal architect of the Washington State Health Care Authority's standardized bid and premium adjustment process, which is one of the more innovative processes that's going on today on behalf of a very large purchaser. In terms of market concentration, the Health Care Authority manages a greater percentage of its market than maybe anybody else in the country, certainly a bigger percentage than the California Public Employees' Retirement System does and maybe some of the other state plans.

Leslie Peters and I used to work together at Coopers & Lybrand. Leslie spent the formative years of her actuarial career with Aetna and then saw the light and moved over to consulting. She's been with Coopers & Lybrand since 1994, where she works on health care reform, Medicaid managed care, and risk assessment and risk adjustment. She has worked on risk adjustment with the Health Insurance Plan of California purchasing cooperative since 1994.

David Sky is with the New Hampshire Insurance Department. He started his career with John Hancock and is currently the actuary from the Life, Accident, and Health Division there, and is the acting director for that division. David's going to give us a little bit of a contrast and give us a regulator's perspective on this. I would also ask any of you in the audience, and I notice a number of people have a variety of experience in states that have some forms of risk adjustment, to feel free to relate those kinds of comments during the session. Risk adjustment is clearly a work-

in-progress; that is, we hope it gets better as we go along.  It certainly has shown progress over the last three or four years, and that's the tone that we want to set today.

**Mr. James S. Matthisen:**  I am going to primarily talk about the employer buying health insurance or a self-insuring link in the health care purchasing chain.  So it is really going to be, roughly speaking, at the top.  There's a bunch of other places you could think about using risk adjustment technology, but for now that's the one I'm going to focus on.  In the work that I've spent most of my time doing, the employee contributions are a key element that's, roughly speaking, a managed competition model where employee contributions are supposed to be set on a level playing field for employees to pick a plan based on its efficiency, its cost, and perceived quality.

To the extent that this is mostly about the Washington State Health Care Authority, they have gone down that managed competition road.  They have almost a perfectly standard plan design.  They do have HMOs and point-
of-service side by side but really very similar plan designs, and there is full insurance for all of those plans.  There's a PPO that fits in alongside with less rich benefits.  Its managed competition employee premium reflects both a benefits design and a theoretical efficiency.  And, of course, they have annual open enrollment.

One of the things that I think makes risk adjustment discussions somewhat dizzying or stupefying is the terms that are around for essentially the same concepts.  I tried to bucket them because there are pretty much two standards—risk assessment and risk adjustment.  Risk assessment, or health status assessment, is just looking at people's characteristics through some measurement method and putting them into cost groups.  Risk adjustment is doing something with those cost groups, paying plans differently or subcapitating groups differently or doing something with the information you have from assessing the population's risk.  There is a move to try to use the term health-based payment for risk adjustment and health status assessment.  I don't know if that'll work because risk is just fun to talk about.  Truly, these aren't really measuring risk or variability.  They're really measuring expected cost.  But, nonetheless, risk adjustment is the one that sounds good to most people.

In terms of a dual standard premium arrangement or a dual standard rate measurement, there's a comparison rate.  This rate is a level playing field premium that plans are hopefully fairly compared on.  In my world, that's for an average population.  We ask the plans to bid assuming they have the entire Washington State employee population, and the rate that they would generate for that population becomes the rate they are compared against their peers on.  Some of the

other terms for this are standard rate, normalized rate, bid rate, and average rate. The second dimension is the payment rate. This is the part that's now trying to pay fairly, given disproportionate risk across health plans. It seems so simple now, and for this group it probably is, but I've spent a lot of my time talking to health plans about the difference between the comparison rate and the payment rate, and I still do run into health plans frequently that say, "Listen, I can do it for $90." They know they have 10% of the population who are all under the age of 25 but still believe they can take all comers at $90, and they really should be rewarded for having essentially a low payment rate as opposed to a good, solid, low, comparison rate.

The project that I've spent a few years on was Robert Wood Johnson funded, and it was academically driven. The University of Washington was the primary contractor working with the Health Care Authority and with Mercer. As such, it had a pretty solid and theoretical research and policy tone. It was robust, both on the statistical end and policy end. I think the policy things are more interesting, but if people have questions later about the statistical details, feel free to ask. I'll go through these fast because I feel like I'm being a little bit wordy. The Health Care Authority as an employer just wanted to be able to compare plans fairly during bid renewals and in all external representations. So, employee contributions, COBRA rates, and retiree premiums for early retirees are all set based on the standardized comparison rate basis. Then they wanted to pay plans fairly. That's the payment basis for predictable differences in health status of enrollees.

I want to take a minute on that word predictable. A lot of times I get in conversations where people say aren't we just going back to experience rating? You know the diagnoses of these people, and you're going to pay more for plans that have sick people. Because of the nature of at least the model that we're using, your diagnoses in year $x$ are used to predict your costs in year $x$ plus 2. So, if you have a multiple trauma, a car accident, surfing accident or whatever, in year $x$, you're not likely to have high year-two costs, but if you're a very ill diabetic, you are. So, it gets to a chronic health status as opposed to acute incidence. Again, we want to encourage plans to enroll all individuals regardless of their health. It is an open enrollment world. Employees theoretically have all choices with or without risk adjustment, but to the extent there's some selection, some cherry-picking, cream-skimming opportunities, we wanted to limit those.

The goal is to have individuals make personal decisions with respect to their plan, quality, and cost. Again, especially with the standardized plan side, a difference in premium reflects either how good that plan is at managing cost or the level of discounts they get. At any rate they're going to balance the cost with a perception

of quality and retain the insurance function of health plans.  We didn't want to make this feel like self-funding, and we wanted those multiple trauma car crashes to be something the health plan takes as a risk.  Minimizing administrative impact, of course, is always a good goal.

We had a pretty thorough confidentiality discussion in Washington state.  We were a little misguided on the sort of practical aspects of confidentiality, thinking that really no one's going to figure it out.  Practically no one would figure it out, but from a stakeholder perspective and from a managing-the-process perspective we ended up getting pretty deep into a unencryptable encryption digestion algorithm.  Once the data leaves the health plan, nobody could ever go back to an individually identifiable Social Security number, and once the eligibility data leaves the employer the same is true.  Just so we could get something done we put in this last policy goal—evolution, not revolution.  Let's make a step.  Let's do something even if it's not perfect.  So, again, the fundamental concept is predictable population costs.  We're not really moving toward experience rating, but we are trying to really account for population differences, and, to just give a couple real examples, in our population we found a range of about plus and minus 20% from the average population.  Early retirees and COBRA people are disproportionately distributed into the PPO plan, and it really did make a significant difference.

Assessment dimensions.  There are different ways to measure health status or health risk.  I don't think anybody is going to go through all those acronyms, but there are a lot of diagnostic-grouping acronyms, diagnostic cost groups, ambulatory care groups, and others, survey-based.  There's a lot of good actuarial work in figuring out measurement time periods and prediction time periods and how to actually implement some of these very good tools, and there are a lot of ways to do prediction.  As I said, our model was statistically research intensive, and we have a two-part model where Part I is the probability a person would have some health care.  Part II is the expected cost of care given that he or she had some health care.  The prediction is Part I times Part II.  It's this very complex, generalized regression with a gamma function.  For those of you who like that kind of stuff there's probably a research paper here that would really float your boat.

On the payment side is really where practitioners and actuaries need to come in and help.  There's been a ton of research.  There are many researchers who are thinking about predictions and risk assessment tools.  But it's trying to design the payment system that I think is at least a step behind the actual tools.  Next, I'll review some of the very basic things.  Prospective is where you sort of get a relativity for a health plan and assume it'll be the same next year and go with it. The next year you get a relativity, so you're always a little bit off.  Concurrent derived

plan relativity in the year of payment. This is what we're doing. We essentially have a big database of the health status buckets for all of the potential people from the measurement period, and then we, after open enrollment, quickly figure out which person is in which plan and which person we can match from our big database of potential health status information. For people who are new hires, they go through a demographic-only model, including a couple, good variables like early retiree status or COBRA status, so that theoretically we're real-time. It takes until about March to fix January through the rest of the year, but we're paying in the year of the enrollment for the person who has the health status. Health status is old, but the eligibility data is real time. There are some retrospective approaches also where at the end of the year there's essentially a settlement approach.

At a very high level, the accuracy of our move from a demographic only to a health status approach, using the Diagnostic Cost Group (DCG), was about three times better. Now, there's a lot of ways to measure and fight about R-squareds and different measurement tools, but what I tell people when they ask is it's about three times better than age and gender only. Interestingly, the results have tracked quite well with intuition. The PPO program and the less managed of the managed care programs have higher health status predictions than demographic predictions. Like me, they are sicker than their age. Also, the plans that have been around longer have had more time, and durational impact and effects, and they are also sicker than their age.

In our study it can be very honestly said that the managed care plans look less efficient and cheap than they used to. That's not really a profound statement. When it really happens people start wondering, thinking, discussing backlashes, things like that. On the PPO side we were doing this managed competition and employee contributions pre-health status. The first year we just had age and sex, and the PPO was 10–12% sicker than average. With worse benefit designs we still had a significant employee contribution in the first year out of the chute, and the PPO lost 30,000 people. The next year along we have the health status in place. The PPO goes from 10–12% to about 19% sicker than average with better measurement. The employee contributions go away, and 25,000 or 30,000 people come back in. So, I don't know if we slowed it, stopped it, or reversed it, but I think that clearly risk adjustment is the root of why they came back in, and the lack of health-status-based risk adjustment is the root of why they left, given the managed competition rules we set up. We kind of did that to ourselves, but it's a pretty significant impact. We have noticed that for the healthy plans that used to be able to come in with a really good-looking bid and still make money and had the potential to shadow pricing, that game is a lot harder to play.

My short list of keys to implementation is culled from a very long list of implementation steps. Enrollment data is important. I guess I took it for granted for a while, but in a lot of my subsequent discussions with other purchasers you need to have a good way to uniquely identify every warm body in your population to do it in a robust way, or you need to develop new models that do it based on incomplete enrollment data. Health plan understanding and buy-in is key. We pretty much were meeting to death on our project. Somewhere between 100 and 150 donuts were consumed per health plan participant, me included. I'm still trying to recover from that. Our project senior management buy-in was important and difficult, not because they didn't want to think about it but because there's a lot to digest. There's a lot to think about, moving from the question: How much do you want to get paid? to the question: How much do you think you should get paid in this hypothetical situation? Furthermore, I'm not going to really pay you that much, but I will for three months, and then I'll adjust it later. That's somewhat challenging. The key to a lot of our health plan and senior management buy-in was simulations. We did a full dry run the year before, and we were still phasing in the health status part of the change.

I talked enough probably about confidentiality. That was a sticky issue, again, not because it's so hard to design a secure process but because it's really important in some political and other situations to do a very thorough job. Audit. Somewhere down the road, and I think we're not there yet, it's going to be important to make sure that the data, in our case a DCG prediction that is generating a $4,000-a-month prediction for a person, is solid. So, I can't tell you I have the exact audit plan. We're doing a lot of work to make sure we design one, but given our phase-in process, we have a little bit of time. Turnover of health industry staff sounds trivial, but we had two plans with new underwriters that hadn't been to the 150-donut meetings and just bid what they wanted to get paid. Then they got a 10% or 15% reduction and I think are still employed.

The financial flow and structure is a challenge. I don't really want to elaborate on that, but it's pretty important to think very carefully about all of the side issues, and in the work I've done it's important to have a purchaser who's kind of in the way of the cash flow. If you're just procuring some rates for some people to go and pay directly, it's a lot harder to think through these systems. The Health Care Authority's approach has them as the big sort of melting pot of money in and money out, and we put in an expense-neutral process to make sure that the money out doesn't change when it's readjusted, but those become really good, meaty, actuarial discussions. Plan delegation and subcapitation and delegation of claims paying. Subcapitation where no claims data comes back is something you'll probably hear about from a few of us today. I didn't list health plan data as one of the keys to

implementation. That might have been an oversight or maybe I was just feeling optimistic at the time.  The approach we took is diagnostic based.  It's inpatient and outpatient.  It's all International Classification of Diseases-9th Revision (ICD-9) data, and it is a big deal to get that.  On the other hand we got it, and it's continuing to get better, and we feel that it is a reasonable thing to ask your health plans to know the diagnoses of the people you're paying them to take care of.

I thought I'd do a quick blurb on other applications, but I don't have a whole lot to say, and I'm not going to elaborate too much, but I do think that at the root of this technology is an understanding of the predictable costs of groups of people.  You can use that knowledge to try to help measure the quality of health plans.  For measuring health care outcomes and population health status over time, you should be able to, and this is a little wishful thinking because data are not going to stay static, but I'm interested to track this project over time and see if some plans help status-based predictions for a closed group of people beat their demographic predictions for a closed group of people and whether other plans' health status is higher than demographic.  Now, it's a little bit of a first step, but I do think that there's some room there for analysis.  This kind of stuff is something that would be good on a report card, providing information to consumers.  We didn't really need that.  We have employee contributions—big, flashing light.  Here's how much it costs for Plan X, Y, and Z.  If you're doing something where you don't have that tool, a scorecard, an efficiency index, could be something that people would want to consider.  And then, of course, there's just going to be a ton of work where the next level between the health plan and whomever they contract with and whoever they contract with and whomever they contract with.  So, I think there's a ton of applications for this technology below this purchaser insurance company level that we're talking about.

Just for kicks I thought I'd do a couple why nots.  I could have saved it for the question-and-answer (Q&A) period, but these are some standard ones.  If I get my people really healthy, you're going to pay me less, so why should I do that?  That's a pretty good question.  The good news is that if you really get your people healthy, none of these methods can predict a zero user.  The lowest they can predict is $30–40 a month for a 40-year-old person.  So, if you really get zero users, you're still winning.  You're just not winning as much.  The real incentive is to capture great data and show a lot of accurate diagnostic data, but manage your plans within that better.  So there may be more advantages to actually embracing the sicker folks and then managing well.

Another negative is that there's a huge administrative burden.  There's no doubt that getting these first few off the ground is a pretty big challenge, but I would say that

the response from a pro-risk adjustment person would be small compared to the dollars that get moved around.  If we start now, it should be actually as easy as running diagnostic-related groups in a few years.

This is a good why not.  We only have 20 people, and your method says that they're all really, really, really healthy.  What are you going to do?  Our answer has been, well, you're a health plan.  This is only 20 people.  We're going to pay you really, really low rates, and the rest of your book of business is your padding for that.  This method, on an individual-by-individual basis predicts better.  If this plan were only 20 people, I think we'd have a different discussion, like why are you in business?

Data are less and less available.  This is the California experience.  This is not yet the Washington experience, but we have a few of those California plans in Washington and they are thinking really hard about what to write into their contracts with subcapitated providers so that they get excellent quality data back.  They are sitting up and noticing.  At this point they don't seem to be pulling the plug and saying we don't want to play.  They're asking, how can we get you the data?

**Ms. Leslie F. Peters:**  I'm going to be talking today about what I see as some important design features of a successful risk adjustment system and then go through a few real-world examples in regard to these successful design features and how their presence or lack thereof affects how well the system works.  I'm sure that many people in the room have heard the statistic that for a given insured population, 20% of the people incur 80% of the costs for that group.  So, given this setup, there's enormous incentive for health plans to get the 80% of the people who only incur 20% of the costs, or, in other words, to not get those 20% of the people who are very high users of health care services.  What this means is that without a risk adjustment system there are only incentives for plans to only get low-risk people.  Successful risk adjustment systems have features that impact plan behavior and have the ability to change the incentives that we were just discussing.

The first thing that I see as an important feature is that there be no limit on the amount of funds transferred between plans.  If you have a limit, say, a percentage of premium or a flat amount per person, then plans with a low-risk population will see that they only have a certain amount to lose, and if that's an acceptable amount for them to lose, then they'll just continue to attract a low-risk population.  Also, if there's a limit on the amount of funds transferred, high-risk plans can be inadequately compensated for the amount of risk that they assume because there's not enough money in the system to fully compensate them for their risk.

The second feature that I see as particularly important is related to the first one in that the system addresses low-risk plans as well as high-risk plans. In a system that only addresses high-risk plans, that looks more like a reinsurance system as opposed to a risk adjustment system. And also if low-risk plans are not included in the system, then they'll just continue to have incentives to remain low-risk plans.

The third important feature is timely administration. In a system where there's a long gap between when a high-cost member is identified or significant claims occurred and when the plan receives any compensation for having this person, then there's going to be less connection between these two events and less impact on plan behavior.

The last important feature is having a complete risk pool. If you have opportunities to have your potential pool of people carved up or split out as if, for example, certain types of groups can opt out of the system, then you'll have a tendency for healthier people to leave your risk adjustment system, and then your system will just serve to transfer money between the various high-risk plans because all the low-risk people are outside of your system.

The first example that I'm going to talk about is the Kentucky high-cost case fund. It's a risk adjustment system that covers small groups and individual insureds in the state of Kentucky. However, association plans can opt out of the system, so association plans are not in the risk adjustment pool. The system covers only nine medical conditions, mostly transplants, AIDS, premature infants, end-stage renal disease, and leukemia, so a relatively short list of conditions are covered. The system is a retrospective. James just discussed that you could have a prospective system, a concurrent system (which is the one he's working on), or a retrospective system. Over the course of the year each health plan pays a certain amount of money into the risk adjustment pool which is 1% of premium. At the end of the year each plan submits qualifying cases for those people who had one of those nine high-cost conditions, and then we calculate a score for each plan, and if Plan X has a score that's higher than the statewide average, they get partially compensated for this excess risk.

In evaluating whether I think this system works or not, I'm going to go back to those key design features that I just mentioned. The first is that there's a fairly long time lag in between when one of these high-cost cases occurs and when the plan gets any money for it. Risk adjustment scores and transfer payments are made generally in July following the year in which the case occurred. This setup has limited impact on plan behavior because the compensation for having one of these high-risk plans

comes upward and a year later from when the case occurred. So, it's not really going to have much impact on health plan premium development or marketing strategies. Also, the list of conditions is very limited. It just includes transplants, AIDS, preemies, et cetera, it doesn't include this middle range of severe and chronic conditions like cancer, diabetes, heart disease, et cetera. Once again, because the list of conditions is fairly limited, a limited amount of funds will be transferred between plans, which will have a limited impact on the incentives in the marketplace. By its design structure it really doesn't address low-risk plans either because plans just pay 1% of premium into the pool, and then funds are taken from this pool a year and a half later and given to these high-risk plans, but for low-risk plans, the maximum that they're on the hook for is the 1% of premium. If they would like to still remain low-risk plans, all they have to do is give up 1% of premium. Once again, it doesn't really have a big impact on the desire to have a low-risk population.

The last feature of this Kentucky program is that the market is segmented. The association plans are allowed to opt out. So, the people left inside the risk adjustment system are the people who can't get cheaper insurance on the outside of the system. The risk adjustment feature of it turns more into just transferring money between the high-cost plans as opposed to transferring money from low-cost plans to high-cost plans. As a consequence, we only have two individual insurers left participating in the individual market in Kentucky.

The next risk adjustment system that I'm going to talk about is the Health Insurance Plan of California (HIPC) that John mentioned. It's a small employer purchasing pool for employers with 2–50 employees, and each employee has a choice of 20 different health plans that they can choose from. There are about 130,000 people in the HIPC right now. One key feature of the HIPC is that participating health plans are not allowed to sell insurance in the open group market for less than they sell it to the HIPC. Even though this is a voluntary purchasing pool, employers are free to find insurance wherever they want. It prevents market segmentation by forcing the health plans to give the best price to the employees in the HIPC.

The risk adjustment feature of the HIPC has several components. Health plans are not allowed to rate for gender, health status, or area. Those are the features that we apply risk adjustment to, and the health status portion of it consists of 136 marker diagnoses, which include a wide range of conditions including cancer, heart disease, multiple sclerosis, et cetera. The important feature of this system is that to qualify as having one of these conditions you have to have had an inpatient overnight stay with this diagnosis. To be counted as a cancer patient you have to have had an inpatient overnight stay for cancer. The risk adjustment portion is

prospective. The Kentucky example I just mentioned is retrospective. This one's prospective.

We take a look at what happened last year and use it to figure out how much people are going to get paid next year. They're similar to James's discussion. There's a two-year gap in between. Each health plan is calculated a score, which is the conglomeration of their area score, their family size score, actually, and their health status score, and we calculate transfer amounts for each health plan, and the health plans are told what their risk transfer amount is going to be in advance of rate negotiations. So, Health Plan X is told you're going to have to pay a dollar per member per month (PMPM), and Health Plan Y is going to be told you're going to receive a dollar PMPM. Health plans know this amount before they enter into written negotiations with the HIPC. They can take this information into account when setting the premiums that they're going to charge in the pool. As a prospective risk adjustment system, it assumes that past distribution of risk among health plans is going to be indicative of the future distribution risk on that health plan.

In evaluating how well the HIPC system is working, when you look at the risk transfer amounts they're very small. About 0.1% of premium gets transferred in the HIPC risk adjustment system, and part of that is related to the fact that we only use inpatient marker conditions. It only measures the risk of people who were hospitalized for 1 of these 136 conditions. So, if you happen to have cancer but were in a managed care plan that was doing its job, you maybe weren't hospitalized. Under this system you don't get counted. The plan doesn't get credit for you unless you are hospitalized. That's seen as a significant drawback of the system, but, as James mentioned this is California, and the standard response that we get from health plans in California is we don't have data. The HIPC is going through some transition right now, so it's unclear whether they will progress forward with a more comprehensive inpatient/outpatient list of conditions.

In evaluating whether this works or not, the risk transfers are small. The fact that we use only inpatient conditions reduces the transfer amounts so the very high-risk plans may not actually get enough money for it to be worth their while to participate in the pool. We had the one and only PPO that was left in the pool I believe two years ago that was informed that they were going to get a risk transfer amount of $15 PMPM, and they dropped out of the pool anyway because they said that our risk is so bad $15 PMPM is not enough. The HIPC's risk adjustment system is perceived as a step in the right direction, but since it's not all that robust, it has a limited impact.

I have a couple of Medicaid examples to talk about next.  A little background on the Utah Medicaid system.  They have mandatory managed care enrollment of their Medicaid members in urban areas.  They have four health plans serving members.  The Utah managed care market is not all that developed.  There's not a lot of competition.  They really only have four plans in the state.  And they don't have any encounter data available yet.  But the state wanted to go forward with some type of risk adjustment mechanism because in the past they had been negotiating rates with the health plans on an annual basis.  Every year they'll sit down with the health plans and negotiate how much they're going to get paid, and the health plans up till this point had been presenting largely anecdotal evidence that they were being selected against.  The state needed some form of measurement that they could use to support the claims or deny the claims by the health plans of their level of risk selection.  So, we modified the method that's used in the HIPC to be used for the Utah Medicaid market.  This is the first year that they're using it.  The rates will go into effect on July 1,1998.  But even with only inpatient diagnosis data we found a very close correlation between historical, anecdotal evidence and the risk adjustment scores.  So, in terms of whether we think this is working we think that it's moving in the right direction, and it actually attaches a concrete answer to statements made by health plans about the risk.

Oregon Medicaid is also implementing risk adjustment next week with their July 1 contracts.  They have mandatory managed care enrollment for all enrollees statewide. There are 13 plans serving 350,000 Medicaid enrollees in the state of Oregon.  They're going to implement two separate risk adjustment systems next week.  One is for the disabled groups.  We're going to use the disability payment system (DPS) which was developed by Rick Kronick at UC San Diego.  The DPS was designed specifically for a Medicaid disabled population group, so that's why it was chosen for this use.  The DPS is constructed based on 2,400 diagnoses grouped into 43 disease categories, and it's a comprehensive risk adjustment system.  Whereas the ones I described previously relied mostly on inpatient data, this relies on complete health plan encounter data, everything from prescription drugs to office visits, as well as inpatient and outpatient.  That's being used for several disabled and related groups that they have.

For the Aid to Families with Dependent Children (AFDC) and related groups, which I guess is now called Temporary Assistance to Needy Families, they are risk adjusting based on the prevalence of maternity and newborn cases in each health plan.  For a lot of these AFDC and related groups there's a high prevalence of maternities and newborns that are more expensive than average, but, more significantly, there's not an even distribution of maternity and newborn cases

among the health plans. So, hence, the desire to risk adjust this component of the capitation.

We've already calculated the risk adjustment scores, and in the Medicaid arena in Oregon there's no negotiation with the health plans. Coopers & Lybrand calculates the rates for the state, and then the state publishes them and says if you would like to participate, this is what you'll get paid. Now they're going to take those average rates and multiply them by the plan's specific risk adjustment score to get the amount that the plan will actually receive. The important feature about calculating these risk adjustment scores in Oregon was that we relied on encounter data submitted by health plans. Oregon has been working for about the last five years on getting good data from their health plans, and this was the first year we thought it was even usable, yet still it resulted in wild fluctuations in results. On the maternity and newborn side the results were so all over the place that they actually put in an artificial floor that said if anybody's score was less than 75%, we're just going to bump you all up to 75% because we think that your scores are so low because your data's so bad. Similarly, on the DPS for the disabled populations we put a 10% corridor. If your score was greater than 1.1 or less than 0.9, we truncated and renormalized a few times to arrive at the final rates. Part of this reliance on data is that when you're calculating risk adjustment scores the score represents the health plan's risk compared to the statewide average. If you're getting data from all the different health plans and one of them has bad data, they will have an abnormally low score and everybody else will have an abnormally high score compared to if you had good data from everyone. All these things put together led us to think that at least in the first year we should put these limits on the risk adjustment scores to get things going the first year and develop a workable system over time.

In conclusion, risk adjustment has some mixed results. Some things are clearly not working. Some things are working OK. But, more importantly, for the things that are working OK, you can identify what the important features are that are making them work and take incremental steps to keep moving the market in the right direction. Success, based on our work in Oregon, clearly depends on good data, but overall I think that risk adjustment will not fix a market that's not working on its own.

**Mr. Bertko:** David will now give us a different perspective.

**Mr David C. Sky:** I'm going to be talking about high-risk pool alternatives. I work for the insurance department, but I have to say that anytime I'm expressing an opinion it's my own and not necessarily that of the New Hampshire Insurance Department. I guess the question would be alternative to what? Why do you need

a high-risk pool? Typically it's in conjunction with health insurance reforms that are providing guaranteed issue rates to certain members of the population who may not have had access to health insurance before, and that will be the context that I'm going to be talking about, what's happened in New Hampshire. With the new health insurance reforms we provide guaranteed access rights to individual purchasers of health insurance, and with these new access rights, there was a risk subsidization mechanism that was established whereby writers of individual health insurance plans would receive a subsidy from other participants in the health insurance market to the extent that they were taking on greater risk than they were taking on before.

Some advantages to this alternative approach would be that it provides the same choice of products that would be available to all individual purchasers of health insurance, and this approach also maintains proper incentives for claims adjudication and management, whereas with a high-risk pool you could have limited choices of products that were available to participants who can only qualify for the high-risk pool, and there may not be proper incentives to adjudicate claims for those in the high-risk pool.

A little bit of background of what's been going on in New Hampshire. The legislators said from a societal perspective we think that discrimination by health status in terms of choice of products and cost of products is inappropriate, and our reforms that went into effect in 1995 really sought to fix that. We have guaranteed issue of all products and portability. You can go from plan to plan with credible coverage. We did provide a concept of grandfathering for products that were purchased prior to reform. Carriers could maintain those blocks separately from the guaranteed issue products, and we did define a small employer as having 1–100 employees. A lot of the people who were typically left to purchase in the individual market could now purchase in the small group market, which really caused some market bifurcation problems. In the level playing field concept Blue Cross/Blue Shield, which was previously exempt from premium taxes, was now subjected to premium taxes because they were no longer a carrier of last resort.

Some market observations. Carriers price their products in this guaranteed issue environment based on the experience of their pool. Blue Cross/Blue Shield had a lot of lead weight, previously being the carrier of last resort, and they found that without any kind of subsidy that went into place in 1995 that they weren't able to compete and they were still stuck with this risk or this poor pool, and they eventually withdrew from the market, which really woke up the insurance department. They represented about 80% of the individual market in terms of

covered lives, and it was pretty much that action that caused the commissioner to hold a hearing about the availability of health insurance in the individual market.

Before 1995, according to our studies, there were about 40,000 covered lives; that is, people who obtained coverage in the individual market. With the reforms and the evolution to date we feel there are about 20,000 people who have coverage through a grandfathered individual policy, which was issued prior to 1995. There are about maybe 5,000 people who are obtaining coverage in this guaranteed issue individual market. If you take the difference between 40,000—that's 25,000. A lot of those people are obtaining coverage as small employers. What's available in the marketplace typically are high deductible indemnity products with $1,000 deductible or higher for the products that are available right now.

If you wanted more information about what's going on in New Hampshire, the department published a paper called "Investigations of the Effect of New Hampshire's Health Insurance Reform Laws," and commissioned a study of what's been going on, which was done by the Center for Health Economics Research. That paper is available from the insurance department. You can send me an E-mail or contact the department. The department wrote a paper called "An Analysis of the Non-Group Market with Recommendations for Change," if you want more background information about what's going on in New Hampshire.

The commissioner's order, written in November 1997, did primarily three things. First it created an association of licensed New Hampshire insurance writers, and this is important because a lot of our domestics were concerned about retaliatory actions of other states if the state itself was going to be involved in assessing carriers for a subsidy mechanism. The commissioner orders an independent association to manage the subsidization mechanism of all licensed health insurance writers. It has a board of five writers and takes a super majority to do anything. Three of the writers are representatives of individual health insurance writers, and two of the writers are representatives of group health insurance writers, the writers who are going to be assessed to fund the subsidies. The next major part of the order is the subsidy determination process, which I'll be talking about in more detail, and the assessment base and determination process, which I'll also be addressing.

The first component of the subsidy determination is a retrospective approach. That means that a subsidy is paid to an individual writer for a given experience year based on that carrier's actual experience in that year. We've heard some talks about prospective versus retrospective. For a prospective, experience of one year would be used as a predictor of experience in the next year; we use a retrospective determination. The next piece of the subsidy determinations process is it only

considers high-cost claims—claims in excess of $25,000.  The predictor of variables that we use are attained age and primary diagnosis category, and if a claimant has more than one primary diagnosis category, carriers are instructed to bundle the claims experience into the diagnosis category that represents the majority of that claimant's claims.  All claims are eventually considered.  Finally, the subsidy is calculated based on the carrier's actual incidence of claims within these predictor categories times a predetermined assumed claim severity, which I'm going to be talking about.  You then subtract from that a theoretical expected incidence times a theoretical expected claim severity.  What's different between what we're doing and what you've heard about before is that typically in a risk adjustment setting you say these plans are participating within this population.  This is the average cost of the population.  I think James referred to it as the big cost or something like that.  To the extent that the competing plans for that population get different categories of risk, we're going to pay it in differently.  We're saying that this market is getting socked with worse risk than it can reasonably handle on its own.  We have to come up with some kind of risk level that we think the market can handle.  That's really a nebulous kind of decision.  We used the SOA's group large claim medical study that was recently published to calculate expected incidence and expected claim severity numbers.  For the severity numbers what we did was we imposed a gamma distribution on claim severity, and for the assumed claim severity we used the 85th percentile, the gamma distribution, replacing that with the carrier's actual claim severity.  We feel what this does, by paying on an assumed claim severity versus the carrier's actual claim severity, is it maintains incentive for proper claim adjudication because they're getting a fixed amount for that diagnosis code, yet it also provides some reimbursement to the extent that they have a greater number of these types of claims than they should reasonably expect to have.

I should also mention Alice Rosenblatt's paper, another Society paper, on risk adjuster techniques.  That paper really goes through a number of the predictor variables and their effectiveness as a risk adjuster predictor variable and a lot of statistical analysis.  We used that in making our decision to incorporate attained age and primary diagnosis as our predictor variables for our model.

The assessment base is all New Hampshire covered lives in all similar lines of insurance or types of insurance plans that provide coverage that's similar to major medical, and that includes group excess loss or stop-loss insurance.  The final subsidy in theory is multiplied by the ratio of the total anticipated subsidies that carriers should be receiving divided by the total base that's available, and hopefully there would always be adequate funds, but if there are not, each carrier's subsidy is reduced proportionately.  Political reality plays a huge part in the determination process, to that extent that subsidy floors and ceilings were also introduced into the

determination process.  A carrier's subsidy is limited so that the loss ratio won't be reduced below 75%, and a floor is introduced so that a each carrier's guaranteed a loss ratio of no worse than 95%.  Ideally, those types of floors and ceilings won't have to be in place, but those were necessary for political buy-in by the health carriers participating in the market.

The similarities in this process to the other presenters is clearly the risk adjuster methods that we're using to determine the subsidy amounts that plans will be getting, and I think the key difference from what the other presenters were talking about is that we have to use a predefined morbidity experience whereas they're using a pool average morbidity based on the participation either in Washington state employee pool or in the HIPC, those purchasing insurance through the HIPC.

There are some possible modifications and enhancements, and actually some of these have been incorporated into New Hampshire House Bill 1411, which is now on its way to the governor, and that's available at New Hampshire's Web site.  The New Hampshire design that was implemented by order includes essentially all individual insured lives as subsidy-eligible.  Perhaps a better design would be to find a way that only a certain percentage of the individual market would be considered  subsidy-eligible lives.  What we've done or what House Bill 1411 would do is allow a limited rating variation attributable to health status of up to a 20% increase in the rates based on health status, and only individuals who are issued a policy with a maximum allowable health status factor would be subsidy eligible.  All other lives would not be eligible for consideration in the subsidy determination process.  What that's supposed to do or the allowed rating variation is supposed to do is provide incentives for consumers to shop around so carriers can't just dump them into the subsidy pool.

I guess other things that you could do to enhance this process would be to have limited product choices or limited open enrollment periods and limit the number of high-risk individuals the plan would have to take.  As you start imposing these other types of limitations, the mechanism starts to look more and more like a high-risk pool and less and less like a risk subsidization mechanism.  That concludes my talk.

**Mr. Bertko:**  We have time for Q&A, and I have prearranged one.  One of the things that we didn't talk about today was the effect on providers and how providers are using risk adjusters.  Harry Sutton, who knows the people up in Minneapolis at Buyer's Health Care Action Group (BHCAG) who are doing this well, is willing to give us a couple of minutes about that.

**Harry L. Sutton, Jr.:** This system that I'm talking about is now in its second year, but it's the fifth year of an organization called the BHCAG in Minneapolis. The system that I'm talking about is also going to be used for Medicaid and for Minnesota Care, which is partially subsidized premiums for lower income individuals, but I'm going to talk primarily about the large employers. Approximately 20 of the largest employers—3-M Company, General Mills, and Dayton-Hudson are the big pushers of this—originally decided to band together to purchase by directly contracting with providers. Now, their initial go-round, which took place in 1994, hired somebody to negotiate deals with the HMOs in Minnesota, and they eventually put themselves with one of the HMOs and got a discount on what their average prices were of about 8%, but all I know is that somebody bid less than everybody else in order to get the business. They have spent three years developing a modified system that uses health risk adjusters to determine the compensation to the provider systems. What they did is set up provider systems or groups. They have some 15 to 20 groups. They overlap in specialty in hospitals, as they almost have to, but they're independent as far as primary care physicians. There's no overlap in primary care physicians.

The system that we use for both the state prospectively and the Buyer's Group is the ACG designed by John Hopkins and partially administered and marketed by CSC Computer Systems. 1997 was the first full year, and it's a three-year contract. The Buyer's Action Group has a request for bids that are supposed to be in by July 1 to get a new administrator, but essentially this is a modified, retrospective adjustment, and it takes six months past each quarter to figure out what happened. They adjust not only by the diagnostic weight of the ACG system now at 4.3, but also each of the provider systems gets their rates adjusted once a quarter. In other words the providers bid, similar to what the first discussion was, an average group of this total population. The amount of money they bid is redivided based on 12 months of previous data as to the degree of difficulty or risk involved in these populations. Suppose in each quarter they bid $100. If they're running $110 which takes you six months to figure out, approximately, after the beginning of the quarter, their fees are lowered by 10%. If they come in under budget that they bid, then they get an increase in fee levels of 10%, so that the rate that each physician gets for a given patient in each system is different. The primary care physicians would all be the same.

Now, sometimes in one of the big ones the specialists weren't willing to take a cut, so all the risk was on the primary care physicians. The ACG system is a diagnostic-based system. It's not a database system looking at claims. This is the theoretical diagnostic-based system between the care systems. Chart 1 gives you some idea of the variation. The size of the groups that are enrolled in these care systems range

from 800 people to 25,000 people.  In total there are 130,000 people, and as the first speaker talked about, it costs approximately 10% of premium to administer this thing, which is one of the problems.  They're trying to get a lower bid on handling the administration.  But it is very data-intensive because of the quarterly reanalysis of each of these systems.  The reason, and it's already caused a change, that this first group is so high, 125%, is that it's a pediatric-only group.  It's an HMO started by Children's Hospitals; therefore, they only take very ill children into their system, so they've had to redo the ACG system to do a better job of predicting cost on pediatrics and high-risk maternity claims.

This is an aggregate experience for a 12-month period.  All of the health risk adjuster systems use the latest 12 previous months of data, and since this is a self-insured plan and a point of service (POS) plan and these companies do not want to be interfered with by ERISA, they're all self-insured, and each employer pays on his own checks for what the claims actually are.  The rating system is dampened by using a $40,000 stop loss cut-off point so that a small system wouldn't get killed, but as far as HMOs go, this is only 250 days per 1,000.  There is a vast amount of data coming out of this system.  It's very expensive to do.  The ACG codes everything, four or five codes per office visits up to, if there is one, plus really detailed codes on the hospital side as well.  We have 18 users of this big computer system.  Every hospital has a system.  Every HMO has a system.  And any large carrier has a system as well.  There are changes pending.  Some of the employers think they pay too much.  The administrative cost is 10%, $15 million a year, approximately.  Part of it is start-up cost.  Once it's in there and you have it in your systems, you can run it more easily, but it is expensive and hard to deal with.

**Mr. Sky:**  I guess I wanted to follow up with Leslie on the Health Insurance Plan of California.  I think you said that the plans in the HIPC have to charge the same price as the non-HIPC plan to small employers, and that's probably for reform purposes, I would guess, but the risk adjustment mechanism is only available to plans in the HIPC, and to my way of thinking I guess there's variation with respect to any plan in the offering of insurance, and the risk adjustment mechanism helps to correct or mitigate the need to plan for variations because the risk adjustment process explains some of the variation or it takes away takes away the explained variation from the claim variation.  Wouldn't plans participating through the HIPC be advantaged relative to plans not participating in the HIPC?  Wouldn't that lead to a single payer system where all health insurance to that segment of the market might only get plans through a HIPC?

**Ms. Peters:**  To the best of my knowledge there has not been any discussion of that topic yet.  The health plans are required to provide proof every year that they're not

underpricing a pool, but up until this point there have been no discussions about whether this is an advantage or not.  It's not particularly perceived as an advantage.

**Mr. Joseph N. Romano:**  Just a general question.  I don't think I heard this in your discussion, James, in terms of the impact of rating areas or geography.  You talked a little bit about the capitation side and the impacts of that, but I'd be interested in your thoughts about how the historical rating areas of the different plans, in terms of whether they cover a particular service area, might impact or would be addressed in a risk adjustment system.

**Mr. Matthisen:**  That's a really good question.  It's one that is frequently asked about our model.  To the extent that employees live disproportionately around the state—healthy people in some areas, retirees in other areas—and to the extent that that has captured their diagnoses, the answer we don't need geographic areas, but to the extent that it's different issues like either rural areas are cheap, which used to be the old truth in Washington, or rural areas are expensive on a fee schedule and a hospital basis which seems to be the evolving truth in Washington, we do not have anything in our system to account for divergent source costs by region.  A lot of health plans want to talk about that, but at this point, as a statewide purchaser, the debate between going to a region-based rating system where someone in one plan would have a different premium in different counties is more negative than the potential bias of having no adjustment by geography.

**Mr. Bertko:**  I have a follow-up question for our panel.  I'm going to put James into working for a government payer rather than a large employer and then switch to the question, is risk adjustment ready for big time for large employers?  Later in the session we have a number of panels that are going to talk about buying the best value health plans.  Do you think risk adjustment is ready to be a part of that when it's a large employer?

**Mr. Matthisen:**  Let me make sure I have it right, plus I'm supposed to restate all questions.  Basically, is risk adjustment ready for big employers?

**Mr. Bertko:**  Yes.

**Mr. Matthisen:**  My personal opinion, sure.  There are some big employers with multiple health plans that are somehow being compared and purchased by employees.  I'm sure that there are a couple big employers who do value their PPO and are having PPO death spiral that they don't quite know what to do about.  It's still a little new, but from everything I've looked at you could today make that decision and in a year or two phase in a big employer health status risk approach.

In terms of saving a ton of money the jury may still be out, but in terms of more equitable, feel-good, accurate, and informed purchasing I think it's very possible.

**Mr. Bertko:** David, did you want to add to that?

**Mr. Sky:** Yes. I don't know if it's ready, but I know employers are ready to the extent that they want to offer their employees other choices in terms of health insurance. If you would want to continue to offer an indemnity plan because it would be a competitive advantage to keep those employees. Employers are really struggling, I think, to find risk adjustment processes that they can use for the plans that they wish to sponsor.

**Mr. James E. Carter:** I actually have a few questions for all of our participants, at least the three primary ones. Some clarification. First starting with James. When you were talking about the small employer market, I didn't know the number of employees or the employee contribution.

**Mr. Matthisen:** There are about 250,000 lives in the risk pool that is being risk adjusted, and the employee contributions aren't very large. On average, per family, it's about $15. I don't have the range memorized, but it's from $0 to about $22 for single coverage, and about three times that for family.

**Mr. Carter:** You also mentioned the fundamental concept that there was a variation of about plus or minus 20%. Is that by group in total or how does that play out?

**Mr. Matthisen:** We have about 16 health plans, and the one that looks healthiest has bad data. The healthiest credible plan is probably about minus 17–18%, and the sickest plan is about 19–20%. I might have exaggerated. Maybe it's plus or minus 17, but we essentially doubled that range of variability by health plans by adding health status and those other demographic variables I mentioned. COBRA status and early retiree status contributed to our spread of risk. Another statistic which compares to one of Leslie's is in our program about 9% of premium is moved one way or another.

**Mr. Carter:** The last question for is for Mr. Matthisen. You mentioned that data is less and less available. Can you explain for what reasons it's less and less available?

**Mr. Matthisen:** What I meant was the move to subcapitating and delegating claims payment without a requirement for data to be fed back. That was not intended to be a specific Washington state comment but more of a global concern. In

Washington state the fact that we're doing this should make the data more and more available.

**Mr. Carter:**  Leslie, you mentioned that there's no limit on the fund transfers.  Is there any thought or possibility of having any reinsurance over $1 million or something?

**Ms. Peters**:  In most of the systems that I discussed the risk transfer amounts are not at all based on what the health plan actually spent.  The scores are based on average cost for a given case.  Average cost for a particular cancer results in a particular score, which is completely independent of whatever the plan spent.  Plans are free to purchase reinsurance voluntarily, but the two issues aren't really connected.

**Mr. Carter:**  I guess this is about Kentucky, and you mentioned the expected impact on plan behavior.  Was there any kind of expectation that was expected?

**Ms. Peters:**  Our group was not involved in the inception of the program.  They actually developed this program by looking at New York's program and pretty much adopted wholesale what was being done in New York.  I don't actually know if there was a particular expectation of what it was going to achieve.

**Mr. Bertko:**  Yes, and just to add, New York's program is changing.

**Mr. Peters:**  Right.  New York is moving to more of like HIPC-style system.

**Mr. Carter:**  You mentioned that the HIPC in California expected underpricing.  How do they judge if they underprice, and how is it enforced?

**Ms. Peters:**  Every year health plans have to certify to HIPC that they are not underpricing.  They present their rating information and plan designs for what they sell in the open market to the people at the HIPC, and if it's not exactly the same plan design, we'll develop plan relativities and compare the difference in benefits to the difference in cost, and if they are underpricing the HIPC, they have two choices.  They can either lower the amount that they're charging the HIPC or they are forced to leave the HIPC.

**From the Floor:**  Is it compared to the average in that there's plus or minus 10%?  Does that compare to the average rate they use outside or, in other words, does it have to be the 0.9 compared to the normal rate?  There's a range they use outside of the HIPC.

**From the Floor:** When you're saying underpricing are you comparing it to the average outside of the HIPC or the low end of that range?

**Ms. Peters:** I'm not sure.

CHART 1
1995 AND 1996 ACTUAL
RISK ADJUSTED PMPMS