Executive Summary

Many of the recent proposals for health care reform have included a system of risk adjustment payments among health plans. The goal of these systems is to remove the financial incentives for health plans to selectively enroll only low-risk individuals and to adequately compensate plans for the risks they enroll. They do this by a two-step process:

1. Measuring the expected health care costs of the individuals enrolled by a plan ("risk assessment")
2. Transferring funds from plans that have less than their share of high-risk enrollees to plans that have more than their share of high-risk enrollees ("risk adjustment").

In response to the need for an effective method of risk adjustment, the Society of Actuaries (SOA) has funded this study of the relative performance of different risk assessment methods and risk adjustment systems.

The study had three main objectives:

1. Compare the predictive accuracy of different risk assessment methods
2. Compare the different risk assessment methods based on other criteria, including administrative practicality, ability to resist manipulation and "gaming" by insurers, and incentives for efficiency
3. Explore the potential for risk adjustment using a list of high-cost conditions.

For the purposes of the study, the SOA developed a detailed data set that described the demographic characteristics, diagnoses, medical utilization, and expenditures for more than 4.5 million individuals (excluding elderly people) over a two-year period. The data included indemnity, preferred provider organization (PPO), and health maintenance organization (HMO) plans, segregated into 19 pools.

Using these data, we tested the predictive accuracy of eight different risk assessment models: a simple age-sex model and seven diagnosis-based methods. We tested the predictive accuracy of these models both prospectively and retrospectively at three levels: individuals, large random groups, and nonrandom groups. We also evaluated these models using other criteria including the feasibility of implementation and the incentives provided. In exploring the practical issues, we simulated a risk adjustment transfer process across plans using the different risk assessment methods. Finally, we developed and tested an alternative risk assessment model using a list of high-cost conditions.

The following conclusions were reached.

All models, including age and sex, perform well for large random groups. If enrollees distribute themselves randomly across plans, then the current risk assessment methods are sufficient. Evidence suggests that this is not the case. Therefore, these models also need to predict well for individuals and nonrandom groups.

An adequate risk assessment method does not need to explain all of the variation in expenditures across individuals in order to prevent risk selection. It only needs to do about as well as a plan can reasonably be expected to do. We assumed an individual $R^2$ of about 0.15 to 0.2 as the standard by which to judge our results. The best diagnosis-based prospective model we tested had an individual $R^2$ of 0.112, well below the 0.15–0.2 accuracy standard. See Table 22.

The best retrospective risk assessment model we tested had an individual $R^2$ of 0.428, well above the 0.15–0.2 accuracy standard. However, all models, including retrospective models, systematically overpredict for select nonrandom groups of enrollees and underpredict for others.

In general, the models overpredicted for persons with low expenditures in the previous year and underpredict for those with high expenditures or inpatient admissions for heart disease or cancer in the previous year. These findings were robust with respect to both health care management type and the population of enrollees studied. The relative performance of the models is very consistent across the 19 pools of data we analyzed.

Thus, our findings indicate that opportunities for profitable risk selection and inequities in payments remain even with the best risk assessment models we tested. The general conclusion to which previous research has already pointed thus still holds: no current risk assessment method can completely remove incentives for risk selecting behavior, whether applied prospectively or retrospectively.

Table 32 summarizes our comparison of the models based on general considerations. While the age-sex model had the lowest predictive accuracy of the eight models tested, it was the best based on the other criteria. The age-sex model is easy to administer, resistant
to manipulation, and provides no incentives for unnecessary care. All diagnosis based models provide more reimbursement for more expensive care for some conditions. This is especially true for retrospective models. See Section VI-D for a thorough discussion. Another disadvantage of diagnosis-based models is that a long time would be required to collect and analyze the necessary data before transfer payments resulting from diagnosis-based models could be made. Transfer payments based on models requiring ambulatory diagnoses appear to be very sensitive to the quality of data for ambulatory care, which seemed incomplete or poor in many cases.

Risk assessment and risk adjustment will play important roles in any health care reform strategy. Our results help to illuminate the relative strengths and weaknesses of different diagnosis-based risk assessment methods including lists of high-cost conditions. Relative to no risk adjustment, these models clearly reduce incentives for risk selection and provide more equitable payments to plans for the risks they enroll. Our pessimistic assessment of the potential for risk assessment and risk adjustment, used alone, brings into focus the need for additional measures to prevent risk selection and ensure that health plans compete on a level playing field.
A Comparative Analysis of Methods of Health Risk Assessment

SOA Monograph M-HB96-1

Daniel L. Dunn, Alice Rosenblatt, Deborah A. Taira, Eric Latimer, John Bertko, Thomas Stoiber, Peter Braun, and Susan Busch

October 1996

Society of Actuaries
475 N. Martingale Road Suite 800
Schaumburg, Illinois 60173-2226
Acknowledgments

We are indebted to a large number of individuals who contributed to the project, especially the four members of the Project Advisory Committee, James C. Hickman, William C. Hsiao, Harold S. Luff, and Joseph P. Newhouse. We also thank the Society of Actuaries, in particular, the Committee on Health Benefits Systems Research:

Kenneth S. Avner, Chair       Richard A. Kipp
P. Anthony Hammond            William R. Lane
William C. Hsiao              Nancy F. Nelson
Francis E. Keenan             David W. Wille
Roland E. King

the Risk Adjusters Research Task Force:

William R. Lane, Chair        Joan E. Herman
Kenneth S. Avner              Lucinda M. Lewis
Stephen D. Brink              Harry L. Sutton, Jr.
Norman E. Crocker             David W. Wille
Donald G. Hamm, Jr.           Ronald M. Wolf

and Thomas P. Edwalds, Project Officer, who all offered advice and support throughout the project. The individuals conducting research on ACG and DCGs—Randy Ellis, Jonathan Weiner, Arlene Ash, Norm Smith, and Greg Pope—provided valuable input on their methods and the preliminary study results. Also, we could not have undertaken such a research project without the rich database contributed by the carriers supplying data to the study, including Employer’s Health Insurance, the Prudential, Health Source, Inc., Sanus/NY Life, Blue Cross and Blue Shield of Florida, Wausau, and four carriers also participating in the HIAA risk adjustment study who remained anonymous to this research team. Finally, we wish to thank Deborah Burke and Deborah Coleman for their untiring clerical, technical, and administrative support during this project.

As always, the contributions of all of these individuals and organizations do not necessarily indicate endorsement of our work or findings. We assume full responsibility for the accuracy and completeness of the information contained in this report.


