RECORD, Volume 28, No. 2*

San Francisco Spring Meeting
June 24–26, 2002

Session 99OF
Predictive Modeling

Track: Health

Moderator: ROBERT BRUCE CUMMING
Panelists: ROBERT BRUCE CUMMING
IAN G. DUNCAN
ELIZABETH V. LEWIS
MARK D. WERNICKE

Summary: Panelists discuss issues and potential application of predictive modeling to underwriting, rating, medical management, disease management, etc. Panelists also discuss practical issues associated with implementing predictive modeling.

MR. ROBERT BRUCE CUMMING: My name is Bob Cumming, and I'm a consulting actuary with Milliman USA in the Minneapolis office. I'm going to cover a bit of background on predictive modeling. Then we're going to move on to Mark Wernicke. Mark is with Humana and he's going to talk about applications of predictive modeling to underwriting and rating of small group business, using the Ingenix predictive model.

After Mark will be Ian Duncan from Landacorp, and he's going to talk about predictive modeling from the perspective of case management—that is, using predictive models to identify potential high-cost patients to direct into care or disease management programs. Finally, Elizabeth Lewis is with Blue Cross/Blue Shield of Minnesota and she's going to talk about some other nontraditional applications of risk adjusters, including how they've used adjusted clinical groups (ACGs) to help out with analyzing cost variation by area to do trend analysis and also to slot providers into different provider tiers.

As background, I am going to define predictive modeling and risk adjustment and discuss how they're similar and how they're different.

Note: The chart(s) referred to in the text can be found at the end of the manuscript.

Predictive modeling is the process of trying to predict future claim dollars for an individual or group of people. This might be done in order to identify whether that person might be a high-risk person to put into disease management, or in order to predict claim dollars for a small group employer for purposes of rating or underwriting. We often try to predict future claim dollars based on factors such as age, gender and past claim expenditures. What's different about predictive modeling is that it uses very detailed claim information to predict claim dollars, relying on very detailed utilization data such as diagnosis and procedure codes as well as pharmacy data.

Risk adjustment refers to the process of trying to predict or quantify differences in health status among populations. You're trying to focus in on what the illness burden of a population is relative to some other population. Risk adjustment is similar to predictive modeling in that it's also generally based on very detailed claim data as well, but it's a little different, and I'll talk about the differences in a second.

A predictive model is something that you use to do predictive modeling. It's designed for predicting claim dollars for that purpose—and again, you generally use very detailed information to drive this—diagnosis and pharmacy data, which is often also used for risk adjusters, but the difference with predictive modeling is that you can use anything. You can use procedure information; you might use utilization data, claim dollars, prior use, anything that helps improve the prediction. You can use any information out there to try to hone in on what a group might look like next year or what individuals might cost next year.

This differs from risk adjusters in that risk adjusters are designed to predict health status. They're supposed to pick up on differences in health status and health status alone. If you're using a risk adjuster for a traditional application, like payment to health plans or payment to providers, you don't want to reflect differences in fee levels or provider practice patterns As a result, with a risk adjuster, you're going to use detailed diagnosis and pharmacy data, but there are items that should not generally be part of a risk adjustment system, such as utilization, claim dollars, or procedure code information. In my mind, that's really how I distinguish the two.

A predictive model is something that was intentionally designed to predict claim dollars and doesn't have any restrictions on what information you can feed into it. A risk adjuster is something that was really designed to predict health status for payment purposes, and as a result there are significant restrictions on what information is appropriate to use for that purpose.

I often put predictive models into a couple of categories, but there is some blurring here. One category is population-based predictive models, where you are trying to predict claim dollars for an entire group of people or entire population. That type of model would be appropriate to use for underwriting or rating purposes. Another category of predictive models, used in particular by the disease management firms,

is focused in on targeted conditions. That is, these are models that predict whether someone might be an expensive heart patient or diabetes patient and so forth, and the purpose behind those models is generally for care management or case management.

I will finish with a mention of some of the firms that have advertised or have developed predictive models.

Integrated Health Care Information Services (IHCIS) has developed a predictive model for underwriting and rating purposes as well as care management purposes, and this model is built off of the Episode Risk Groups (ERGs), since this company developed some of the logic behind ERGs for Symmetry. IHCIS has developed a risk adjuster that is really for payment purposes. They also have a separate model that is developed for predictive modeling that uses additional input that increases your predictions. Ingenix has developed a set of predictive models, and Mark is going to talk about some of those models today and how they're used at Humana.

Landacorp also has models that they use for disease management purposes, as well as models that are appropriate for underwriting and rating. Medical Scientists has some products like this, as does TruRisk. DxCG recently released a model that they are classifying as a predictive model. It uses more than just diagnosis code information to make predictions. A number of insurance companies have developed their own models for rating purposes. They have home-grown models that they've developed internally that they use for rating and underwriting purposes.

MR. MARK WERNICKE: I'm the small group chief actuary at Humana, and I want to share with you how we've used a predictive model, a modeling tool, to do renewal underwriting for smaller block of business.

The tool is called Procise Predict, which Ingenix produced. I'll use the term Ingenix, although that's not technically correct, as I go through this. We actually installed it in the spring of 2001 and have been using the tool to do renewals for groups of 1 to 99 people on an automated basis since that time.

Before getting started, let me tell you just a little bit about Humana. We're a regional managed care company in the central part of the United States. In other words, we don't do business on either coast, but we've got a lot of business in all the other states. We've got about one million small group members. The average case size is 10, so that means we renew about 4,000 small groups every month.

The bottom line is that we now use drug data, as well as age and sex information to do the renewals. Let me tell you about our test results. Before we installed the Ingenix software, we actually purchased it on a trial basis and we looked at 1997 and 1998 data in two of our larger states, Texas and Indiana, using what our underwriters classified as tiers at that time. "Preferred" in our terminology means

the better health risk, manual is like the standard case and then S1, S2 and S3 are different variations of substandard. So the ratings here are what the underwriters have put in place.

Then we looked at actual data and how those claims rolled out and completed in 1999 and 2000. Based on how they rolled out, we knew what the actual per month per member (PMPM) claims came in by tier. We went back then and we looked at 1996 and 1997 drug data, ran the Ingenix tool and had that model go ahead and distribute those same cases by preferred, manual and substandard tiers. We forced the exact same number of groups within each cell and then we looked at the actual results. As seen in Table 1, the model preferred group came in with lower PMPM claim costs than the original preferred group, and the model substandard group came in with higher PMPM claim costs than the original substandard group. This basically showed us for the Texas and Indiana business that the Ingenix model did a better job of predicting future health care costs than the underwriters did.

Table 1

Test Market Results

	Underwriting Tier	Procise Predict™ Tier
Preferred	$109.96	$103.69
Manual	$129.24	$120.54
S1	$140.83	$154.58
S2	$191.20	$139.02
S3	$182.00	$217.95
Total	$128.05	$128.05

We also learned that the underwriters had a tendency to look at any case having a large claim and simply calling that case substandard. So the larger cases in the small group segments, the cases of 40, 50, 60 lives, were more likely to get called substandard. In other words, they were less preferred or manual. Ingenix didn't view that. Ingenix recognized the fact that a large claim for a larger group would have more premium dollars to spread that over. So, we saw a switch in the proportion of larger cases that were classified as better health risks using Ingenix. Now we've had it in place for a while, and a year later this is kind of a comparison

of actual renewal results (Tables 2 and 3). The benchmark refers to how the underwriters had historically classified the business and the proportion of the group that fell into the various health status tiers. In terms of groups, what has happened is that Ingenix has more groups showing up as substandard. More of the micro groups actually show up as substandard. So on a group basis, Ingenix has more groups being substandard, but on a member basis, if you distribute that by members and you follow that logic, the Ingenix methodology actually produces more preferred members and fewer substandard members. That's the difference in how the results pan out.

Table 2

Distribution of Groups - May 2002 Renewals

	Benchmark	Procise Predict™ Tier
Preferred	41.7%	40.8%
Manual	27.6%	20.4%
Substandard	30.7%	38.8%

Table 3

Distribution of Members - May 2002 Renewals

	Benchmark	Procise Predict™ Tier
Preferred	35.7%	43.0%
Manual	28.8%	25.2%
Substandard	35.5%	31.9%

So why have we used Ingenix? We think it does a better job of predicting costs versus premium than the underwriters did. We also learned, as I've pointed out, that it does a better job of appropriately assigning risk by case size, whereas the underwriters have a tendency not to account for extra premium. It also accounts well for age and sex demographics.

I have a chart of some history of Humana data by case size over time (Chart 4). We've run every member through the Ingenix model and then sorted them by case size. Not surprisingly, as case size decreases, the health status or the health risk of those cases increases, so our one- or two-life business clearly has a much higher risk and a much higher health status score than the other states. That's a quick summary of how we're using Ingenix, what the results are and why we think it makes sense to Humana.

MR. IAN DUNCAN: I'm an actuary and vice president of strategic development for a company called Landacorp. A couple of years ago, I went to work for a small startup in New Jersey. We were an actuarial consulting firm and we did predictive modeling, and we sold our company to Landacorp, which is a provider of many hospital-based case management systems and disease management systems. So they assembled an interesting group of different skill sets with us on the actuarial and prediction side. They also do automated disease management.

I'm going to talk about a couple of things. One of them is obviously prediction, and another is how we use prediction in the disease management and care management business, how it all relates, how you can automate it and so forth.

I want to set the stage with a couple of quotations that I think are interesting and help to frame what I think prediction is about. We've talked about it in a risk context, but there's more to prediction than just risk. If you read what Leonard Schaeffer, chairman and CEO of WellPoint Health Networks has to say about focusing on the patients who need care, he says there something like eight percent of our members account for 70 percent of our costs. We've all heard this. Eight percent of WellPoint's members, however, amounts to over a million people, and none of us has the resources to manage that many people, however needy they may be. So you need some form of triage, some form of focus to find those people who really need the help and who are going to produce some kind of return on your investment. In our company we're very concerned and interested in trying to focus on those members who need help and who are going to produce an ROI.

An anonymous health-care executive said that getting the right members involved in a program is really important. We had this program in which we identified 2,600 people who were at risk; we predicted that they needed our care. At the end of the day only 76 people enrolled. You can have the very best prediction and the most sophisticated data mining, but if you're only going to enroll 76 people, it's just not worth it. So you need to expand your risk predictions by a whole set of other kinds of tools and various ways to make sure that you drive people, that you manage people, and you drive that ROI.

I don't think that there are necessarily very clear definitions of what people mean by prediction when they use that word, and we come up against people in the marketplace who say, "Well, we're predicting that this group is going to have high costs," when we think they're actually segmenting the data into different population groups. As an example, I can identify everybody in a population who is, for example, diabetic. It's certainly true that that sub-population is going to have higher costs and exhibit greater risk than the population at large in any year, but I don't think of that as being particularly predictive. I would call that identification or segmentation.

We think of prediction in our company as the process of assigning some kind of probability or some kind of risk score to the people who are going to experience whatever outcome it is that we're talking about. It could be hospitalization, it could be cost dollars or any one of those things.

Table 4 shows some numbers that get to the idea of what's different about segmentation and prediction. This is a population from a 300,000-member regional HMO for 1998 and 1999. We've separated people into three categories based on their historical cost. You find about 90 percent of the members are not using resources—they are very low utilizers. There's a moderate group who experienced, on average, between $5,000 and $6,000 in costs, and then you have these people who have some kind of acute episode and are experiencing very high costs.

Table 4

Distribution of Costs and Population by Category - example

	Distribution of Members and Claims Projection Period			
Historic Period Group	Historic Period Cost	$0 - $2	$2 - $25	$25+
Acute				
$0 - $2	$ 324	$ 327	$ 5,368	$ 46,836
87%		90%	10%	0%
		90%	*64%*	*40%*
$2 - $25	$ 5,658	$ 668	$ 6,599	$ 47,811
12%		55%	40%	5%
		10%	*34%*	*40%*
$25+	$ 49,032	$ 847	$ 9,609	$ 58,489
1%		26%	46%	28%
		0%	*2%*	*20%*
TOTAL	$ 1,230	$ 355	$ 5,851	$ 49,377

In a traditional case management environment in which the medical management people are looking at who to manage in any period, they would tend to focus on the high-cost populations. It's a small enough population—you don't have the resource constraint problems. When you talk to people who do this for a living, you learn that they know all of their members, they know the members' spouses, they know the members' children. They know these members very well because they're managing them quite intensively.

But look at what happens to those people in the next year. Twenty-six percent of the people who had very high costs in the prior year in fact regress totally down to the means. Their costs in the second year are under $1,000, on average. About half of this group experienced costs on average around $10,000, and only 28 percent of the population stayed at about the same level that they did in the first year. So if you're managing in an unfocused, untargeted way without doing some kind of segmentation or prediction as to who's going to have the events in the second year, I would say that you're probably wasting some of the resources that could be devoted elsewhere in the plan.

Where would you devote those resources? Well, you wouldn't naturally think to look in the 87 percent of the population who are low-cost utilizers in the prior year, and yet if you look at the numbers, forty percent of the high-cost population in year two actually came from a population that on average cost $300 in the prior year. In our predictions, we're trying to find those people who aren't generally on the radar screen, but who are still going to account for a fairly significant amount of

dollars in the second year. We do that through a fairly intensive data warehousing process.

Some of you may have been through session 11L yesterday, "Data Mining Techniques in Actuarial Modeling." We used the same kinds of tools and algorithms that the instructor was talking about. What makes our process a little different is that we're not applying these tools directly to the administrative claims data as we get it. We create a data warehouse that includes a fair amount of derived variables. What do I mean by this? Derived variables include things like the principal diagnosis of the individual, the patient's principal treatment regimen and the patient's medical position ratio. Medical position ratio implies the percentage of prescriptions that they could have filled over a time period versus what they, in fact, did fill. These are all markers that indicate the degree of various behavioral factors, like compliance, that are quite important when it comes to being able to predict who out of any group is going to have an adverse event in the following year.

We often ask if this is a clinical process or a statistical process, and the answer is, it's both. You obviously need clinical people involved because you're going to be building definitions of diseases, definitions of treatment patterns, finding gaps in care and so on. All of these are clinical events. At the same time, there's a lot that you can do with data mining, as the instructor pointed out yesterday.

We primarily use three different statistical techniques for our modeling. We use stepwise regression models; we use a decision tree model that's a home-grown version of the pi-squared automatic interaction detection algorithm; and we also have our own home-grown version of a neural network.

When you talk to your audiences about this kind of issue, you often get asked: "I've got all this data, you've got whatever algorithm and model, why can't you just hook up your model to my data, press a button, come back in a week and you will have found all my high-risk people?" Unfortunately, it doesn't work like that, as we all know. It takes a lot of analysis and a lot of development of models, testing hypotheses and so forth.

The output from our process is a prioritized or ranked list of members. There are going to be a few people who have a high probability of having whatever the event is that you're trying to predict, and as you increase the number of people, the probability goes down until eventually you get to the average incidence of whatever it is in the population.

What we then do with that data, I think, is really important. Up until this point, we've been doing prediction or identification, but this is of no value unless you do something with that information. So we have a process and Chart 2 shows how that process works. What we have done through our data mining and prediction process is identify people who are at risk and are a disease burden. But there's a

second dimension to this, and that has to do with information about the patient that you just cannot get through the administrative data. Through disease management programs of various types, you can interact with the high-risk members and weed out all of the people who are there as false positives. You're never going to be able to deal with that issue through the data; you need some other kind of intervention to identify those people. Find out what their self-management skills are and then get to where you really want to be with this kind of program, which is get them back, get them compliant, have them do the right thing and reduce their potential risk.

Here are two examples of prediction. Chart 3 is an example of a project that we worked on for one of our customers who wanted us to be able to take our prediction from the administrative data and marry it with real-time incoming pre-authorization data as people were being admitted to the hospital. The reason for this is that they wanted to be able to prioritize the case management nurses' task list on a daily basis. We developed some models that assigned a risk score of one, two or three to people based on the combination of historical claims data that we had available, as well as information coming in about what the admitting diagnosis was, what institution they were being admitted to, what day of the week and so forth. It's not a great model. The percentage of people who we were able to correctly identify isn't as good as we'd like it to be, but it shows that you can deal with the problem in real time.

The second one is one of my favorites (Chart 4). We studied drug interactions in a 50,000-member population. About one percent of that population is on two drugs that, in combination, are very highly predictive of adverse events. Based on the data, about 22 percent of the members who are on that combination of drugs had a hospital admission in the following month. This is not the following year; this is the following month. So, through this kind of analysis and this kind of work, we're able to update our models and include into our derived variables additional algorithms that we then use to update and refresh what we're doing on a regular basis.

We put this all together, because, as I said, we're very ROI focused in what we call the population risk management savings model. The purpose of this model is to allow our customers to identify and focus on the key variables that really drive ROI in a successful program.

We've talked about the identification of at-risk populations: how you do that, how you stratify the populations, how you identify the level of risk and so forth.
But in addition to that, there are some other very important factors that you need to take into account. You're going to be doing interventions on these populations, so you need to know how much it costs to deliver an intervention. There are different types and intensities of interventions. Are they automated? Are they nurse-based? How productive are the individuals or the resources that you're putting into interventions? How successful are you at connecting with the target

patients? Do you connect with 76 out of 2,600, or can you get more people to enroll in your program? And finally, how effective are your interventions? We have an interactive model where you can change the different variables and you can see how, for example, intervention effectiveness is going to affect your return on the program, compared with, say, the rate at which you can get people to engage in your program.

I'll end by showing you some more results. I said earlier that the result of the prediction process is what we call a yield curve (Chart 5). The yield curve simply ranks all members of a population from the most risky down to the least risky, and assigns probability to each one of those. The most risky people in that cohort, on average, are experiencing about a one in two probability of having the adverse event, and for the population as a whole it works out to be about one in eight or about 12 percent.

We're often asked, "What factors are in your model? What factors are driving you to say that those are the risky members?" Well, here are a couple of them that tend to come up from time to time (Table 5). In this particular instance, we have a high-risk targeted population, and the gender breakdown between males and females is about the same as in the population as a whole. The average age, however, is much higher. The average number of co-morbid diseases present in the population is much higher. Because this is going to be sent to a customer who said he uses it for intervention, I put the prevalence of a couple of diseases out there.

Table 5

Comparison of Target and Non-Target Patients

Characteristics of Patients with 1999 Cost < $2,000

	Male	Female	Average Age	Avg. Num. of comorbid diseases	Disease Prevalence: Diabetes	Disease Prevalence: CHF
High Risk, Targeted, No Intervention N=1,107	58.9%	41.1%	58.7	2.9	59.3%	2.0%
Not-Targeted (Low Risk) N=171,071	54.6%	45.4%	32.2	0.5	1.7%	0.1%

What kind of outcomes did we have with that particular set of data? Well, what we have here is our targeted group (Table 6). This is made up of the 1,100 people that are the riskiest people in the population. Average cost in the historic period was about $1,200. Average cost of that population in the prediction period is slightly less than three times the historic period. For those members who had claims over $2,000, there are about five times their historic costs and that compares with the population as a whole, $430 going to about $1,100 and about the same amount for those people who do experience an adverse event, about the same amount of average cost. The point here is that you've improved your chance of finding through prediction. You've improved your chance of finding those high-risk people by about 3.5 to 1, compared with drawing those people at random out of the population.

Table 6

Outcomes of Targeted Patients

Year 2000 Outcomes of Patients with 1999 Cost < $2,000

		1999	2000, all members	2000, members w/claims >$2,000
High Risk, Targeted, Untouched by Intervention N=1,107	Avg Cost	$ 1,278	$ 3,176	$ 6,602
	% with Cost >= $2,000	0.0%	39.7%	100.0%
Not-Targeted (Low Risk) N=171,071	Avg Cost	$ 432	$ 1,108	$ 6,033
	% with Cost >= $2,000	0.0%	12.2%	100.0%

19

I think that this has been a quick and brief overview of how we put together programs, how we do prediction and how we think about prediction. What we're doing is very focused on return, very focused on the need to combine certain elements, the data element, the prediction element and the execution. Without those three things working together, you're not going to have a successful program.

MS. ELIZABETH LEWIS: We're going to switch gears a little bit. I'm going to talk more about the risk adjustment part of it and what my company did. We took a risk adjuster and, as Bob said, we used ACGs and did some nontraditional applications. Nothing was very complicated, but we were very happy with the results. We did an area factor analysis on our small group block. We used it in

some analysis to break down our increases into utilization, morbidity and cost. We're currently working on using ACGs to look at the relationships between benefit levels and morbidity, and then also one of our groups came up with a benefit design in which they used ACGs as well.

First we'll talk about what we do with area factors. Here's a little background in case you don't do business in Minnesota. In small group, which is two to 50 lives, they allow us to have three different geographic regions for area, and then they can be 20 percent apart. For years we did it all the same way: we'd get the PMPMs by county, and we'd age and sex adjust them in the traditional way and draw the lines to where it made sense for the area factors to be.

So one year we did that and we sent it in as usual and the regulator sent it back and said she was looking for "improved actuarial analysis." I can laugh at that because I'm not the one that was working in small group that year. What her point was, and it was an excellent point, was that in Minnesota you can have plus or minus 25 percent of where you set a group's rate. So by putting people in that range, we were kind of already taking health status into account, and if that varied by area and we weren't adjusting it out, we could be double-hitting people.

We pulled a bunch of data by member, their claims, their month of exposure, then we got their ACG score. If you're not familiar with ACG, it looks at the ICD-9 (International Classification of Diseases) codes and puts them into categories based on what claims they've had, their age and their sex. It's actually on an individual basis, not on the episode of care, but a member-by-member score. The other thing that we collected, of course, was the zip code, because you really can't do a lot with area factors without that.

After that it was just a simple math exercise where we calculated an average ACG by county and overall, and then we adjusted each county's PMPM to the average ACG in the state so that they were all on a level playing field, and we were left with PMPMs by county where the differences were only for cost and utilization. So now we could reflect the health risk in our other variable, as we should.

Now we have all these PMPMs by county that are nicely risk adjusted, and we had to get them into three factors. So just to start, we put them into 31 regions. We have a lot of Minnesota maps that we colored in and broke down into pieces where we get 50,000 to 60,000 member months in a region, and the counties are touching each other. Then we sorted them from high to low, and this is where we met with our sales department and said, "Okay, we have to draw two lines to come up with three factors for the ones that are on the edge. What do you know? What makes sense? We didn't want to be too disruptive. We didn't want to cause any more problems with competition, which is always going on.

The results were interesting, comparing our age and sex adjusted to our ACG adjusted. About half of the counties stayed in exactly the same area with the

method, which shouldn't be too surprising when there were only three to begin with, so there's not going to be a lot of movement. About 16 percent of the counties moved to a higher area when we health status adjusted them, and that means their morbidity was better than average, so when we removed the impact of that, they were actually in a higher-cost area, so they had to go up. Nobody had to move more than two areas. A third of the counties actually got to go to lower areas and a few of those actually got to go down in two area factors, so sales was thrilled about that and our regulator was thrilled as well.

The next thing we did was trend analysis. With trend, we did the typical thing: year to year you run the PMPMs, you age and sex adjust it, and you see how much is it going up. We could split out what portion was due to the cost portion of it and what part was due to utilization, but we didn't really know where the case mix was, or what the change in morbidity mix was, or anything like that. We started to see double-digit increases, and they were going on for a while. We didn't think they were going to stop, so we said we need to get more information; we need to change our methodology. That's where the risk adjuster came in again. So what we did here is, as before, we determined the portion of trend that was due to the change in the cost. . We used a repricing methodology on all the claims and we repriced 1999 using 1998 fee levels so that we had new PMPMs and could determine the amount of the cost difference that is attributable to the change in our provider contracts.

Then this is where the risk adjuster came in. It's a simple process. Member by member, we assigned them an ACG and we looked at between the two years how that average ACG changed. That was the change in our morbidity. Then, what was left over, of course, was the change in utilization.

What we learned from this is very interesting. Age sex accounted for only one percent of the cost increase, and we found that morbidity was 30 to 40 percent. The range is because we looked at it from different market segments and products. The open access was on the higher end of the change in morbidity. So, of course we saw that and thought, "What in the world could be going on here? Could people really be getting that much sicker?" A new thing we're working on is asking if it is something with the level of benefits. We have seen a lot of people in the last few years adding preventive services, and we could tell in other trend studies that we had done that the preventive services had increased where we were seeing increases in the PMPM. So we're thinking a lot more people are getting the preventive services, and as a result things are being found earlier. Is that what's driving the increased morbidity? We're trying to dig into that more to see if that's what's going on. Now, of course, if that's the case, hopefully in the long run that will help us. If they find it sooner, maybe they'll solve the problem sooner. Also, if the morbidity is actually the problem, an enhanced disease management program may help in this situation.

The last thing we have done with ACGs was for one of our largest self-funded

accounts and their consultants, who decided they wanted to have a benefit plan with tiers in which you get the richest benefits when you use the most efficient providers. What's interesting about this is not the process, but what we learned in the case. Each health plan that wanted to be part of this sent in their data, all kinds of detailed data. The consultants then did all the ACG adjusting for the differences between the status between the plans, and each plan and provider was assigned to a different benefit case based on the results. What's interesting is, because we're all in Minnesota and we all have different contracts with different providers, there are a couple of providers that actually fell into more than one tier based on what health plan they were with, which was interesting to explain.

So what we learned was that providers did want to be in the first tier. We actually had a couple of our big hospital groups come back and say that they didn't want to be in the third tier and wanted to talk about what we're charging, so we worked with them on that. The other thing we learned that was interesting is that at first our health-status-adjusted PMPMs were the lowest of anyone else that was participating with this group. The difference was pretty close to ten percent and we thought, "Wow, what in the world are we doing? We should lower all our rates." Then we found out that our system only sorts a few ICD-9 codes per member. So we fixed that immediately and after that we were all pretty much in line.

Going forward, I know we have a research department in our company, and they do a lot of stuff with risk adjusters. They are the ones that are working on the analysis of what's going on. For example, are the benefits driving utilization, which then makes the morbidity look worse? They're also helping us with the utilization factors that we have for deductibles and co-pays and such. Right now, we use data we buy from the consultants. We figured that we are a large company and we have all this data in our system, we should be able to develop our own utilization factors off that. So, we're going to risk adjust that information to get our own utilization factors, and we're hoping to do that soon.

MR. ROBERT LYNCH: Using the predictive modeling for the underwriting, you had the system where you were using pharmacy and you were comparing that with what your underwriters were doing. I'm not clear on what your underwriters were doing. Were they using a medically based debit underwriting system? Were they pulling claims over certain dollar amounts? What were they doing in your comparison method there?

MR. WERNICKE: What our underwriters did, and I apologize for not going over that, was they tried their best to look at just large claim data predominantly. If there had been a number of pregnancies, they would not assume those were recurring. Based on that large claim data, they made assumptions as to whether those sorts of claims were recurring or not and based on that information, really a subjective process, not an objective process, put groups into the various health status tiers and that's how it worked.

MR. LYNCH: So I guess the second part of my question, then, is I gather that there wasn't an effort to use diagnosis codes or common procedural technology (CPT) codes instead of prescription data.

MR. WERNICKE: We talked about that at some length, and at the time that we started to use the Ingenix model, we looked at one or two other firms that had predictive modeling tools, and none of them had that sort of data available. Some of them thought they were close, but we kept waiting and it didn't really ever pan out. Ingenix is about to release a new tool where they also use, the Clinical Care Groups (CCGs) and they're basically trying to incorporate diagnosis information along with drug information and age and sex data. So I think we'll be interested in looking at that as it rolls out over the next six months to a year.

MR. LYNCH: So I gather from that you were using the drug NDC codes, because that data was more readily available and more reliable.

MR. WERNICKE: It was more readily available. It was immediately accessible; there was no three-month claims lag, etc. That's right.

FROM THE FLOOR: This is sort of a follow up to that last question. Did you try and do any comparison of the efficiency of the Ingenix method versus a more traditional underwriting approach using credibility and large claim pooling?

MR. WERNICKE: Not directly. Like I referenced, we just looked at our Texas and Indiana business and what our underwriters had done, and they thought that was a good process, if you will. We can argue whether it's a good process or not and compare it to Ingenix. We didn't look at any other sort of methodologies.

FROM THE FLOOR: I agree with Mr. Duncan's concept of focusing in on the high users. With all the concerns that people have over Health Insurance Portability and Accountability Act (HIPAA) privacy regulations, do you have any idea how that might be done, so that if you do identify people that need some help, you don't run the risk of helping them and turning around and getting sued by them?

MR. DUNCAN: I was hoping to avoid questions about HIPAA, of which I am not an expert. I think that we're a vendor to health plans, which means we're always a step behind the health plans, so we're protected to some degree. When people sign up for HMO coverage, I believe, in some instances, they are giving the HMO the right to offer them programs. Certainly that seems to be what our customers are doing. We're beginning to run into some odd situations, like one potential customer who is absolutely insistent that the data not leave their premises. Why, it's not clear, but it just seems that, when pressed, they say we just think employees are going to be more comfortable if they know the data is still physically resident. I think this is very tough, and there are probably people who are better qualified to talk about HIPAA right now than I am.

MR. JOHN ALBERT: This is directed to Ms. Lewis. How significant were the changes in health status by area, and are you allowed to rate by health status?

MS. LEWIS: They weren't very significant at all, because, like I said, we're allowed to use a 20 percent difference between the areas. We actually only used ten percent because they're not that different. Probably part of the reason for that is that the metro area is forced to be the highest area, and when you look at the results it is not the highest area, so I think that compresses everything. Our regulator is the one who asked us to use this, so she was quite open. She's always very forward-thinking in asking for this kind of thing. We don't use it on a group-by-group basis exactly. We set our area factors after adjusting for the health status, but then group-by–group, we don't actually look at that, we just look at their zip code at that point.

Chart 1

Small Group Ingenix Scores

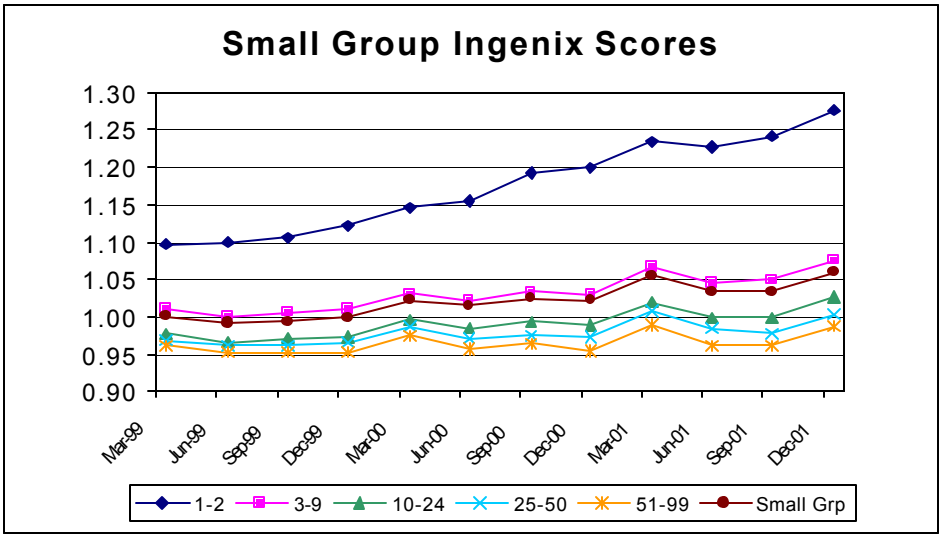


Chart 2

Prediction and Targeting leverage information about patients and their profiles

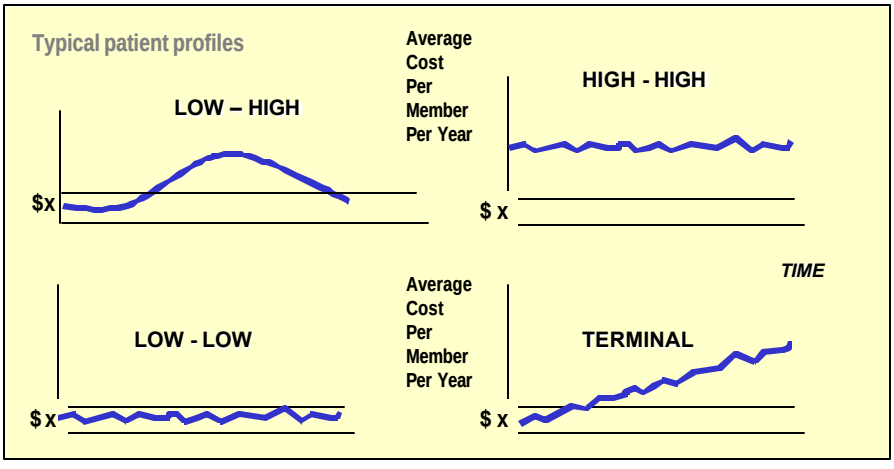
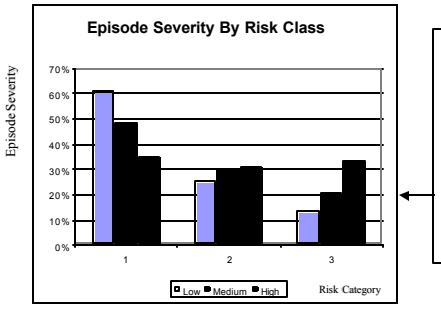


Chart 3

Prediction Example: Severe Hospital Episodes



Interpretation:
- Patients are stratified into Low, Medium and High Risk
- Approximately 15% of Low-risk patients experienced Long stays
- Approximately 35% of High-risk patients experienced Long stays.

Hospital admission surveys combined with patient claims histories stratify newly-admitted patients into three risk categories; Risk categories are highly predictive of the relative length of hospital episodes and the probability of readmission.

10

Chart 4

Prediction Example: Hospitalizations from drug-drug interactions

Prescriptions and Subsequent Hospitalizations in 1999

	Patients*	Not Hosp Final Day of Month	Hosp Next Month	%
SSRI	41,837	41,756	1,310	3.1%
Warfarin	14,749	14,614	2,102	14.4%
SSRI / Warfarin	598	583	129	22.1%

Prescriptions and Subsequent Hospitalizations in 1999

	Patients*	Not Hosp Final Day of Month	Hosp Next Month	%
Glyburide	20,984	20,914	1,365	6.5%
Ciprofloxacin	12,142	12,042	700	5.8%
Glyburide / Cipro	250	245	38	15.5%

Presence of certain drugs in combinations is HIGHLY predictive of hospitalizations – as early as the NEXT MONTH!

11

Chart 5

Typical Yield Curve