

## Assessing Predictive Modeling Tools for Pricing & Underwriting

Ruth Ann Woodley, FSA, MAAA and Marilyn Schlein Kramer

In today's world of soaring medical costs, margins for HMOs and insurers are being squeezed more tightly than ever. The present point in the underwriting cycle requires underwriters to focus on increasing membership and margins. Passing along "standard" price increases to "low risk" accounts will not be sustainable in the coming years. Thus, actuaries and underwriters are looking for tools that can improve pricing accuracy, and lead to higher earnings. Predictive modeling, a tool many companies are familiar with from its uses in disease management, intuitively seems promising.

Risk selection represents another area where the application of predictive models seems promising. Carriers have always struggled to understand how the morbidity of their enrollment from a "slice" case compares to that of the other plan offerings, and how membership within a given market compares to the population in total. The introduction of consumer driven plans and the possibility that these plans attract only the healthiest employees makes this an even more pressing issue today.

Traditional underwriting uses a risk model based on age, sex, and prior cost, with some additional features on occasion, such as geographic region and industry. Predictive modeling comes from health services research and adds the additional predictive factor of diagnoses. Predictive models use information included in medical and pharmacy claims (diagnoses, procedures, drug type/dosage, etc.) to estimate future costs at the individual member level. While the terminology and approach of predictive modeling may be different from pricing models underwriters have traditionally used, underwriters should not dismiss adapting predictive models for use in pricing.

Generally speaking, groups under 250 lives are too small to have a fully credible experience base under traditional models. Past claim costs can be an unreliable starting point for pricing these groups because of the phenomenon of "regression to the mean<sup>1</sup>." Research has shown that as many as 75% of members with very high or low costs in one year will move back towards average costs in the subsequent year. This is part of the rationale for the common sales pushback that the health event driving the prior year's claims is over and will not recur. Without knowing the reason for the cost, it is difficult for underwriting to refute this pushback. As a result, underwriters turn to medical underwriting, which can be costly and inconsistent. Predictive models can be used instead to separate those members whose high costs are driven by chronic conditions, and therefore likely to continue, from those driven by random or one-time events.

Predictive models offer underwriters a low-cost means of setting cost estimates based on illness burden and prior treatments. They can predict instances of "regression to the

---

<sup>1</sup> For a discussion of regression to the mean and its implications for health actuaries, see: Henry Dove and Ian Duncan: "Actuarial Issues in Care Management Interventions". Paper 2 of a series *An Introduction to Care Management Interventions and Their Implications for Actuaries* sponsored by the SOA Health Section and available at <http://www.soa.org/ccm/content/areas-of-practice/health/research/part-1-introduction.?g11n>.

mean” and conversely situations when a member might become higher cost in the future. Increasingly, employers and benefits consultants are using them to evaluate health plan performance and set pricing levels. Finally, the results of predictive models are more intuitive than some other common rating factors like SIC codes.

But much of the large body of existing literature on the validity of predictive models can be frustrating to actuaries and underwriters. Until recently there was no research making a case for the underwriting value of predictive models. This has begun to change, with articles from the Society of Actuaries’ *Health Section News* in August 2003<sup>2</sup> and August 2005<sup>3</sup>. While these articles a huge step towards building the case that predictive models can improve underwriting accuracy, a comparative analysis is still needed of the tools available that considers both technical and practical features.

Most predictive modeling research has assessed how well the tools do in targeting patients for disease management or risk adjusting payments among large employers and health plans. It has not asked or answered the right questions for underwriting applications, so companies may need to do additional analysis and thinking in order to choose the best model to use in their underwriting.

#### **Technical Analysis:**

Most research on predictive models, and the 2002 SOA Study<sup>4</sup>, performs analysis at the wrong level for underwriting applications. Generally speaking, the research has evaluated whether the models identify those few individuals who will have the highest future cost, and whose costs can be most effectively reduced, so that their care can be better managed and those risks avoided.

In contrast, pricing tools are more concerned with costs at the employer group (or block of business) level, than at the member level. Commonly-used evaluation metrics like R-squared, which measures what percentage of the variation in results is explained by a given model, usually have very low values at the individual level. So readers conclude that these tools only explain 10-20% of the variation in results. That sounds like a poor result to actuaries and underwriters who are used to thinking about entire accounts or blocks of business – but traditional age-sex factors or a prior year’s claims experience don’t explain much of the variation in a single individual’s results either!

What is needed is additional analysis that compares predictive ability across tools for groups of 20, 50, 100, and more lives to see how results vary at various case sizes. Statisticians use the Grouped R-squared to measure the ability of the models to explain variation in expenses at the group level (account, employer, provider, etc.). It should be no surprise to actuaries and underwriters that predictive models do better at the group

---

<sup>2</sup> Ellis, Randall J., Kramer, Marilyn Schlein, Romano, Joseph F., Yi, Rong, 2003. Applying Diagnosis-Based Predictive Models to Group Underwriting. *Health Section News*: 1, 4-7.

<sup>3</sup> Winkelman, Ross, 2005. Optimal Small Group Renewal Methods. *Health Section News*: 12-15.

<sup>4</sup> Cumming, Robert B., Knutson, David, Cameron, Brian A., Derrick, Brian, 2002. A Comparative Analysis of Claims-based Methods of Health Risk Assessment for Commercial Populations. Conducted by Milliman USA and Park Nicollet Institute Health Research Center. Sponsored by the Society of Actuaries.

level as opposed to the individual member level. For example, in one typical study of the power of predictive models to identify individuals for disease management, individual R-squared results for age/sex, experience and diagnosis models were 2%, 6% and 9% respectively, compared to Grouped R-squared results (using 2-percentile cost groups) of 17%, 50% and 82% respectively.<sup>5</sup> Additional research is still needed on how these results would compare for real employer groups rather than groups of people designed to have similar costs.

Many studies truncate very high dollar claims above a threshold of \$25,000 or \$50,000. One reason this is done is to prevent a small percentage error in the predicted claim value from having a disproportionate impact on the R-squared calculation. For example, if the model predicts \$750,000 of claims for a member whose actual costs turn out to be \$850,000, that model still has done a good job of identifying that member as high risk. However, that \$100,000 of error may be very significant in pricing, especially for a relatively small group, if the costs are recurring. Truncating also removes the impact of large random claim events from future projections, but predictive models may be better able to distinguish these from large claims that are to some extent ongoing. Finally, the highest cost claims are the most important for pricing some products, like employer stop loss.

It is also common in the published research to focus on groups of members with similar claim costs (for example, the highest 10%) when choosing the best tool to predict who the future highest cost members will likely be. This approach does address the problem of focusing on member-level rather than group-level results, but does not use the kind of heterogeneous groups that exist when pricing real customers.

Finally, an analysis of whether predictive models can improve pricing accuracy needs to be put into the correct business context. The predictive power of these models must be compared to current business tools like age-sex factors, SIC loads, and pricing from prior experience. These different techniques should also be tested in combinations, not just independently. This kind of work will help users define the optimal underwriting process for different business segments. And researchers should continue to move beyond hard to understand statistics like R-squared and attempt to analyze the impact of changes to quoted rates on business results like profit margins and close ratios.

### **Practical Analysis for Underwriters:**

Research on the power of predictive models in identifying members most likely to benefit from case management also deviates from pricing and underwriting needs for reasons driven by its practical applications. Case management is focused primarily on the small percentage of members who will have the highest costs rather than the overall population. While this may be helpful in underwriting, it does not address the equally important issue of identifying members and groups who will have the lowest future costs to ensure these are not over-priced and thus lost.

---

<sup>5</sup> Ash, Arlene S., and Byrne-Logan, Susan, 1998. How Well do Models Work? Predicting Health Care Costs. Proceedings of the Section in Statistics and Epidemiology, American Statistical Association: 42-49.

Some tools may focus on correctly predicting costs for patients having those conditions where disease management interventions are available and most effective, and as a result could sacrifice predictive power across all members. The output format of a tool is an important quality as well. The model may only produce a binary indicator of whether a patient is a good case management candidate, but underwriters need some sort of relative risk score in order to derive an expected future cost or cost distribution.

Many academic studies indicate that models should be built not to reflect cost differences among different treatment choices, since these could influence physician behavior when their compensation is tied to the results of a risk model. However, these variations in treatment patterns may lead to real cost differences that should be captured by the tools when they are used for underwriting.

Some practical questions are the same whether you are considering a tool for case management or for underwriting. Is the data required available easily and on a timely basis? Does the model's design make sense from your underwriting experience? Can you see "inside" the model or is it a "black box?" Can the model be used across varying lines of business?

But some important questions for the actuary or underwriter are not usually asked for case management uses. How will use of the tool impact the underwriting work flow? Will it reduce variation in performance among underwriters? Can it reduce administrative costs and if so, by how much? Can I explain to brokers and customers how the predictive model impacted their rates? How do I apply regulatory restrictions to the output, and is the prediction still valuable once those are taken into account? And how do I reflect varying benefit designs and provider contracts?

All of these technical and practical issues have to be factored into any decision of whether, and which, predictive model to use for underwriting. As a result, you may need to do additional analysis beyond that available in the literature today. Choosing a predictive model for pricing and underwriting is a very different decision than choosing one for case management, and some of these perspectives may lead you to different conclusions.

### **Check List for Reviewing Predictive Models for Underwriting:**

#### Technical

- Are R-squared results presented at different employer group sizes?
- Are R-squared results compared to other underwriting methods?
- Are large claims truncated in the analysis? If so, at what level and with what impact on the model's predictive power? And how are the excess claims allocated back to the overall pricing?
- Are the groups being presented in the analysis similar to real customers?
- Can you evaluate the impact on business metrics (close ratio, profit margin, etc.)?
- How, if at all, does the model use credibility?

- How, if at all, does the model incorporate traditional rating factors without “double-counting” these?
- How does the model handle new members or members with no prior claim history?
- How does the model account for incomplete incurred claims?
- How does the model take into account the lag between the experience claim period used and the underwriting year whose costs are projected?

#### Practical

- Does the model predict well for low and high cost members and low and high cost/risk groups?
- Does the product generate a credible risk score that can be inexpensively applied?
- Does the model reflect real cost differences from area contracting and treatment patterns?
- Does the model integrate into your underwriting process? For example, is the data required by the model available in time for your renewal processes and is the output in a usable format?
- Will the model reduce administrative costs? By how much?
- Can you explain the model’s design and rate impact, especially rate increases, to others – sales, marketing, employee benefits managers, consultants, brokers?
- How will you adjust results for regulatory limits?
- Can you use the model across products and benefit designs?

#### Software Design/Implementation

- Can you buy the models and link them into your current systems or do you need to purchase an entire reporting system?
- Is the software compatible with your IT environment? For example, if you use Excel, does the software produce results that can be easily exported to Excel?
- What is the model’s implementation time, including needed customization?

Ruth Ann Woodley is a consulting actuary and Vice President of Ruark Consulting, LLC, in Simsbury, CT.

Marilyn Schlein Kramer is President of DxCG, Inc., a predictive modeling company based in Boston, MA. DxCG is a business unit of ISO, Inc.