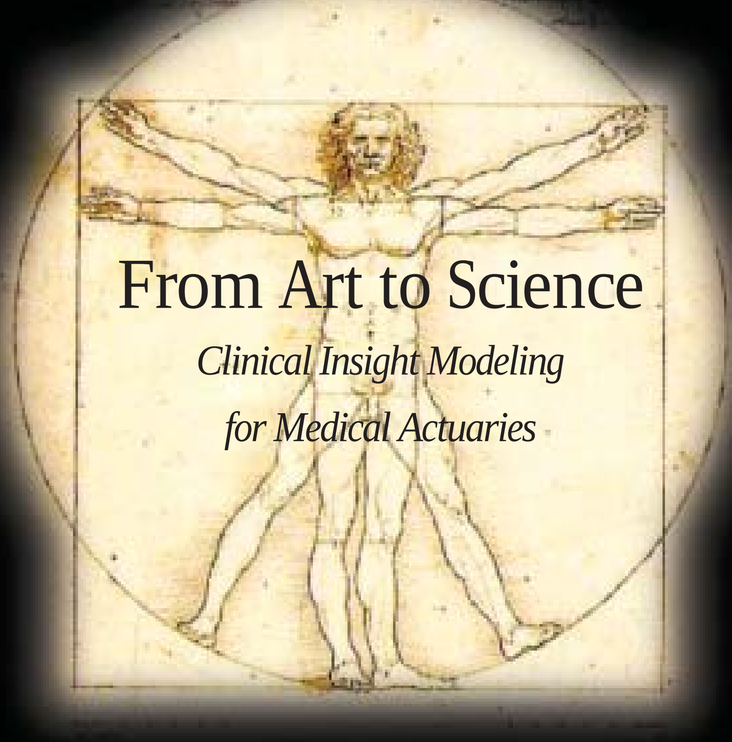


# From Art to Science

## *Clinical Insight Modeling for Medical Actuaries*

**By Harry Poteat, M.D.**

***Combining clinical insight with actuaries' understanding of mathematical patterns in the population works fairly well for predicting catastrophic risks in medical care. But it could work even better.***

*Actuaries who practice in the medical arena* tell me that what they do is more art than science. As a physician who builds predictive models for HMOs and insurance companies, I'd have to agree. In catastrophic risk prediction, standard statistical models often don't apply. Technology moves so fast that what made medical treatments expensive 5 years ago may not make them expensive today. And even if the types of expense (say transplant or neonatal) remain similar, the case rates and severity for cases in these areas are constantly in flux. So how do we turn an art into a science?

First, build reproducible, objective processes into our models. Second, make sure we use all the predictive information we have at our disposal, especially epidemiology. And third, validate our models to see how accurate they are. I call this loose confederation of ideas and concepts "clinical insight modeling."

To make a process reproducible and objective, you have to take advantage of today's computer technology. Our predictive model software is a Markov state transition (MST) model. The MST model is designed to function in data-poor environments and uses a benchmark database we build through a process we call triangulation.

For data-rich environments we use powerful computers to look at each piece of data carefully and see if the data forms patterns that let us predict future costs of individual patients (claimants). We call this data mining process rational artificial intelligence (RAI).

### Epidemiology Can Help

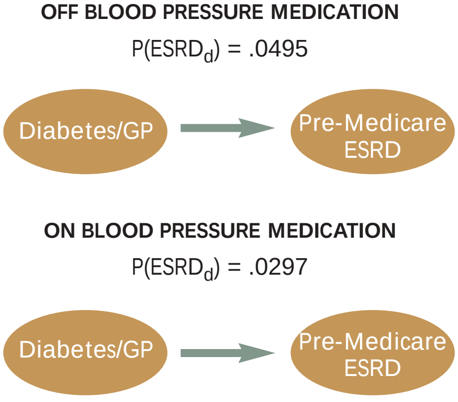
In many ways, our industry uses all the predictive modeling information available to us to get answers. It's not at all uncommon, for example, to engage a nurse in renewal underwriting for small groups, and many reinsurance firms use doctors and nurses to manage cases for them. This clinical insight is then combined with the actuaries' understanding of mathematical patterns in the population and experiences gleaned with other similar populations, as well as industry trends and costs over time. This isn't a bad system, but it could be better.

**TABLE 1**

| GEOGRAPHIC REGION | ESRD CASES PER 100,000 MEMBERS |
|---|---|
| Northeast | 4.21 |
| Mid-Alantic | 6.16 |
| Southeast | 6.85 |
| West | 4.34 |
| Midwest | 4.72 |
| Southwest | 6.56 |
| All United States | 6.38 |

*Legend: Claims data such as that present in the MedStat market scan database, expert opinion, and the epidemiology of diabetes are used to predict case rates for renal failure in diabetics. Obesity, advanced age, and being of certain race or ethnicity predispose to the disease. Estimates shown are from the MST models triangulated database for commercially insured populations.*

**FIGURE 1:** Predicting Catastrophic Events




$P(ESRD_d)$ = Probability of progressing from gross proteinuria to ESRD

The environment for figuring out case rates for populations is often data-poor. Clients are often reluctant to provide data about patients in their health plans, and asking for data uses up relationship points. So how does clinical insight modeling work with this conundrum?

One discipline that's underutilized in predicting future case rates is epidemiology—the demographic study of disease and its prevalence in the population. Epidemiologists, for example, have long known that people of different ages and races get diabetes at different rates. Left untreated or under-treated, diabetes often progresses to end-stage renal disease (ESRD) and kidney failure.

In addition, health plans have long known that younger people on dialysis get transplants while older people tend to stay on dialysis. Health plans also know that after 33 months or more, the liability for these patients can be shifted to the government payer, the Centers for Medicare and Medicaid Services.

An MST model provides a reproducible, verifiable estimate of ESRD case rates caused by diabetes. Epidemiological risk factors for diabetes can be translated into an overall case rate for the disease, and knowledge of disease progression can give us some ideas about case mix severity.

A good model can also relate catastrophic risk to levels of medical management. It has long been known that control of blood pressure in diabetics radically reduces their risk for ESRD and that good control of the sugar levels in their blood also correlates with fewer cases of kidney failure. These relationships are well established, yet it's unlikely today when actuaries build up projections for specific excess of loss that they're incorporating such information as adjustments to their models.

Because the MST models are Markov state transition models, they require some knowledge or understanding of the probability an event will occur. If, in addition, one knows how a new drug or treatment will affect that probability, it's possible to explicitly model the impact of this new technology.

For example, the probability of a diabetic developing kidney failure when not taking a specific medication is 4.9 percent per year; when taking the medication the probability is 2.9 percent per year. This is important information for the actuary, because knowledge of how the diabetic population in an analyzed group is being managed will influence future cost predictions. And kidney failure is very expensive.

## Validation

Validation is difficult for a number of reasons. First, it forces us to admit that no model is perfect. The problem with reproducibility is that it allows us to make the same mistake twice if we're not careful. And the great blessing of subjective human error is that we never have to admit to making the same error twice because no two subjective opinions are ever identical.

This is a trap that in some ways can make us artisans and not scientists. In addition, confronting our own lack of perfection is very dangerous for our sales efforts. Given the immense variability of biological systems, the unpredictability of everyday life, and the irrational decisions consumers and their doctors make every day, validation leaves us to confront imperfection head on. But if we're to progress as a science, this must be part of the process.

**TABLE 2:** Validation at a Commercial Client

| Condition | Actual Pm/pm | Predicted Pm/pm | % Difference |
|---|---|---|---|
| Transplant | $ 0.72 | $ 0.70 | +1.6% |
| Neonatal | 8.38 | 8.62 | –2.9% |
| Heart disease | 14.80 | 15.31 | –3.4% |
| Diabetes | 1.81 | 1.24 | +31.4% |
| $N...N_{23}$ | | | |
| Total pm/pm | $130.22 | $130.39 | –0.1% |

FIGURE 2: Diabetics in Medicaid Population Decreasing 40%

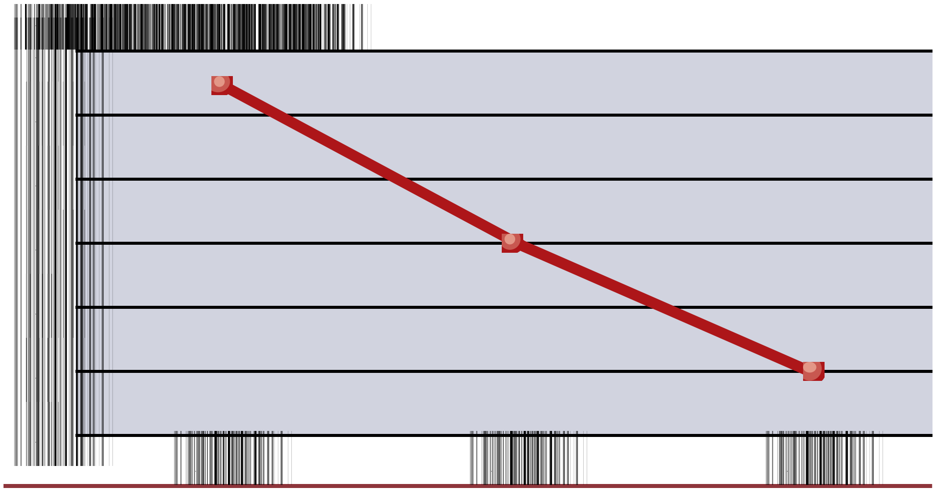


FIGURE 3: Total Spending for Diabetics Decreasing 6%

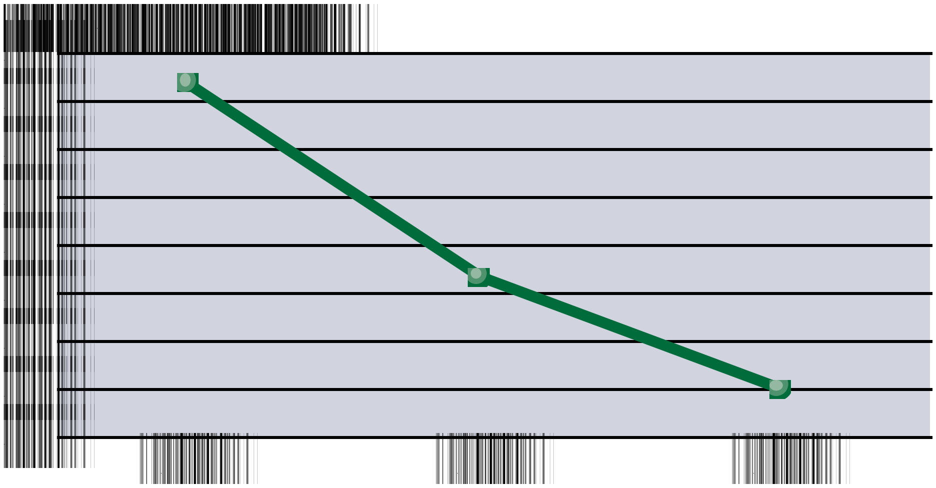
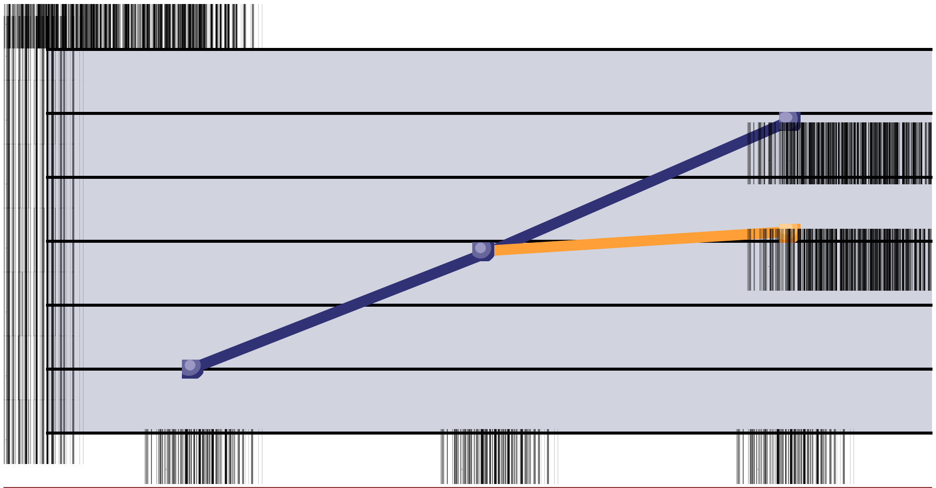


FIGURE 4: PM per Costs with and without Management



By necessity, most modelers use the past to validate their models of the future. Imagine using year 2000 data to predict the consumption of ciprofloxicin ("Cipro") in 2001-2002. Imagine validating a prediction for the use of mechanical hearts in 2005 based on year 2000 data. In either case, validating one's models against the past seems subject to potential inaccuracy.

To avoid the trap of using the past as an inexorable guide to the future, models should combine claims experience, expert opinion, and medical literature with historical data. We call this triangulation.

A recent example of medical cost projections in a data-poor environment is summarized in Table 2.

In the real-life example below, we show how an MST model was used to project case rates and severity with and without aggressive cost containment measures. Figure 2 shows that the number of diabetics under management is expected to decrease over time. In fact, this has been the clients' experience. The management cohort has decreased from death and disenrollment.

The second figure shows a projection of total medical spending for diabetics, which has been similar to client experience.

Left unmanaged, the natural progression of diabetes causes an increase in case severity, while the total number of cases in the cohort decreases. Market forces also increase pharmacy costs as new drugs to treat the disease are released into the market. Actual client experiences after introduction of an aggressive management program are shown below (yellow line).

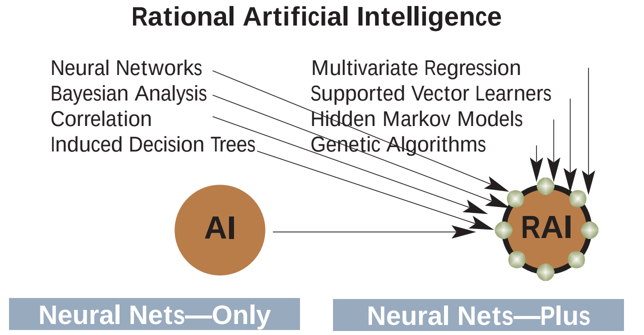
The figure above indicates a projected increase without cost containment (35 percent over 2 years) in PMPM costs (purple line).

Clinical insight modeling using MST may be effective in determining a disease-specific trend. In this client example, the models would benefit

**TABLE 3:** Association of Prior Year's Claims with Future Year's Claims in Dependents

| 1998 | | 1999 | |
|---|---|---|---|
| Case Costs 1998 | No Cases 1998 | Cases over $10,000 in 1999 | % of Group |
| <$1.50 | 6,144 | 174 | 18% |
| $1.50–$800 | 16,667 | 226 | 24% |
| $801–$12,000 | 7,198 | 406 | 42% |
| >$12,000 | 518 | 155 | 16% |
| | 28,527 | 961 | 100% |

**FIGURE 5:** Technology Overview




from comparison to diabetes costs in other Medicaid programs to rule out fluctuations in trend not associated with management initiatives.

#### Working in Data-Rich Environments

A data-rich environment is one that includes renewal underwriting or case managers whose job it is to improve health and reduce patient costs before these become catastrophic. These skilled nurses are a limited resource and must be allocated to those patients likely to become high-cost next year. As we see in Table 3, being expensive in the prior year is rarely a guide to becoming expensive in the subsequent year. The table below indicates that fully 45 percent of the patients who cost more than $10,000 in 1999 cost less than $800 in 1998.

One of the problems with the world in general and our medical industry specifically is that the people who sell beer can tell you more about the likelihood that somebody you're insuring will buy their beer than you can tell about whether that person is likely to get sick or become high-cost next year. It's somewhat humiliating to realize that beer companies are better at mining data about their drinkers than we as doctors are in mining data about the patients and individuals our health plans insure.

Why is data mining hard in the health insurance industry? We don't use central data repositories as often as we could, and we must be satisfied with silos of information divided into pharmacy, provider, and institutional costs. And our industry has

**TABLE 4**

| Model Type | Computing Time | Non-Linear | Complexity |
|---|---|---|---|
| Logistic regression | Low | Poor | Moderate |
| Neural net | High | Good | High |
| Decision Tree | Moderate | Fair | Low |
| The Nth Model | N' | N'' | N''' |

***LEGEND:*** *Let the Nth model be an amalgamation or "hybrid" of several other models.*

been slow to innovate. The impact of government regulation is also salient. So let's examine a recent innovation.

The first generation of data-mining products in health care relied nearly exclusively on the use of a predictive technology called neural nets. Neural nets attempt to duplicate the way the brain thinks about problems and pattern recognition. Neural nets are excellent at dealing with nonlinear relationships.

For example, the relationship between catastrophic neonatal risk and median income is decidedly nonlinear. As sad as it may be to talk about, very poor families have very few expensive babies. Often this is an issue of not being able to access emergency medical care when it's needed. Poor families often live in cities and while they may not be able to afford co-pays for prenatal care, they'll get access to emergency medical care when it's needed.

As a result, this "high-end" of the lower class has a high rate of catastrophic neonatal cases. Rates for catastrophic neonates are the lowest in the middle class, in part because these individuals take pregnancy as a serious responsibility. Rates increase again for the upper classes, many delaying childbirth into the late 30s and early 40s and some using fertility drugs to aug-

**FIGURE 6:** Model RAI Disease Associations with Future Expense

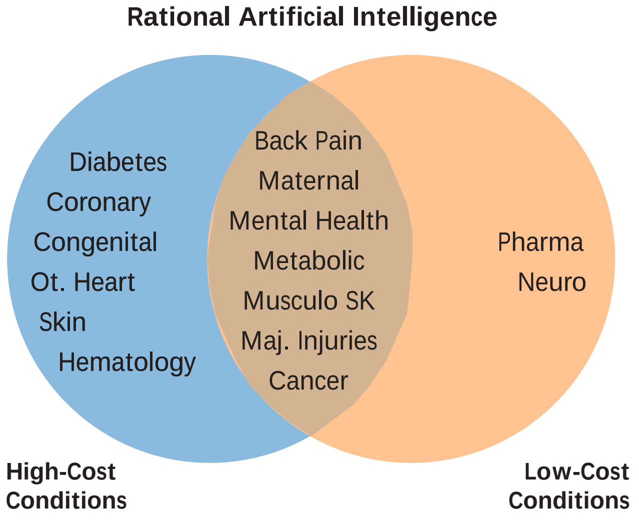


**TABLE 5**

| Class | Patients | Mean PMPM | Median PMPM | CV | P Value |
|---|---|---|---|---|---|
| 1 | 4,428 | $1,072.00 | $702.00 | 157.74% | <.0001 |
| 2 | 5,600 | 808.33 | 308.50 | 423.80 | <.0001 |
| 3 | 6,217 | 461.85 | 350.00 | 121.23 | <.0001 |
| 4 | 2,666 | 630.42 | 336.00 | 260.14 | <.0001 |
| 5 | 7,559 | 262.08 | 112.00 | 478.31 | <.0001 |
| Total | 26,470 | | | | |

***LEGEND:*** *Historical validation of a rational artificial intelligence model for identifying high-cost cases in a population of Medicaid patients. Class 1 has the highest risk, class 5 the lowest risk. Mean and median per member per month (PMPM) payments in the year following application of the decision rules are shown. CV is the coefficient of variation. It's the mean divided by the standard deviation.*

ment their chances of having children. Fertility drugs often lead to several children being born at the same time. Children born as part of a multiple birth cohort have lower birth weights and often end up in the neonatal intensive care unit.

Is there a relationship between median family income and neonatal catastrophic risk? Yes. Is the relationship linear? No. Is the relationship likely to be found using traditional modeling methods? No. Is it likely to be found using a neural net model? Yes. As Table 4 shows, neural nets excel at dealing with nonlinear problems, yet they take a long time to come up with the answer. And the answer is relatively complex. Therefore, it's hard to derive clinical or underwriting insight. It's a black box.

Commercial systems, which use only neural net technology, take months to train.

The idea behind rational artificial intelligence or AI (some have called it second-generation AI) is that you don't use neural net technology to solve every problem. Linear problems can be solved using a minimum of computing power. Often we find that the number of years a person has had diabetes is directly proportional to his risk for developing catastrophic complications from the disease, a linear problem. Age and risk for prostate cancer are linear over certain broad ranges.

So in a rational world we use the simple models to attack these simple problems, and the complicated models to solve the difficult problems such as median income and catastrophic neonatal risk. The result is that we can build predicative models faster than with first-generation AI. We can keep RAI models updated and current more easily. We can keep the models simple enough to draw underwriting inferences from them.

### RAI in Action

The medical director of a *Fortune* 100 company wants to know what's driving medical costs at his company. He's willing to accept the hypothesis that last year's high-cost employees and dependents weren't going to be responsible for the majority of his high-cost cases in the coming year (see Table 3). One can use RAI to build a model to predict which patients were going to exceed $10,000 in costs in the subsequent year.

With access to administrative claims, pharmacy, workers' compensation, and disability claims for 27,000 covered lives, one can see a strong correlation between patients costing less than $801 a year and being low-cost in the subsequent year. If the majority of the patient's medical bills were for pharmaceutical or neurological claims, this correlation was particularly strong.

In contrast, for patients costing more than $801 and having the majority of their medical costs for congenital anomalies, heart disease, blood disease, or skin diseases, there was a strong correlation with high cost in the next year (see Figure 6).

These clinical insights prompted the medical director to lower the co-pays for certain drug classes that he deemed likely to have potential economic benefit to his company. Among these were antidepressants.

In a second example, a state government is anxious to assign those patients likely to have very high costs in a subsequent year to case managers and have them looked after by disease management companies. Given rather complex interactions among socioeconomic status, disability, and ethnicity, the group feels that using RAI to find patients for management would be effective.

Here, many of the standard predictive elements predisposing to high costs were uncovered—having had a hospital admission, or having exceeded a specific cost threshold in a prior year. The RAI is successful in finding a small cohort of patients who, at least using historical validation, were predicted to account for a much higher level of medical costs. This separation of patients is superior to that which is achievable using standard statistical methods.

In summary, clinical insight modeling is the:

- Use of models that can explicitly account for changes in medical technology and treatments;
- Use of all available data, especially epidemiological information about disease states;
- Use of models that provide reproducible results and can be validated, with validation information being disclosed and published periodically.

We continue to evolve medical cost prediction from an art to a science and hope to make clinical insight modeling a key part of such a change in our chosen field. ●

**HARRY POTEAT, M.D.**, IS VICE CHAIRMAN AND CIO OF MEDICAL SCIENTISTS, INC. IN BOSTON.