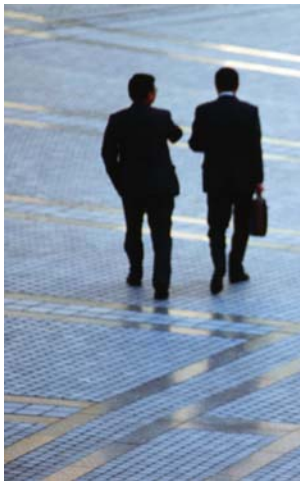
SOCIETY OF ACTUARIES



A Comparative Analysis of Claims-Based Tools for Health Risk Assessment

by: Ross Winkelman, FSA, Principal & Consulting Actuary, Milliman ■ Syed Mehmud, Actuarial Assistant, Milliman
Peer reviewed by: Leigh Wachenheim, FSA, Principal & Consulting Actuary, Milliman

April 20, 2007

Actuaries
Risk is Opportunity.℠

Table of Contents

Acknowledgments

We are grateful to the many people who supported this work with their time and expertise. Special thanks to Steven Siegel, the research actuary for the Society of Actuaries, for his insight and also for coordinating our work with the vendors and the Project Oversight Group.

The Society of Actuaries Project Oversight Group provided technical review, guidance and support: Bill Lane, FSA, Chair; John Bertko, FSA; Tim Connell, FSA; Ian Duncan, FSA; Rafi Herzfeld, FSA; Darrell Knapp, FSA; David Knutson; Ken Lau, FSA; Jeanne Nallon; Bernie Rabinowitz, FSA; Craig Schmid, ASA; Steven Siegel, ASA; and John Stark, FSA.

We would also like to express gratitude to all the software vendors. Without their cooperation and involvement, this important research would not have been possible. In addition, thanks to Leigh Wachenheim for her insight and peer review, and Sherri Norton for her efforts to coordinate and complete the many rounds of edits necessary for the report.

Finally, we would also like to thank the teams that completed the 1996 and 2002 studies for laying such a strong foundation for our work. The researchers from those projects are listed to the right:

1996 Research Study

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

2002 Research Study

Robert B. Cumming, FSA
David Knutson
Brian A. Cameron, FSA
Brian Derrick

"We learned a great deal from everyone involved in this effort and enjoyed the collaboration immensely."

Ross Winkelman & Syed Mehmud
April 9, 2007

This Society of Actuaries research project builds on the work done for the 1996 and 2002 claims-based health risk assessment research projects. The purpose of this study is to evaluate the predictive accuracy of the commercially available claims-based risk assessment tools under different sets of conditions and with different sets of available information. It also provides some information on the tools' ease of use and other qualitative characteristics. Given the number of possible uses of risk adjusters, and the many different measures available to evaluate risk adjusters, this report does not attempt to identify which model is the best. It is intended primarily to provide useful quantitative information to assist individuals in selecting the appropriate risk-adjustment model for their given circumstances.

The substantial increase in the number of models available in the marketplace is primarily due to an increase in the number of models being offered by each vendor, but new vendors are also present in the marketplace. Overall, the models have become more tailored to the situation for which they are being used and more sophisticated in general.

Throughout this report, the risk-adjustment models are grouped together based on the similarities of their input data sources. This categorization allows for appropriate comparisons since the input data that a risk adjuster uses is a defining characteristic and often the first consideration a purchaser makes in narrowing down the choices for a particular risk-adjustment application. The abbreviations shown in the Inputs column in the tables are defined at the beginning of the results section of this report.

Table I.1, repeated in the results section of this report, summarizes the numeric R-squared and MAPE results of the study for the prospective (predicting future 12-month cost), nonlagged (without data or prediction lag) models.

In Table I.1, and throughout the report, "offered" refers to models as they were provided by the software vendors. "Optimized" means that the models were calibrated to the population and data used in the study, and prior costs were added as an independent variable. The term "optimized" is used in the context of the optimization methods that could be reasonably employed by most end users (including the researchers), not the methods that vendors could use to optimize their own models with the addition of a single (or several) prior cost input variable(s). It is also important to note that the results in this report (including results for models where prior costs were added) are based on member-level analysis, not analysis at the employer-group level. The parameters and results of optimal methods will change as the group size, type of population, data, and modeling conditions change.

TABLE I.1 R-Squared and MAPE for Prospective Nonlagged - Offered vs. Optimized (Recalibrated, with Prior Cost, 250k Claim Truncation)

			Offered Models		Optimized Models w/ Prior Costs	
Risk Adjuster Tool	**Developer**	**Inputs**	**R-2**	**MAPE %**	**R-2**	**MAPE %**
ACG	Johns Hopkins	Diag	19.2%	89.9%	23.0%	86.2%
CDPS	Kronick / UCSD	Diag	14.9%	95.3%	24.6%	85.6%
Clinical Risk Groups	3M	Diag	17.5%	90.9%	20.5%	86.6%
DxCG DCG	DxCG	Diag	20.6%	87.5%	26.5%	82.5%
DxCG RxGroups	DxCG	Rx	20.4%	85.3%	27.1%	80.7%
Ingenix PRG	Ingenix	Rx	20.5%	85.8%	27.4%	80.9%
MedicaidRx	Gilmer / UCSD	Rx	15.8%	89.6%	26.3%	81.9%
Impact Pro	Ingenix	Med+Rx+Use	24.4%	81.8%	27.2%	80.6%
Ingenix ERG	Ingenix	Med+Rx	19.7%	86.4%	26.5%	81.2%
ACG w/ Prior Cost	Johns Hopkins	Diag+$Rx	22.4%	85.6%	25.4%	82.1%
DxCG UW Model	DxCG	Diag+$Total	27.4%	80.4%	29.1%	78.3%
Service Vendor		**Inputs**	**R-2**	**MAPE**	**R-2**	**MAPE**
MEDai	MEDai	All	N/A	N/A	32.1%	75.2%

* The offered MEDai model was not tested in the study.

As shown in Table I.1, the optimized models perform very well (in the prior study, the greatest prospective R-squared was 21.8 percent). The MEDai methodology included in the study produces the highest R-squared and lowest MAPE among all models. The DCG model produces the highest R-squared and lowest MAPE of the diagnosis input data models. The RxGroups and PRG pharmacy (Pharmacy NDC-based) models generally had good measures, especially considering that they only use pharmacy data. MedicaidRx performs surprisingly well once it is calibrated for the study's commercial population and a prior cost variable is added, given that it was developed for a Medicaid population. The DxCG Underwriting Model performed well in the underwriting model category (those that include prior costs as inputs in offered model).

Predictive ratios included in the report show the ratio of predictions to actual costs by disease category and cost percentile. Table I.2 shows the predictive ratio results by medical condition:

TABLE I.2 Predictive Ratios by Medical Condition in 2003 (Offered Nonlagged Prospective, 250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	88.4%	100.0%	96.7%	103.1%	99.6%	92.3%
CDPS	Diag	95.0%	73.4%	84.8%	76.4%	67.3%	92.5%
Clinical Risk Groups	Diag	85.1%	94.7%	99.7%	99.5%	91.5%	89.0%
DxCG DCG	Diag	93.3%	98.3%	98.6%	103.2%	86.4%	95.9%
DxCG RxGroups	Rx	95.5%	76.9%	97.9%	89.4%	89.2%	88.6%
Ingenix PRG	Rx	94.9%	93.9%	98.2%	89.7%	79.6%	87.1%
MedicaidRx	Rx	90.1%	94.9%	92.7%	79.1%	90.8%	94.0%
Impact Pro	Med+Rx+Use	97.6%	115.4%	96.4%	99.8%	95.1%	98.0%
Ingenix ERG	Med+Rx	90.0%	99.2%	94.8%	92.9%	80.0%	91.9%
ACG w/ Prior Cost	Diag+$Rx	92.5%	109.0%	95.8%	97.5%	103.6%	91.0%
DxCG UW Model	Diag+$Total	93.2%	84.9%	91.1%	90.7%	103.6%	94.6%
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai*	All	N/A	N/A	N/A	N/A	N/A	N/A

* The offered MEDai model was not tested in the study.

Predictive ratios closer to 100 percent indicate higher accuracy. The results vary considerably by medical condition category. The Impact Pro model has the best predictive ratios for three of the medical condition categories. The ACG system has the best predictive ratio for two of the medical conditions and Clinical Risk Groups has the best ratio for diabetes. The pharmacy input only models have less accurate predictive ratios relative to the other models for heart disease.

The predictive ratio results by disease category highlight the importance of choosing a model that uses grouping criteria consistent with the intended application, especially where disease specific analysis is being employed.

Table I.3, on the opposite page, shows the predictive ratio results by cost percentile.

The predictive ratio results by percentile show the limitations in risk-adjuster predicted costs for the highest- and lowest-cost individuals. In general, results change significantly as cost percentile ranges change, and ranked results are different than in prior tables although MEDai had the best predictive ratios in multiple categories. Of the diagnosis input models, Clinical Risk Groups performed well for all but the middle two cost percentile categories.

The results presented in the Executive Summary represent a small subset of the full study results. Results under a large number of other conditions and scenarios are presented throughout the results section of this report and in Appendix A.

For all but one product, the researchers used the models and created the predictions in their offices. During the period of this study, MEDai did not have a product that could be tested in the researchers' offices. Therefore, MEDai was provided the calibration data and the input information for the testing phase. The other models may (or may not) have performed much better if the representatives from those companies had been given the opportunity to tailor and calibrate their

models to the population and data used in the study. In this report, MEDai is characterized as a service vendor as opposed to a software vendor and is illustrated separately, in fairness to the other vendors. MEDai provides models other than the one included in this study. Additional MEDai models (offered, concurrent, without prior costs, etc.) were not included in the study because of the logistics necessary to ensure a level playing field.

The 2002 SOA risk-adjuster study focused primarily on payment adjustment, although underwriting applications were discussed. This new study addresses the underwriting applications of risk adjusters in more depth. In particular, the effects of adding prior cost as an additional independent variable as well as incorporating data and prediction lag are quantified and discussed. The inclusion of a prior cost independent variable increases the accuracy of the models significantly and dampens differences in predictive accuracy between the models. Modeling data and prediction lag causes predictive measures to worsen overall, although less so for the prescription drug models that rely upon NDCs (national drug codes).

There are many important considerations in using a risk adjuster in a business situation where small differences in the tool and implementation method can have a substantial impact on the stakeholders in the health insurance marketplace. Readers should use the results in the tables in the Executive Summary carefully and are encouraged to review the full report for a complete understanding of how the different models performed under various conditions. Also, while the number of models has increased to address their many uses, it is important to consider what adjustment or customization is worthwhile in a particular situation.

The study was structured so that the playing field would be as level as possible. Vendors were given the opportunity to review and comment on the results of their particular products and to review the report prior to publication. Finally, the participating vendors were also given the opportunity to post their comments about the study methodology and report on the SOA Web site, www.soa.org.

Where appropriate, the study and this report have followed the structure of the 2002 study for consistency. The major differences in the methodology for this study were the addition of the lagged model testing, the addition of aggregate prior costs as an independent variable and different methods for recalibrating the models.

Disclosure Statement

Milliman is a consulting firm, and its technical work sometimes includes the direct use and review of risk adjusters and their application. Milliman has no ownership interest in any of the products tested. Milliman holds an Ingenix ERGs license, and has incorporated Ingenix products within MedInsight (a Milliman product). Milliman also holds a DxCG license for use and as a distributor, and has incorporated DxCG products within MedInsight. Milliman also has used CDPS and MedicaidRx in various offices. MEDai is a client of the Atlanta office of Milliman. Johns Hopkins is also a client of Milliman, but not for consulting services concerning risk-adjustment. The researchers who worked on this study were not involved with any client work for risk-adjuster vendors.

TABLE I.3 Prospective Optimized (Recalibrated, with Prior Costs), Nonlagged Predictive Ratios by Cost Percentile Groupings (Cost Groupings Defined for 2004)

	Percentile Ranges							
Risk Adjuster Tool	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
ACG	27.1%	46.7%	69.6%	99.1%	146.5%	249.9%	544.2%	8433.1%
CDPS	24.2%	43.8%	67.8%	98.6%	150.4%	256.7%	546.1%	8537.4%
Clinical Risk Groups	28.4%	49.2%	73.0%	103.5%	150.4%	238.8%	488.7%	6808.8%
DxCG DCG	25.2%	45.6%	70.4%	101.1%	149.7%	248.5%	528.7%	7780.7%
DxCG RxGroups	24.9%	48.0%	75.0%	105.4%	151.3%	237.3%	482.6%	7177.5%
Ingenix PRG	25.0%	48.0%	74.5%	104.4%	150.6%	238.0%	489.1%	7426.9%
MedicaidRx	24.2%	46.4%	73.4%	106.2%	155.8%	243.8%	478.5%	6773.7%
Impact Pro	29.7%	50.6%	74.9%	103.6%	149.5%	235.0%	470.1%	6587.2%
Ingenix ERG	24.3%	46.1%	73.6%	107.4%	156.4%	245.1%	482.0%	6226.3%
ACG w/ Prior Cost	27.2%	51.7%	76.5%	102.1%	141.7%	230.3%	510.3%	8146.4%
DxCG UW Model	26.8%	50.9%	77.4%	107.6%	150.4%	229.0%	452.4%	6427.8%
	Percentile Ranges							
Service Vendor	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	29.5%	52.5%	78.0%	106.5%	145.4%	216.2%	411.9%	5592.5%

SECTION II. Introduction

Definition of Adjustment

To provide a framework for this study, risk-adjustment is defined as the process of adjusting health plan payments, health care provider payments and individual or group premiums to reflect the health status of plan members. risk-adjustment is commonly described as a two-step process. The first step involves risk assessment, which refers to the method used to assess the relative risk of each person in a group. The relative risk reflects the predicted overall medical claim dollars for each person relative to the claim dollars for an average risk person. The second step in the risk-adjustment process is payment or rate adjustment, which refers to the method used to adjust payments or premium rates in order to reflect differences in risk, as measured by the risk assessment step. It is common to refer to a particular risk assessment method as a risk adjuster.[1]

Background: Why Is Risk-Adjustment Important?

Health claims–based risk assessment and adjustment tools are used in a number of applications, including the following:

- Renewal rating and underwriting of individuals and employer groups
- Provider capitation and risk-based reimbursement
- Health plan payment, especially in government programs such as Medicare and Medicaid
- Care management, for identifying and categorizing high-cost and/or highly impactable patients
- Assisting government agencies and consumers in accurately comparing competing insurance products.

The predictive models included in this report are also used for purposes other than risk-adjustment including trend analysis, rating and medical management.

Risk-adjustment is a powerful and much needed tool in the health insurance marketplace. Risk adjusters allow health insurance programs to measure the morbidity of the members within different groups and pay participating health plans fairly. In turn, health plans can better protect themselves against adverse selection and are arguably more likely to remain in the marketplace. Higher participation increases competition and choice.

Risk adjusters also provide a useful tool for health plan underwriting and rating. They allow health plans to predict more accurately future costs for the members and groups they currently insure.

Finally, risk adjusters provide a ready, uniform tool for grouping people within clinically meaningful categories. This categorization allows for better trend measurement, care management and outcomes measurement. The risk adjuster structure, like benchmarks for service category utilization, allows different departments within an insurance company to communicate with each other. In particular, medical management and actuarial and finance professionals can measure the impacts of their care management programs.

Other Considerations in Selecting a Risk Assessment Model

This study focuses on evaluating the predictive accuracy of health-based risk assessment models. While improved accuracy is the primary reason for implementing any health-based risk-adjustment model, other criteria should be considered when selecting a model. These include the following (in no particular order):

- Ease of use of the software
- Specificity of the model to the population to which it is being applied
- Cost of the software
- Transparency of the mechanics and results of the model
- Access to data of sufficient quality

[1] R. B. Cumming, D. Knutson, B. A. Cameron, and B. Derrick, "A Comparative Analysis of Claims Based Methods of Health Risk Assessment for Commercial Populations." A research study sponsored by the Society of Actuaries. May 24, 2002. This subsection is substantially the same as the referenced report; the current report provides additional detail and updates the definition of risk-adjustment.

- Underlying logic or perspective of a model that makes it best for a specific application
- Whether the model provides both useful clinical as well as financial information
- Whether the model will be used mostly for payment to providers and plans or for underwriting, rating and/or case management
- Reliability of the model across settings, over time or with imperfect data (models that are calibrated and tested on a single data set and population may or may not perform well on different data sets/populations)
- Whether the model is currently in use in the market or organization and
- Susceptibility of the model to gaming or upcoding.[2]

The study included testing of models using lagged data. Other real world conditions faced by health plans or other stakeholders using risk adjusters include rating restrictions from small group regulation and the impact of employee and group turnover. The researchers involved in this study also completed a separate study on the effects of real world conditions on predictive performance, entitled the "Optimal Renewal Guidelines" study.[3] This study was focused on small group renewal rating, but the results are helpful in considering real world conditions encountered in other situations. Some results from this study are included and discussed in Section VII of this report, "Limitations and Factors Impacting Risk Adjuster Performance."

Important Notes

A number of competing methods are used to perform health risk assessment using diagnosis, procedure and/or pharmacy data. The number of methods that could be included in this study was restricted because of the availability of resources and time. In addition to the vendors and products included in this study, other vendors and products are currently available in the marketplace. The performance of these other products has not been evaluated, and the exclusion of a particular product from this study does not indicate any judgment about that product's performance or characteristics.

Research Team

The research team was comprised of consultants from Milliman. Ross Winkelman, FSA, and Syed Mehmud were the primary investigators for this study. Leigh Wachenheim, FSA, peer reviewed the analysis and report. Significant contributions were also made by other Milliman consultants, including Jonathan Shreve, FSA, Craig Johns, PhD, Paul Sahkrani and Karan Rustagi.

Contact information for the lead researchers is provided below:

Ross A. Winkelman, FSA
Consulting Actuary
Milliman, Inc.
1099 18th Street, Suite 3100
Denver, CO 80202-1931
e-mail: Ross.Winkelman@Milliman.com

Syed Mehmud
Actuarial Assistant
Milliman, Inc.
1099 18th Street, Suite 3100
Denver, CO 80202-1931
e-mail: Syed.Mehmud@Milliman.com

Additional contact:

Aree Bly, FSA
Actuary
Milliman, Inc.
1099 18th Street, Suite 3100
Denver, CO 80202-1931
e-mail: Aree.Bly@Milliman.com

[2] Cumming et al., "A Comparative Analysis of Claims Based Methods." This subsection is substantially the same as the referenced report; the current report updates the criteria for model selection.

[3] Conclusions and excerpts from this study have been published. Please feel free to contact the researchers of this study for copies of the excerpts or for more information.

SECTION III. Study Design

The number of approaches that can be used for risk-adjustment has been increasing over the last decade. This study focuses on models that use medical diagnosis codes and/or pharmacy codes in administrative claim data in the assessment of risk. For this study, 12 health risk assessment models were evaluated, including four diagnosis-based models, three pharmacy-based models, two models based on diagnosis and pharmacy data and three models that use prior cost data.

The risk-adjustment models have changed in the following primary ways from those available in the marketplace during the 2002 study:

- Some companies are offering a greater number of model variations than previously offered to address the variety of applications for which the models are being used. For instance, several companies now offer models based on claims data with and without data and prediction lag, at different claims truncation levels (i.e., pooling), and for specific purposes (provider payment versus underwriting). The model variations evaluated in the study do not include all of those available from the vendors represented.
- The modeling techniques have become more sophisticated; some vendors are using techniques to capture nonlinear relationships including neural networks and clustering methods.
- Some models now incorporate prior costs directly in their predictions. Use of prior costs is not appropriate for all circumstances (provider payment and premium risk-adjustment are two obvious examples), but including them is not only potentially appropriate, but also greatly enhances a model's predictive capability for a number of actuarial and underwriting purposes.

The following models were evaluated:

- Adjusted Clinical Groups (ACGs) Version 7.1 (with prior year's pharmacy cost as input)
- Adjusted Clinical Groups (ACGs) Version 7.1 (without prior year's pharmacy cost as input)
- Chronic Illness and Disability Payment System (CDPS) Version 2.5
- Clinical Risk Grouping (CRG) Version 1.4
- Diagnostic Cost Groups (DCGs), RiskSmart Version 2.1.1
- Episode Risk Groups (ERGs) Version 5.3
- Impact Pro
- MEDai
- MedicaidRx
- Pharmacy Risk Groups (PRGs) Version 5.3
- RxGroups, RiskSmart Version 2.1.1
- Underwriting Model, RiskSmart Version 2.1.1.

Inclusion of Medicare's Hierarchical Condition Categories (HCC) model was considered but not included because of concerns with the project scope and technical support during the Medicare bid season.

The ACGs, CDPS, DCGs and CRG use diagnosis data available from administrative claim records. MedicaidRx, RxGroups and PRGs use pharmacy data. The ERGs, Impact Pro, MEDai and DxCG underwriting model use diagnosis and pharmacy data. The model versions referenced above were the most recently available when the study began in May 2006.

The following briefly describes each of the risk adjusters. These descriptions are summarized from documentation provided by the software vendors. Where appropriate, the descriptions are substantially similar to those included in the 2002 report.

Adjusted Clinical Groups (Vendor: Johns Hopkins University, School of Public Health)

Adjusted Clinical Groups (ACGs) is a diagnosis-based risk assessment model developed by Jonathan Weiner and other researchers at the Johns Hopkins University. The ACG System includes a suite of predictive models developed to identify high cost cases. ACG Case-Mix System 7.1 was used for this study. The model incorporates the morbidity-based ACG categories; selected, high-impact, disease-specific Expanded Diagnosis Clusters (EDCs); and diagnostic indicators of the likelihood of future hospitalizations and of being medically frail.

The concurrent model used in this study is based on an actuarial cell approach (ACG actuarial cells are clinically defined, mutually exclusive groupings of patients that have a similar level of risk) as opposed to being regression based. All else being equal, this approach usually lowers predictive accuracy. However, actuarial cells are recommended by the ACG Team for payment applications based on their characteristics with respect to implementation, understanding and stability.

Chronic Illness and Disability Payment System

The Chronic Illness and Disability Payment System (CDPS) is a diagnosis-based risk assessment model developed by Richard Kronick and other researchers at the University of California, San Diego. CDPS Version 2.5 was used for this study. This model was originally developed for use with Medicaid populations, including disabled and Temporary Aid for Needy Families (TANF) populations. The CDPS model is an update and expansion of a prior model developed by Kronick and published in 1996 called the Disability Payment System (DPS). The DPS model was developed for the Medicaid disabled population.

The CDPS model assigns each member to one or more of 67 possible medical condition categories based on diagnosis codes. Each member is also assigned to one of 16 age/gender categories. For each member, the model predicts total medical costs based on the medical condition categories and age/gender category assigned. The model provides two sets of risk weights: one set calibrated for a TANF population and another set calibrated for a disabled population. In this analysis the weights for the TANF population were used, since a TANF population is more similar to the commercial population used for this analysis. The model also provides different sets of risk weights for adults and children, both of which were used for this analysis.

Clinical Risk Groups (Vendor: 3M)

CRG Version 1.4 was used for this study, which was released by 3M in 2006. CRG is a diagnosis-based risk assessment model. CRGs can be used for risk-adjustment in capitated payment systems and as a management tool for managed care organizations (MCOs). The design and development was influenced by the Medicare Inpatient Prospective Payment System (PPS). Every enrollee is assigned to a single risk group based on clinical criteria. CRGs offer the user the choice of three models for both prospective and retrospective applications. All have about 1,100 unique groups. Since CRGs are clinically based, they are designed to serve as the foundation of management systems that support care pathways, product line management and case management.

Diagnosis Cost Groups (Vendor: DxCG)

Diagnosis Cost Groups (DCG) is a component of the RiskSmart Models, which is a product of DxCG. DCG research began in 1984 at Boston University, with numerous refinements and extensions implemented under the leadership of Arlene Ash and Randall Ellis of Boston University in the subsequent 20 years. DCG is a diagnosis-based risk assessment model with many variations depending on the type of population being analyzed (commercial, Medicaid, Medicare), source of the data (inpatient only versus all encounters) and purpose of the model (payment versus explanation).

For the purpose of this analysis, RiskSmart Version 2.1.1 was used. The DCG model is a commercial all-encounter model used to identify the total payment (medical cost and pharmacy cost) both prospectively and concurrently. In the prior study, there was no model to predict the total payment concurrently.

DxGroups are fundamental building blocks of DCG models. All diagnosis codes are grouped into 781 clinically homogeneous groups (DxGroups). These groups are further mapped into 184 hierarchical condition categories. Each patient is also assigned to one of 32 age/gender categories. The model predicts the total medical cost for each patient based upon the HCC and the age/gender category.

Episode Risk Groups (Vendor: Ingenix)

The Episode Risk Groups (ERGs) is a risk assessment model developed by Symmetry Health Data Systems, a subsidiary of Ingenix, Inc. ERGs are based on the Episode Treatment Groups (ETGs) models, also developed by Symmetry, which group medical services into episodes of care. The ERGs were developed and released in 2001. Those used in this analysis are based on Version 5.3 of the ETGs.

The ERG model assigns each member to one or more of the 120 possible medical condition categories (called episode risk groups) based on diagnostic and procedural information available on medical and pharmacy claims. An ERG profile for each member is created by considering age, gender and the ERGs to which they have been assigned. Prospective and retrospective risk scores are assigned using that profile.

Impact Pro (Vendor: Ingenix)

Impact Pro was developed by IHCIS, which is a subsidiary of Ingenix, Inc. This is a combination reporting system and risk-adjustment algorithm, incorporating enrollment information, medical claims and pharmacy claims. The system groups claims into unique episodes of care and other diagnosis-based Impact Clinical Categories (ICCs). These categories describe a member's observed mix of diseases and conditions and underlying co-morbidities and complications. The ICCs are further grouped into homogenous risk categories ("base-markers"). Each member may be grouped into one or more base-markers and one demographic marker. The risk weights are then output, specific to several different possible applications and settings (i.e., truncation levels).

MEDai (Vendor: MEDai, Inc.)

Risk Navigator Clinical™ is a predictive modeling solution and reporting tool developed by MEDai, Inc. Risk Navigator Clinical™ forecasts cost, inpatient stays, emergency room visits, Rx cost and savings utilizing medical and pharmacy claims, demographics, lab results and health risk assessments (HRAs). Individual predictions per member are made using a combination of clinical factors including disease episodes (Symmetry ETGs), drug categories, age, sex, insurance type and other risk markers such as timing and frequency of treatment or diagnosis.

Risk Navigator Clinical™ utilizes two years of data to construct, refine and test models. Gathered and validated data are run through MEDai's prediction engine, Multiple Intelligent Tasking Computer Heuristics (MITCH), which incorporates linear and nonlinear technologies.

MedicaidRx

MedicaidRx is a pharmacy-based risk assessment model developed by Todd Gilmer and other researchers at the University of California at San Diego. The model was originally designed and intended for a Medicaid population and is an update and expansion of the Chronic Disease Score model developed by researchers at Group Health Cooperative of Puget Sound.

The MedicaidRx model assigns each member to one or more of 45 medical condition categories based on the prescription drugs used by each member and to one of 11 age/gender categories. Based on the medical conditions and age/gender categories, the model predicts the overall medical costs for each member. The model includes separate sets of risk weights for adults and children.

Pharmacy Risk Groups (Vendor: Ingenix, Inc.)

Pharmacy Risk Groups (PRGs) is a pharmacy risk assessment model developed by Symmetry Health Data Systems, a subsidiary of Ingenix, Inc. Version 5.3 of PRGs was used for this study. The building blocks of PRGs are a patient's mix of pharmacy prescriptions and how a drug relates to other drugs prescribed for the patient. Each NDC is mapped to one of 107 PRGs. A PRG profile for each member is created using the age, gender and PRGs to which they are assigned. Using the PRG profile, a member's prospective or retrospective risk score is computed.

RxGroups (Vendor: DxCG, Inc.)

RxGroups is a component of the RiskSmart Models (a product of DxCG). For the purpose of this analysis, RiskSmart Version 2.1.1 was used. RxGroups is a pharmacy-based risk assessment model released in 2001 that was developed by researchers and clinicians from Kaiser Permanente, CareGroup of Boston and Harvard Medical School. This model classifies NDCs into 164 mutually exclusive categories (called RxGroups) based on each drug's therapeutic indication. Each patient is also assigned to one of 32 age/gender categories. The model predicts the

total medical cost for each patient based upon the RxGroups and the age/gender category.

Underwriting Model: RiskSmart (Vendor: DxCG, Inc.)

The RiskSmart underwriting model is a new addition to the RiskSmart Models, a product of DxCG, and was released in 2006. For the purpose of this analysis, RiskSmart Version 2.1.1 was used. The underwriting model is used to help underwriters assess employer groups with health care coverage for renewal and price-setting purposes before claims have fully matured. The model incorporates claim lag into its predictions by providing a six-month lag between the end of the baseline period and the prediction period. The underwriting model uses HCCs, disease interactions, age/gender categories and a prior cost variable to predict future medical costs. The underwriting model is different from most models in that it includes a prior cost variable to help with its predictions. It also has a variety of truncation options ($25,000, $100,000 or $250,000).

Study Methodology: 50/50 Split Design with Offered and Calibrated Weights

Each risk adjuster was analyzed using up to 10 scenarios (some scenarios were not practical, possible or appropriate for some models). Each scenario was run using no claim truncation and claim truncation at $100,000 and $250,000. Calibration refers to adjusting the model coefficients to the data and population used in the study. Adding prior costs as an independent prediction variable to the prospective models was a separate step. The following scenarios were analyzed:

- Prospective Model with Offered Risk Weights (without data and prediction lag)
- Prospective Model with Offered Risk Weights (with data and prediction lag)
- Prospective Model with Calibrated Risk Weights (without data and prediction lag)
- Prospective Model with Calibrated Risk Weights (with data and prediction lag)
- Prospective Model with Calibrated Risk Weights (without data and prediction lag)—including prior costs
- Prospective Model with Calibrated Risk Weights (with data and prediction lag)—including prior costs
- Concurrent Model with Offered Risk Weights (without data and prediction lag)
- Concurrent Model with Offered Risk Weights (with data and prediction lag)
- Concurrent Model with Calibrated Risk Weights (without data and prediction lag)
- Concurrent Model with Calibrated Risk Weights (with data and prediction lag)

These scenarios represent different approaches to implementing the risk adjuster model. The following section describes the major differences between the scenarios.

Claim Truncation

For each application the results were analyzed using three scenarios for truncating large claims: truncate large claims at $100,000, at $250,000 and no truncation. The truncation applies to total claim dollars for a given member for 2004 (or 2003 for concurrent predictions). Also, in cases where a model took prior cost information as input, the cost was appropriately truncated, and the model was rerun for the corresponding analysis.

Truncation of large claims is common when analyzing the predictive accuracy of risk adjusters for a variety of reasons, including the following:

- Truncation limits the impact of outliers. This should provide more stability in the results when calibrating the models and when analyzing predictive accuracy.
- Large claims for a given person are generally not predictable. Accordingly, some researchers argue that they should be removed or limited when doing the analysis.
- Truncation simulates the impact of reinsurance or stop loss at those levels.
- Some measures of predictive accuracy are overly sensitive to large claims.[4]

[4] Cumming et al., "A Comparative Analysis of Claims Based Methods." This subsection is substantially the same as the referenced report; the current report updates truncation levels and adds prior cost explanation.

Prospective vs. Concurrent

A prospective application of a risk adjuster involves using historical claims data to predict medical claim costs for a future period. A concurrent (*or retrospective*) application involves using claims data from a period of time to predict medical claim costs for that same period. Concurrent applications involve estimating the health status of individuals regardless of the underlying cost structure, since actual costs are available for concurrent time periods. In this study the prospective models use diagnosis and pharmacy data from 2003 to predict total medical claim costs for each member for 2004. The concurrent model uses diagnosis and pharmacy data from 2003 to predict total medical claim costs for each member for 2003 (the first year in the study data period). The concurrent application is slightly different from the prior SOA study. In that study, data for 1998 and 1999 were available, and the concurrent models were evaluated on 1999 data (the second year in the study data period).

Offered vs. Calibrated Risk Weights

For each risk adjuster there is a risk weight for a given medical condition category. The risk weight reflects an estimate of the marginal cost for a given medical condition relative to the base cost for individuals with no medical conditions. The offered risk weights are the standard risk weights that are provided with the risk adjuster software.[5] Adjustments to the offered risk weights were developed for the calibrated analysis.

With and Without Data and Prediction Lag ("Lagged" and "Nonlagged")

In this study *lagged* scenarios refer to scenarios where the combination of data lag and prediction lag are present. Claims take several months on average to be paid and, in some instances, can take much longer (up to several years). Data lag refers to the situation where a health plan is missing paid claims data, because it is not available when the risk-adjustment analysis is being performed. Additionally, in many applications there is a delay between the data paid-through date and the beginning of the prediction period (this is referred to as prediction lag). For the nonlagged scenarios, data incurred in 2003, paid through August 2005 was used to run the models. For the lagged scenarios, data incurred in January through August 2003, paid through August 2003 was used. Incomplete data cause predictions to be less accurate in general, but accurately reflect the environment in which the actuary and underwriter must work for many situations. Pharmacy data–based models are less adversely affected by data lag than medical (and medical plus pharmacy) models because pharmacy data are paid more quickly (this helps mitigate data lag, but not prediction lag).

Including Prior Costs as a Predictor

Using prior aggregate costs as an explicit, contributing predictor in models is not appropriate for provider or health plan payment purposes. However, for actuarial and underwriting purposes, including prior costs significantly improves the models' performance. Some models include prior costs in their products—namely, the DxCG underwriting model, the MEDai model used in the study and the ACG prior cost model. For other models it was added as an independent variable. Out of necessity, including prior costs was done as part of the calibration step under the "With Prior Costs" scenarios.

Steps in Study Methodology

The analysis can be described briefly by the following steps:

- Step 1 - Separating the data set into two equal-sized subsets: (1) a calibration subset and (2) a validation subset
- Step 2 - Assigning individual-level risk scores using each risk adjuster (the score for a particular member reflects an estimate of the relative cost for that member)
- Step 3 - Regression analysis: performing a linear regression using the calibration data subset to determine adjustments to the offered risk weights (for the recalibrated analyses only)
- Step 4 - Applying calibrated risk scores: applying the adjustments calculated during Step 3 to the validation data set in order to compute a calibrated score (for the recalibrated analyses only)

[5] Cumming et al., "A Comparative Analysis of Claims Based Methods."

- Step 5 - Analyzing results: analyzing the predictive accuracy using the validation data set to compare the score (i.e., predicted claims) of each member or group of members to actual claim dollars.

Each of these steps is described below.

Data Description

The study used data from MedStat Marketscan. The data set consisted of ICD9, CPT4 and NDC codes and associated amounts for a two year continuously enrolled, comprehensive major medical population, with approximately 620,000 members and about three billion dollars in annual claims.

For the concurrent nonlagged analyses, the classification period (which is the same as the prediction period) spanned claims incurred from Jan. 1, 2003 through Dec. 31, 2003, but paid through Aug. 2005.

For the concurrent with-data-lag analyses, the classification and prediction period spanned incurred claims from Jan. 1, 2003 through Aug. 31, 2003 but paid through Aug. 31, 2003.

For the prospective nonlagged analyses, the classification period spanned incurred and paid claims from Jan. 1, 2003 through Dec. 31, 2003 and the prediction period spanned incurred claims from Jan. 1, 2004 through Dec. 31, 2004, but paid though Aug. 31, 2005.

For the prospective with-data-lag analyses, the classification period spanned incurred claims from Jan. 1, 2003 through Aug. 31, 2003 but paid through Aug. 31, 2003 and the prediction period spanned incurred claims from Jan. 1, 2004 through Dec. 31, 2004, but paid though Aug. 31, 2005.

Table III.1 presents a comparison of the demographic distribution of the study population against that of a distribution typical insured population (referred to as the "Reference" population in the table). The Reference population was derived from the *Milliman Health Cost Guidelines*, 2006 edition. As illustrated in the table, the demographic distribution of the study population exhibits a greater proportion

The population underlying the study had the following characteristics:

TABLE III.1 Demographic Characteristics of Study Population Compared to Reference Population

Demographic Category	% of Total		% of Category	
	Study	Reference	Study	Reference
Male, To 25	0%	2%	1%	7%
Male, 25-29	1%	3%	4%	11%
Male, 30-34	2%	4%	5%	13%
Male, 35-39	2%	5%	6%	15%
Male, 40-44	3%	5%	10%	16%
Male, 45-49	5%	5%	15%	15%
Male, 50-54	7%	4%	20%	13%
Male, 55-59	8%	2%	24%	7%
Male, 60-64	5%	1%	15%	4%
Demographic Category	**% of Total**		**% of Category**	
	Study	**Reference**	**Study**	**Reference**
Female, To 25	0%	2%	1%	6%
Female, 25-29	1%	3%	4%	10%
Female, 30-34	2%	4%	5%	13%
Female, 35-39	2%	5%	6%	14%
Female, 40-44	4%	5%	10%	16%
Female, 45-49	6%	5%	15%	15%
Female, 50-54	8%	4%	20%	13%
Female, 55-59	10%	3%	24%	8%
Female, 60-64	7%	2%	16%	5%
Demographic Category	**% of Total**		**% of Category**	
	Study	**Reference**	**Study**	**Reference**
Child, 00-01	1%	3%	5%	7%
Child, 02-06	4%	7%	14%	20%
Child, 07-18	16%	21%	61%	59%
Child, 19-22	5%	5%	21%	13%

of individuals at older ages (50 years plus) than the Reference population. In addition, the demographic distribution of the study population exhibits relatively fewer children. The implication of the demographic differences is that the study

likely has placed more emphasis on the predictability of chronic illnesses than might be expected with other population distributions. This can also be seen in the error calculations presented later in this report. For the purposes of this study, this likely emphasizes differences in the predictive power of the various software packages.

For the cost groupings, the population size is readily apparent since individuals are placed in percentiles. For the disease groupings, the number of people in each cohort varies depending on when the individuals were identified with the condition. However, for the nonlagged, prospective analysis, Table III.3 shows the number of individuals by disease cohort during 2003.

TABLE III.2 Geographical Characteristics

Region	Members
Northeast	43,330
North Central	392,743
South	128,436
West	52,301
Unknown	873
Total	617,683

TABLE III.3 Members by Disease Category

Condition Category	Unique Members
Asthma	6,806
Breast Cancer	2,299
Diabetes	19,690
Heart Disease	19,270
HIV	170
Mental Illness	22,421
Total	70,656

Step 1. Separating the Data Set into Two Equal-Sized Subsets

A 50/50 split design was used for the study to allow for the development and testing of calibrated risk weights. Specifically, each member was randomly assigned to one of two subsets: (1) the calibration data subset and (2) the validation data subset, placing half of the population in each subset. This design was used to avoid over-fitting the data, which could exaggerate the goodness of the fit and various other measures of predictive accuracy (Cumming et al. 2002).

Step 2. Assigning Individual-Level Risk Scores Using Each Risk Adjuster

Each member is assigned a risk score (based on certain medical condition categories, including drug therapy categories and age/gender categories) by each risk adjuster model. Each risk adjuster model (except for CRGs and MEDai) produces a set of indicator variables (0 or 1) representing the condition and age/gender categories assigned. 3M's Clinical Risk Groups puts each member into one (or more) of about a thousand risk categories. MEDai produces a set of 1,000+ indicator variables, including medical condition, drugs, age/gender and prior cost categories. Some of these indicators are 0/1 and other are continuous variables (such as prior cost). For the prospective analysis, the indicator variables are based on either 2003 or 2004 diagnosis and pharmacy data as indicated. For the concurrent analysis, the indicator variables are based on 2003 diagnosis and pharmacy data.

Step 3. Regression Analysis (Recalibrated Scenarios)

For recalibrated scenarios, the prior study calculated new risk weights by regressing demographic and condition indicators on total actual claims for the calibration segment of the data. This study proceeded in a slightly different manner. Adjustments to the offered risk weights were calculated by regressing demographic and condition indicators on the difference between actual total claims and the offered risk-adjustment predictions. In general, to calculate the adjustments to risk weights for a particular risk adjuster, the following multivariate linear regression model was used ("Bin" indicates the age/gender or condition category(s) assigned to a particular individual):

$$Y_{\text{Actual}} - Y_{\text{Prediction}} = \sum_{i=1}^{A} \alpha_i \times \text{Age Bin}_i + \sum_{i=1}^{B} \beta_i \times \text{Condition Bin}_i$$

where

Y_{Actual} = Total actual allowed claims (including medical and pharmacy)
$Y_{\text{Prediction}}$ = Total predicted allowed claims (including medical and pharmacy)
α_i = The regression coefficient that specifies adjustments to the demographic-based risk prediction
β_i = The regression coefficient that specifies adjustments to the condition-based risk prediction.
y = The regression coefficient for prior cost (if the scenario includes prior cost)

For the "With Prior Cost" scenarios, prior costs were added at the same time the models were recalibrated (since most of the offered models did not use prior costs, it was not appropriate to add prior costs without recalibrating). Therefore, for the scenarios where prior cost was included as a predictive variable, the calibration equation included a prior cost term as shown in the equation below:

$$Y_{\text{Actual}} - Y_{\text{Prediction}} = \sum_{i=1}^{A} \alpha_i \times \text{Age Bin}_i + \sum_{i=1}^{B} \beta_i \times \text{Condition Bin}_i + \gamma \times \text{Prior Cost}$$

where (in addition to the variable definitions from prior equation)

γ = The regression coefficient for prior cost (if the scenario includes prior cost)

A linear regression is performed to determine a set of adjustments that best fits the calibration data set. These adjustments are specific to the condition and demographic variables, and are therefore applied to the individual-level risk score output by the software. Both the software output score and this adjusted or calibrated score are then multiplied by the average per member per year (PMPY) cost (from the calibration set) to obtain an offered and calibrated prediction, respectively.

A separate calibration analysis was performed for each level of claim truncation (none, $250,000 and $100,000) and for lagged versus nonlagged scenarios. Also, separate calibrations were performed for the prospective and concurrent scenarios. Yet another set of calibrations was performed by including prior cost as a prediction variable. Accordingly, there are up to 24 sets of calibrated predictions for each risk adjuster.

Calibrations for concurrent scenarios differed slightly in that they did not include demographic variables as predictors. It is undesirable to assign risk to a member who did not incur claims, and including demographic indicators in the recalibration method used in the study would result in a nonzero score being assigned to members without claims.

The adjustments recognize the credibility of the observations by dampening the adjustments according to the *p*-value. Lower *p*-values indicate that the statistical credibility of the result is higher. The study used a credibility factor equal to $(1.0–p\text{-value})^{5.95}$ for adjustments to the offered predictions. Therefore, a *p*-value of 0.01 would result in a credibility weight of 94.2 percent. Alternatively, a *p*-value of 0.50 would result in a credibility weight of 1.6 percent. The adjustments calculated from the regression were multiplied by the credibility weights to calculate the ultimate adjustments to the offered prediction (this convention assigns the complement of the credibility to no adjustment from the offered risk weight/score).

A number of other adjustments are commonly employed in developing a final set of risk weights for actual implementation. These other adjustments can include removing variables that are not statistically significant, smoothing the age/gender risk weights, blending developed risk weights with the "offered" risk weights, combining variables in the payment model, calibrating the risk weights after removing any variables, clinical review of the relationships, testing the stability of the risk weights with different claim truncation levels and testing the stability of the risk weights using subsets of the data (Cumming et al. 2002). This study does not include any of these further adjustments. It was concluded that further risk weight (without prior costs) recalibration would likely only provide marginal improvement because most of the vendors have already spent considerable time calibrating their models to a commercial population.

The methods used in the study to add prior costs as an independent variable were fairly straightforward and are consistent with the approach generally taken by health plans (although prior costs are usually added at the employer group level, and employer group level analysis was not a component of this study). More sophisticated approaches would likely result in improved accuracy, but were not practical for this study (or for most end users). Those approaches might include varying the weight of prior costs depending on the specific condition(s) present (chronic versus acute) and/or the age and gender of the individual, among others.

Calibrations were not carried out on the CRG adjuster because this software puts each individual into one risk category, rather than an array of condition and age/gender variables. Adding a prior cost variable was still possible and was carried out.

As stated previously, MEDai provided Milliman researchers with their set of calibrated predictions. MEDai also presented a version of the predictions that were not calibrated to the data set provided to them. However, those offered predictions are not presented in the study because of the special accommodations made to include MEDai.

DxCG uses the MedStat Marketscan data for all plan types, including all enrolled members to develop and calibrate their models (this study used continuously enrolled members from the Comprehensive Major Medical plan design subset of the same MedStat data).

Step 4. Applying Calibrated Risk Scores

Each member in the validation data subset is scored using the indicator variables described in Step 2 and the corresponding offered risk weights. These weights are then adjusted using the process described in Step 3. The adjusted risk weights and indicator variables are then used to create adjusted (calibrated) risk scores.

Step 5. Analyzing Results

As a final step, the predictive performance of the models is analyzed by comparing predicted cost (risk score multiplied by average allowed cost in calibration data set) to actual experience (as measured by the allowed cost). This comparison is done for both individuals and groups of individuals as described later.

Measures Used to Analyze Predictive Accuracy: Individual and Nonrandom Groups

Three measures were used to compare the predictive accuracy of the risk adjusters examined in this study. In general, these measures compare actual claim dollars with predictions from the risk adjuster models. This comparison is performed on two levels: (1) by individual and (2) by group.

Measures of Predictive Accuracy: Individual Level

The individual-level measures of predictive accuracy include individual R-squares and mean absolute prediction error (MAPE).

Individual R-squared in this context is described as the percentage of the variation in medical claim costs explained by a risk adjuster model. Variation refers to the difference in medical costs for a given individual compared to the average medical cost for all individuals (Cumming et al. 2002). The formula for R-squared is

$$\text{R-Square} = 1 - \frac{\sum (\text{Actual} - \text{Predicted})^2}{\sum (\text{Actual} - \text{Average of Actual})^2}$$

where the summation is over the entire sample.

It is important to note that this formula is a derived form of the basic R-square formula, and that the derivation holds if the prediction is based on the least-squares algorithm. In the case of this study, the derivation does not hold as predictions are based on grouping algorithms, clinical meaningfulness, etc. Therefore, what is presented here carries the statistical essence of the R-squares, but is not strictly an R-square calculation.

Mean absolute prediction error is calculated in a similar fashion. It is defined as the ratio of the absolute value of the prediction error to the sample size. Prediction error is defined as the difference between actual medical costs and predicted costs. The formula for MAPE is

$$\text{MAPE} = \frac{\sum |\text{Actual} - \text{Predicted}|}{\text{Sample Size}}.$$

Different arguments are made regarding the merits of alternative methods for measuring goodness of fit. Individual R-squared is a standard statistical measure for assessing model results and is commonly used for measuring predictive accuracy of risk adjusters. It is a single summary measure on a standardized scale of 0 to 1, where 0 indicates that the model explains 0 percent of the variation in cost for each individual and 1 indicates that the model explains 100 percent of the variation (i.e., 100 percent accuracy in the predictions). The standardized scale helps with comparability between studies. However, there still are many potential issues associated with comparing individual R-squared from one study to another that may make the comparisons inappropriate or invalid. These issues include differences in the data sets, study design and data quality.[6]

Individual R-squared has certain drawbacks. Because it squares each prediction error, it tends to be overly sensitive to the prediction error for individuals with large claims. According to the 1996 study, "because R^2 squares the errors of prediction, it can be greatly affected by a relatively small number of cases with very large prediction errors. Given the typical distribution of health expenditures across individuals, where a small number of individuals have relatively large expenditures, this is a concern for their analysis" (Dunn et al. 1996). This is one of the reasons for truncating large claims when individual R-squared is used as a measure of predictive accuracy.[7]

The mean absolute prediction error is also a single summary measure of predictive accuracy. On the positive side, it does not square the prediction errors and, so, is not overly sensitive to large claims. However, it is not expressed on a standardized scale, so comparisons across studies are difficult to make. Therefore, for purposes of this study, we have expressed MAPE as a percentage of the average PMPY cost.[8]

Measures of Predictive Accuracy: Group Level

A group-level measure of predictive accuracy involves adding up the total predicted claims for a group of individuals and comparing that value to the actual claims for the same group. This comparison provides a measure called the predictive ratio. A predictive ratio that is closer to 1.0 indicates a better fit. The predictive ratio is the reciprocal of the common actual-to-expected (A to E) actuarial ratio.

The methods for calculating a predictive ratio can differ primarily in how the groups are defined. There are two general approaches: (1) nonrandom groups and (2) random groups. Nonrandom refers to grouping individuals based on selected criteria. The common criteria used for analyzing risk adjusters include groups based on medical condition or amount of claim dollars. Nonrandom groups can also be defined based on other criteria, such as members of a particular employer group. This is sometimes referred to as using real groups. Random groups refer to groups created by selecting individuals at random from the study data set.[9] We used nonrandom groupings in this study as explained in the next section.

[6] Cumming et al., "A Comparative Analysis of Claims Based Methods." This subsection through the footnote reference is substantially the same as the referenced report; the current report includes minor wording changes.

[7] Cumming et al., "A Comparative Analysis of Claims Based Methods." This paragraph is substantially the same as the referenced report; the current report removes a reference to a previous study.

[8] Cumming et al., "A Comparative Analysis of Claims Based Methods." This paragraph is substantially the same as the referenced report; the current report adds a note about MAPE.

[9] Cumming et al., "A Comparative Analysis of Claims Based Methods." This subsection is substantially the same as the referenced report; the current report indicates grouping used for current study.

Nonrandom Groups Used for This Study

This study uses nonrandom groups based on three criteria: (1) medical condition in 2003, (2) medical condition in 2004 and (3) ranges of medical claim dollars for 2004.

The medical conditions used for this study include breast cancer, heart disease, asthma, depression, diabetes and HIV. The medical conditions were determined using medical diagnosis codes and an adjustment for false positives (a single instance of a relevant code was sufficient for inpatient claims, whereas two or more instances were required on outpatient claims). It should be noted that this approach might create a bias in favor of risk adjusters that are based on diagnosis data. A risk adjuster that distinguishes among individuals based on particular criteria (e.g., diagnosis codes) may tend to perform better when predicting expenditures for groups of individuals determined using the same type of criteria (Cumming et al. 2002).

For different medical conditions, the performance of the risk adjuster models may change significantly. For a given medical condition, a risk adjuster will naturally tend to perform better on this test if it has a medical condition category that matches more closely with the definition of the medical condition used in this study. The diagnosis definitions used in this study appear in Table III.4.

TABLE III.4 ICD-9 Definitions for Condition Category Cohorts

Condition	ICD-9
Breast Cancer	174-174.9
Heart Disease	390-398, 402, 404-429
Asthma	493-493.9
Mental Illness	290-298.9, 300-312.9
HIV	042
Diabetes	250.1, 250.10, 250.11, 250.12, 250.13, 648.0, 648.00, 648.01, 648.02, 648.03, 648.04, 648.8, 648.80, 648.81, 648.82, 648.83, 648.84, 250.0, 250.00, 250.01, 250.02, 250.03, 250.2, 250.20, 250.21, 250.22, 250.23, 250.3, 250.30, 250.31, 250.32, 250.33, 250.40, 250.41, 250.42, 250.43, 250.5, 250.50, 250.51, 250.52, 250.53, 250.60, 250.61, 250.62, 250.63, 250.70, 250.71, 250.72, 250.73, 250.8, 250.80, 250.81, 250.82, 250.83, 250.9, 250.90, 250.91, 250.92, 250.93, 362.0, 362.0, 362.01, 362.02, 362.1, 775.1, 790.2, 790.21, 790.22, 790.29, 253.5

Grouping Individuals Using Base Year vs. Prediction Year Information

There are two alternate approaches in determining the nonrandom groups. One approach uses claim information from the base year (i.e., 2003) to define the group. The other approach uses claim information from the prediction year (i.e., 2004) to define the group. Different years were used to define the groups based on the scenario.

Measures that use groups based on claim information from the prediction year may be more useful when analyzing risk adjusters for applications such as underwriting or rating, identification of patients for case or disease management, provider profiling and provider payment. These types of measures help answer questions such as: How well can the risk adjuster predict claims for the next year? How well can the models predict who will have a large claim next year?

Measures that use groups based on claim information from the base year may be more useful when analyzing risk adjusters for applications such as health plan payment. These types of measures help us answer questions such as the following: If a health plan (directly or indirectly) selected members based on their claim history (i.e., past medical conditions or expenditures), then would the health plan receive a fair payment for the upcoming year?

Throughout this report, the risk-adjustment models are grouped together based on the similarities of their input data sources. This categorization allows for appropriate comparisons since the input data a risk adjuster uses is a defining characteristic and often the first consideration a purchaser makes in narrowing down the choices for a particular risk-adjustment application. The abbreviations shown in the Inputs column in the tables are defined as follows:

Code	Description
Diag	ICD-9 Diagnosis Codes
Med	ICD-9 Diagnosis Codes and Procedure Information
Rx	Pharmacy NDC Codes
$Rx	Prior Pharmacy Cost
$Total	Prior Total Cost
Use	Measure of Prior Utilization, but not Prior Cost
All	All of the above

The MEDai process produced the best R-squared (and MAPE) fit. During the period of this study, MEDai did not have a product that could be tested in the researchers' offices. Therefore, MEDai was provided the calibration data and the input information for the testing phase. The other models may (or may not) have performed much better if the representatives from those companies had been given the opportunity to tailor and calibrate their models to the population and data used in the study. In this report MEDai is characterized as a service vendor as opposed to a software vendor and is illustrated separately, in fairness to the other vendors. MEDai provides models other than the one included in this study. Additional MEDai models (offered, concurrent, without prior costs, etc.) were not included in the study because of the logistics necessary to ensure a level playing field.

Table IV.1 shows R-squared results for the offered models (not customized for the population and data used in the study) and optimized models (optimized indicates that the predictions were calibrated for the population and data, and prior costs were included as a prediction variable). Higher R-squared values indicate a model with a better fit. The tables that follow this one help to further explain the results of the study in more depth. A primary objective of Table IV.1 is to present a high level overview of the results for the benefit of the reader. Some of the offered models include prior costs (denoted by "$" in the Inputs column). A prior cost independent variable was added to all of the optimized models.

TABLE IV.1 R-Squared for Prospective Nonlagged (Offered vs. Optimized) by Truncation Level (Offered Compared to Recalibrated, with Prior Costs)

		Offered Models			Optimized Models (Include Prior Costs)		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	20.8%	19.2%	16.2%	24.2%	23.0%	20.2%
CDPS	Diag	17.6%	14.9%	12.4%	27.4%	24.6%	21.2%
Clinical Risk Groups	Diag	19.3%	17.5%	14.9%	21.5%	20.5%	18.4%
DxCG DCG	Diag	22.3%	20.6%	17.4%	29.7%	26.5%	22.9%
DxCG RxGroups	Rx	23.8%	20.4%	16.8%	30.6%	27.1%	23.4%
Ingenix PRG	Rx	25.0%	20.5%	17.2%	30.9%	27.4%	23.7%
MedicaidRx	Rx	19.3%	15.8%	12.9%	29.7%	26.3%	22.7%
Impact Pro	Med+Rx+Use	26.3%	24.4%	21.3%	29.3%	27.2%	24.0%
Ingenix ERG	Med+Rx	23.7%	19.7%	16.2%	30.0%	26.5%	22.8%
ACG w/ Prior Cost	Diag+$Rx	25.6%	22.4%	18.7%	27.7%	25.4%	22.1%
DxCG UW Model	Diag+$Total	31.3%	27.4%	23.6%	33.1%	29.1%	25.2%
Service Vendor	**Inputs**	**100K**	**100K**	**250K**	**None**	**100K**	**250K**
MEDai*	All	N/A	N/A	N/A	35.7%	32.1%	27.6%

* The offered MEDai model was not tested in the study.

Including prior costs in the prediction is appropriate only in some circumstances such as renewal underwriting. Prior costs are obviously not appropriate for recognizing risk differences in capitation payment.

The pharmacy-only models generally performed well in both the offered and optimized models. The MedicaidRx model has a relatively low R-squared for the offered model, which would be expected given that it is intended for a Medicaid population, and the study used a commercial population. The optimized models show significant improvement over the offered models, which is primarily due to the addition of prior costs as an independent variable. (The optimized Impact Pro error measures improved less than other models that do not include prior costs.) This cause of improvement is evidenced by the smaller improvement from offered to optimized predictions for models that include prior costs in the offered model.

TABLE IV.2 MAPE for Prospective Nonlagged (Offered vs. Optimized) by Truncation Level (Offered Compared to Recalibrated, with Prior Costs)

		Offered Models			Optimized Models (Include Prior Costs)		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	87.7%	89.9%	90.4%	84.6%	86.2%	86.6%
CDPS	Diag	93.4%	95.3%	95.8%	83.7%	85.6%	86.3%
Clinical Risk Groups	Diag	88.7%	90.9%	91.4%	85.2%	86.6%	87.0%
DxCG DCG	Diag	85.3%	87.5%	88.0%	80.5%	82.5%	83.2%
DxCG RxGroups	Rx	82.9%	85.3%	85.9%	78.7%	80.7%	81.4%
Ingenix PRG	Rx	83.4%	85.8%	86.4%	78.9%	80.9%	81.5%
MedicaidRx	Rx	87.3%	89.6%	90.2%	79.9%	81.9%	82.6%
Impact Pro	Med+Rx+Use	79.3%	81.8%	82.4%	78.7%	80.6%	81.2%
Ingenix ERG	Med+Rx	84.1%	86.4%	87.0%	79.1%	81.2%	81.9%
ACG w/ Prior Cost	Diag+$Rx	85.1%	85.6%	85.6%	80.3%	82.1%	82.6%
DxCG UW Model	Diag+$Total	80.1%	80.4%	80.4%	76.1%	78.3%	78.9%
Service Vendor	**Inputs**	**100K**	**100K**	**250K**	**None**	**100K**	**250K**
MEDai*	All	N/A	N/A	N/A	73.0%	75.2%	75.6%

* The offered MEDai model was not tested in the study.

R-squared improves substantially when actual costs are truncated (as expected), although some models show more improvement than others.

Table IV.2 is similar to Table IV.1, except that MAPE results (as a percentage of total average actual costs) are shown instead of R-squared results. Unlike R-squared, a lower MAPE is more desirable.

MAPE calculations reduce the impact of misestimates on outliers as compared to R-squared calculations. MAPE results may be more appropriate to review for purposes such as small group renewal underwriting; where state regulations limit allowable rating action, outliers are less important. Predicting outliers within small groups with more precision may not be helpful depending on state regulations because some states substantially limit how much a company can vary rates from the average due to health status.

The results for MAPE are relatively similar in terms of the order of performance of the different models. For the optimized models, the MEDai and DxCG underwriting models had the lowest MAPE (indicating better performance), while the offered CRGs and CDPS models had the highest MAPE.

Comparison of Results to Prior (2002) SOA Study

Table IV.3 shows a comparison of the R-squared results of this study to the R-squared results of the 2002 study.

The truncation levels, while different between the two studies, are relatively comparable because differences in cost levels between the two studies can be explained in terms of overall medical care cost trend (i.e., $50,000 is comparable to $100,000) and data sampling. The sample was restricted to individuals having comprehensive benefit–type coverage, to allow for the homogeneity of the sample and ease of comparability. While a $200,000 truncation level would have been more comparable to the $100,000 level used in the prior study, $250,000 is used because several of the models included that truncation level in their offered models, and not a $200,000 level.

Two of the notable differences highlighted in Table IV.3 are as follows:

- The models are generally performing better than they did in the prior study. This is likely due to improvements in the models themselves and improvements in data coding.
- RxRisk is not included in this study. Limited resources dictated focusing on the more recently updated and widely used adjusters. The copy of RxRisk that was received indicated that it had not been updated since March 2002.

Prospective, Offered, Without Prior Cost

Table IV.4 shows the R-squared and MAPE results of the models that do not use prior costs.

As shown in Table IV.4, the Impact Pro model performed the best under both MAPE and R-squared. Ingenix PRG also performed well, especially for R-squared at 100k truncation. From Table IV.2, it can be seen that the Impact Pro model results under MAPE did not change much from the offered model to the optimized

TABLE IV.3 Comparison to 2002 Study of Offered Weight R-Squared Prospective Nonlagged by Claims Truncation Level

	2002 Study			Current Study		
Risk Adjuster Tool	**50K**	**100K**	**None**	**100K**	**250K**	**None**
ACG	N/A	N/A	N/A	20.8%	19.2%	16.2%
CDPS	13.4%	12.5%	10.3%	17.6%	14.9%	12.4%
DCG	19.5%	18.0%	14.3%	22.3%	20.6%	17.4%
MedicaidRx	11.6%	9.8%	7.1%	19.3%	15.8%	12.9%
RxGroups	20.6%	18.1%	13.4%	23.8%	20.4%	16.8%
RxRisk	17.5%	14.8%	11.1%	N/A	N/A	N/A
ERG	21.8%	19.3%	14.6%	23.7%	19.7%	16.2%

TABLE IV.4 R-Squared and MAPE Prospective Nonlagged Offered (Without Prior Cost) by Claims Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	20.8%	19.2%	16.2%	87.7%	89.9%	90.4%
CDPS	Diag	17.6%	14.9%	12.4%	93.4%	95.3%	95.8%
Clinical Risk Groups	Diag	19.3%	17.5%	14.9%	88.7%	90.9%	91.4%
DxCG DCG	Diag	22.3%	20.6%	17.4%	85.3%	87.5%	88.0%
DxCG RxGroups	Rx	23.8%	20.4%	16.8%	82.9%	85.3%	85.9%
Ingenix PRG	Rx	25.0%	20.5%	17.2%	83.4%	85.8%	86.4%
MedicaidRx	Rx	19.3%	15.8%	12.9%	87.3%	89.6%	90.2%
Impact Pro	Med+Rx+Use	26.3%	24.4%	21.3%	79.3%	81.8%	82.4%
Ingenix ERG	Med+Rx	23.7%	19.7%	16.2%	84.1%	86.4%	87.0%
ACG w/ Prior Cost*	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model*	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai*	All	N/A	N/A	N/A	N/A	N/A	N/A

* These models include prior cost as input.

TABLE IV.5 R-Squared and MAPE Prospective Nonlagged Offered vs. Recalibrated (Without Prior Cost, 250K Truncation)

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**Offered**	**Recalibrated**	**Change**	**Offered**	**Recalibrated**	**Change**
ACG	Diag	19.2%	19.6%	0.4%	89.9%	88.8%	-1.1%
CDPS	Diag	14.9%	17.7%	2.8%	95.3%	91.9%	-3.4%
Clinical Risk Groups*	Diag	17.5%	N/A	N/A	90.9%	N/A	N/A
DxCG DCG	Diag	20.6%	21.3%	0.7%	87.5%	87.0%	-0.5%
DxCG RxGroups	Rx	20.4%	20.5%	0.1%	85.3%	85.3%	0.0%
Ingenix PRG	Rx	20.5%	21.2%	0.7%	85.8%	85.6%	-0.2%
MedicaidRx	Rx	15.8%	17.7%	1.9%	89.6%	88.4%	-1.2%
Impact Pro	Med+Rx+Use	24.4%	25.6%	1.2%	81.8%	81.6%	-0.2%
Ingenix ERG	Med+Rx	19.7%	20.0%	0.3%	86.4%	86.1%	-0.3%
ACG w/ Prior Cost**	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Offered**	**Recalibrated**	**Change**	**Offered**	**Recalibrated**	**Change**
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input.

model, which was recalibrated and prior costs added. This is somewhat surprising, although the Impact Pro model is intended for an underwriting system. Therefore, the Impact Pro model appears to capture measures of prior use, even if not directly. MedicaidRx and CDPS were not intended for a commercial population, and the offered predictive measures reflect this.

Comparison of Offered and Recalibrated Models

Table IV.5 shows how the predictive measures changed with recalibration for the prospective, nonlagged models that do not use prior costs.

The greatest improvements after recalibration are for CDPS and MedicaidRx. In addition, the improvement in several models is relatively small. The models with modest changes either have been designed to be very robust or were calibrated on a data set similar to the one used in the study (and vice versa for the others).

The recalibration is fairly straightforward. The approach differed slightly from the approach used in the prior study. Adjustments to the originally offered demographic and condition weights were calculated rather than completely new replacements for the offered weights. This approach was more straightforward mechanically than the prior study's approach since some tools do not provide offered weights easily (the calculated adjustment was credibility adjusted using the *p*-value of the statistical tests). The "Study Design" section includes a more detailed description of the recalibration process.

Comparison of Results Using Lagged and Nonlagged Data

Table IV.6 shows how results changed with lagged models.

As shown in Table IV.6, the increase in performance with complete nonlagged data is significant. A few of the vendors offer models within their product suite that include consideration of lag—for example, DxCG underwriting models and Impact Pro.

The commercial pharmacy risk adjusters perform better than the diagnosis only models with lagged data. The DxCG DCG and ACG models are most affected by lag and complete data.

Appendix A includes values for the optimized models.

Concurrent and Comparison to Prospective

Table IV.7 shows the results for the offered concurrent models. It would not be appropriate for concurrent models to consider costs for the period (that would be a fairly easy model to build!).

The DCG model performs best under both R-squared and MAPE. Models that use prior cost as an input variable have "N/As" in the table as well as other models that do not output a concurrent risk score by design.

Table IV.6
* Model includes prior cost as input.

Table IV.7
* These models do not include a concurrent option.
** These models include prior cost as input.

TABLE IV.6 R-Squared and MAPE Offered Prospective Lagged vs. Nonlagged (Without Prior Cost) 250K Truncation

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**Lagged**	**Nonlagged**	**Change**	**Lagged**	**Nonlagged**	**Change**
ACG	Diag	14.5%	19.2%	4.7%	93.7%	89.9%	-3.8%
CDPS	Diag	11.9%	14.9%	3.0%	98.8%	95.3%	-3.5%
Clinical Risk Groups	Diag	14.1%	17.5%	3.4%	93.9%	90.9%	-3.0%
DxCG DCG	Diag	15.1%	20.6%	5.5%	91.6%	87.5%	-4.1%
DxCG RxGroups	Rx	18.0%	20.4%	2.4%	87.4%	85.3%	-2.1%
Ingenix PRG	Rx	18.0%	20.5%	2.5%	87.8%	85.8%	-2.0%
MedicaidRx	Rx	13.6%	15.8%	2.2%	91.7%	89.6%	-2.1%
Impact Pro	Med+Rx+Use	21.4%	24.4%	3.0%	85.5%	81.8%	-3.7%
Ingenix ERG	Med+Rx	16.9%	19.7%	2.8%	88.7%	86.4%	-2.3%
ACG w/ Prior Cost*	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model*	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Lagged**	**Nonlagged**	**Change**	**Lagged**	**Nonlagged**	**Change**
MEDai*	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE IV.7 R-Squared and MAPE Offered Concurrent Nonlagged by Claims Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	29.4%	29.7%	27.4%	73.0%	75.0%	75.4%
CDPS	Diag	35.5%	32.9%	31.0%	79.0%	80.6%	81.0%
Clinical Risk Groups	Diag	47.1%	43.3%	39.9%	68.6%	70.5%	70.9%
DxCG DCG	Diag	57.2%	51.8%	49.8%	61.6%	65.0%	65.4%
DxCG RxGroups*	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG*	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	32.1%	28.1%	24.6%	77.2%	79.1%	79.6%
Impact Pro*	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	46.5%	42.4%	38.6%	65.8%	67.7%	68.2%
ACG w/ Prior Cost**	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+$Total	N/A	N/A	N/A	N/A	N/A	**N/A**
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE IV.8 R-Squared Offered Nonlagged (Without Prior Cost & 250K truncation) – Prospective vs. Concurrent

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**Prospective**	**Concurrent**	**Change**	**Prospective**	**Concurrent**	**Change**
ACG	Diag	19.2%	29.7%	10.5%	89.9%	75.0%	-14.9%
CDPS	Diag	14.9%	32.9%	18.0%	95.3%	80.6%	-14.7%
Clinical Risk Groups	Diag	17.5%	43.3%	25.8%	90.9%	70.5%	-20.4%
DxCG DCG	Diag	20.6%	51.8%	31.2%	87.5%	65.0%	-22.5%
DxCG RxGroups*	Rx	20.4%	N/A	N/A	85.3%	N/A	N/A
Ingenix PRG*	Rx	20.5%	N/A	N/A	85.8%	N/A	N/A
MedicaidRx	Rx	15.8%	28.1%	12.3%	89.6%	79.1%	-10.5%
Impact Pro*	Med+Rx+Use	24.4%	N/A	N/A	81.8%	N/A	N/A
Ingenix ERG	Med+Rx	19.7%	42.4%	22.7%	86.4%	67.7%	-18.7%
ACG w/ Prior Cost**	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Prospective**	**Concurrent**	**Change**	**Prospective**	**Concurrent**	**Change**
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

Table IV.8
* These models do not include a concurrent option.
** These models include prior cost as input.

Table IV.9
* Model could not be recalibrated consistently with other models.
** These models include prior cost as input.

TABLE IV.9 R-Squared and MAPE Prospective Recalibrated Nonlagged (Without Prior Cost vs. With Prior Cost) 250K trucation

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**w/out Prior**	**with Prior**	**Change**	**w/out Prior**	**with Prior**	**Change**
ACG	Diag	19.6%	23.0%	3.4%	88.8%	86.2%	-2.6%
CDPS	Diag	17.7%	24.6%	6.9%	91.9%	85.6%	-6.3%
Clinical Risk Groups*	Diag	N/A	20.5%	N/A	N/A	86.6%	N/A
DxCG DCG	Diag	21.3%	26.5%	5.2%	87.0%	82.5%	-4.5%
DxCG RxGroups	Rx	20.5%	27.1%	6.6%	85.3%	80.7%	-4.6%
Ingenix PRG	Rx	21.2%	27.4%	6.2%	85.6%	80.9%	-4.7%
MedicaidRx	Rx	17.7%	26.3%	8.6%	88.4%	81.9%	-6.5%
Impact Pro	Med+Rx+Use	25.6%	27.2%	1.6%	81.6%	80.6%	-1.0%
Ingenix ERG	Med+Rx	20.0%	26.5%	6.5%	86.1%	81.2%	-4.9%
ACG w/ Prior Cost**	Diag+$Rx	N/A	25.4%	N/A	N/A	82.1%	N/A
DxCG UW Model**	Diag+$Total	N/A	29.1%	N/A	N/A	78.3%	N/A
Service Vendor	**Inputs**	**w/out Prior**	**with Prior**	**Change**	**w/out Prior**	**with Prior**	**Change**
MEDai**	All	N/A	32.1%	N/A	N/A	75.2%	N/A

Table IV.8, on the opposite page, compares the R-squared and MAPE values for the prospective and concurrent models.

The concurrent model performance appears to be correlated with the level of data included in the models. The prospective models are also obviously affected, but the impact is greater for the concurrent models. This outcome is intuitive because it is easier to predict total current expenditures (medical plus drug) with information on both the medical diagnoses a person has and the drugs they are taking than to try to predict both aspects of costs with only one of the types of data. Prospective predictions are less precise and, therefore, having all of the data is less helpful.

Impact of Adding Prior Cost to Recalibrated

Adding prior costs as an independent prediction variable increases accuracy for most models significantly (especially those that do not already reflect prior costs). Where health plans use risk adjusters in renewal underwriting, they generally use prior costs at the employer group level in combination with the aggregated individual risk-adjustment predictions to develop the renewal rate for the group. Evidence suggests that the credibility or weight assigned to prior costs should increase as group size increases. Therefore, if the risk-adjustment software includes a measure of prior cost in the individual predictions, it is important to consider how this affects the weight that should be applied to aggregate prior costs. Modeling the accuracy of the different models on employer groups was outside the scope of this study (but is listed as an area of recommended future study). In general, we would expect the relative differences in accuracy between the models to decrease as group size increases.

Table IV.9, on the opposite page, shows the impact of adding prior costs to the recalibrated models that do not include prior costs.

As shown above, the MEDai process outperforms the other models by a significant margin. In addition, the pharmacy models benefit a great deal by the addition of prior costs. In fact, the MedicaidRx model outperforms three of the commercial models on R-squared, and four of the commercial models (commercial meaning only available with licensing fee, meaning that CDPS is not a commercial model) on MAPE once prior cost is added.

SECTION V. Grouped Results by Medical Condition

Grouped results are presented using predictive ratios, which are simply the ratio of the average predicted cost to the average actual cost for a particular group of individuals. Predictive ratios closer to 100 percent are desirable.[10] As shown in the table below, predictive ratios are generally less than 100 percent, which is somewhat expected since risk adjusters generally underpredict costs for higher cost individuals. This is an important tendency since it affects applications like Special Needs Plans for chronically ill individuals in Medicare Advantage.

Individuals are assigned to the condition categories based on the presence of those conditions in either 2003 or 2004, depending on the scenario. For example, Table V.1 below groups individuals according to the presence of the respective medical condition in 2003 (and is labeled as such: "by Medical Condition in 2003"). For all of the prospective models, the predictive ratios are for 2004 predictions and 2004 actual costs (however, they vary in what year the condition categories are defined). For all of the concurrent models, the predictive ratios are for 2003 values (not technically predictions since they are concurrent) and 2003 actual costs.

Prospective—2003 Medical Condition

The first section of the grouped results shows predictive ratios for six selected medical conditions in 2003 (see Table V.1).

TABLE V.1 Predictive Ratios by Medical Condition in 2003 (Recalibrated Nonlagged Prospective without Prior Costs, 250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	98.3%	90.9%	96.2%	100.8%	99.1%	98.0%
CDPS	Diag	97.1%	81.3%	97.7%	93.5%	94.9%	91.1%
Clinical Risk Groups*	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	93.5%	91.1%	97.5%	96.0%	92.5%	98.5%
DxCG RxGroups	Rx	95.2%	72.4%	95.7%	86.7%	84.0%	89.2%
Ingenix PRG	Rx	93.0%	73.2%	96.0%	86.3%	85.6%	87.4%
MedicaidRx	Rx	91.9%	74.0%	95.2%	78.8%	84.7%	88.1%
Impact Pro	Med+Rx+Use	99.3%	97.5%	98.3%	97.0%	101.6%	97.8%
Ingenix ERG	Med+Rx	97.3%	92.6%	99.4%	94.5%	81.5%	92.3%
ACG w/ Prior Cost**	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input.

[10] An interesting question was posed by William Gilmore, ASA, MAAA, of Blue Cross Blue Shield of Mississippi. Mr. Gilmore noted that the average member prediction was very close to the average member cost, based on his use of a risk adjuster in practice. However, the average male and female predictions were not equal to the average male and female member cost (respectively). The differences were relatively small, but still material. This issue was investigated and its findings confirmed. The result is logical because condition category weights are usually not specific to a demographic category (gender or age), but are instead optimized across the entire population. This is done for reasons of credibility and parsimony. A chance to test the change in predictive measures resulting from overall demographic adjustments was not available. A very small improvement in predictive measures with this change would be expected. Maybe more importantly, the results would be sound across age/gender categories, which would help when explaining them to others within an organization.

TABLE V.2 Predictive Ratios by Medical Condition in 2004 (Recalibrated Nonlagged Prospective without Prior Costs, 250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	71.6%	63.8%	83.7%	60.1%	71.9%	70.8%
CDPS	Diag	69.2%	57.5%	84.1%	55.1%	63.3%	65.7%
Clinical Risk Groups*	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	68.2%	64.6%	84.4%	57.7%	66.0%	70.5%
DxCG RxGroups	Rx	68.2%	64.6%	84.4%	57.7%	66.0%	70.5%
Ingenix PRG	Rx	74.1%	52.9%	86.8%	58.3%	60.8%	69.5%
MedicaidRx	Rx	72.6%	53.6%	87.1%	57.9%	63.0%	68.2%
Impact Pro	Med+Rx+Use	73.9%	65.2%	88.6%	58.8%	57.7%	69.2%
Ingenix ERG	Med+Rx	73.9%	65.2%	88.6%	58.8%	57.7%	69.2%
ACG w/ Prior Cost**	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models included prior cost as input.

Prospective—2004 Medical Condition

Table V.2 shows predictive ratios for the same medical conditions based on the presence of that condition in 2004. As shown in this table, the predictive ratios worsen when 2004 costs are used to group individuals. This is due to individuals with these medical conditions in 2004 having higher average costs and a larger variance in costs than those with these medical conditions in 2003. Higher average costs and a larger variance in costs cause the predictive ratios to worsen.

Impact Pro, Ingenix ERG and ACG performed well relative to the other models under the predictive ratio measure. An interesting observation is that predictive ratios for pharmacy-only adjusters vary noticeably with diseases and are generally not as close to 100 percent as the diagnosis models (this is more prominent in the analysis using 2003 claims to define condition groupings). This outcome is not surprising since a diagnosis-based criterion was employed for creating the disease groups rather than one based on NDC codes. This example further highlights the importance of appropriate tool usage. When considering the choice of adjuster for purposes of stratifying the population into cohorts, that choice should be based on whether the desired definitions of the cohorts are reflected in the adjuster grouping mechanism.

The performance generally improves considerably for the concurrent models compared to prospective results with medical conditions in 2004.

Concurrent—2003 Medical Condition

TABLE V.3 Predictive Ratios by Medical Condition in 2003 (Recalibrated Nonlagged Concurrent without Prior Costs, 250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	103.2%	102.5%	88.8%	91.3%	41.0%	100.6%
CDPS	Diag	104.7%	76.5%	87.1%	83.8%	80.1%	80.2%
Clinical Risk Groups*	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	92.9%	98.4%	93.0%	95.8%	83.3%	94.7%
DxCG RxGroups	Rx	85.8%	79.7%	89.4%	75.2%	67.6%	79.6%
Ingenix PRG**	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	85.8%	75.9%	90.1%	65.0%	73.2%	79.9%
Impact Pro**	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	92.5%	96.6%	93.7%	89.8%	74.8%	85.2%
ACG w/ Prior Cost***	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model***	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai***	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models do not include a concurrent option.

*** These models include prior cost as input.

Prospective with Prior Costs—2003 & 2004 Medical Condition

Tables V.4 and V.5 show predictive ratios for disease-based groups in 2003 and 2004, respectively, using a prospective application of the risk adjuster models (optimized by recalibrating and including prior cost).

As expected, the predictive ratios for the concurrent models generally improved compared to the prospective models without prior costs. In addition, the predictive ratios exceed 100 percent more often. This is expected given the variation in actual costs for these conditions.

TABLE V.4 Predictive Ratios by Medical Condition in 2003 (Recalibrated Nonlagged Prospective with Prior Costs, 250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	99.0%	91.0%	100.1%	105.6%	115.5%	99.2%
CDPS	Diag	93.2%	86.6%	99.7%	96.8%	96.4%	94.6%
Clinical Risk Groups	Diag	96.5%	110.2%	110.0%	115.8%	109.3%	101.8%
DxCG DCG	Diag	95.8%	90.2%	99.2%	96.2%	99.3%	100.3%
DxCG RxGroups	Rx	101.2%	79.6%	99.0%	97.0%	94.6%	96.8%
Ingenix PRG	Rx	97.9%	80.0%	98.4%	96.4%	93.5%	94.9%
MedicaidRx	Rx	97.9%	84.2%	98.7%	95.2%	96.3%	96.8%
Impact Pro	Med+Rx+Use	100.8%	99.9%	99.5%	98.6%	106.5%	100.0%
Ingenix ERG	Med+Rx	99.8%	92.6%	101.0%	97.8%	92.7%	97.2%
ACG w/ Prior Cost	Diag+$Rx	100.7%	101.0%	100.5%	102.5%	119.1%	100.1%
DxCG UW Model	Diag+$Total	99.1%	93.1%	100.7%	97.6%	107.3%	101.0%
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	104.4%	93.3%	102.6%	97.9%	96.1%	99.7%

TABLE V.5 Predictive Ratios by Medical Condition in 2004 (Recalibrated Nonlagged Prospective with Prior Costs, 250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	72.1%	64.2%	86.9%	62.6%	83.4%	71.7%
CDPS	Diag	68.7%	61.4%	85.7%	57.7%	66.5%	68.5%
Clinical Risk Groups	Diag	75.2%	65.4%	89.6%	59.9%	66.0%	73.1%
DxCG DCG	Diag	76.7%	57.3%	88.4%	60.9%	68.1%	72.5%
DxCG RxGroups	Rx	76.7%	57.3%	88.4%	60.9%	68.1%	72.5%
Ingenix PRG	Rx	74.6%	57.6%	88.0%	60.5%	67.8%	72.6%
MedicaidRx	Rx	74.4%	60.7%	88.1%	59.4%	69.1%	71.3%
Impact Pro	Med+Rx+Use	76.7%	71.9%	89.0%	62.6%	77.6%	71.9%
Ingenix ERG	Med+Rx	76.7%	71.9%	89.0%	62.6%	77.6%	71.9%
ACG w/ Prior Cost	Diag+$Rx	75.3%	70.3%	88.2%	62.3%	85.6%	73.9%
DxCG UW Model	Diag+$Total	75.3%	70.3%	88.2%	62.3%	85.6%	73.9%
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	79.8%	66.8%	91.1%	62.0%	70.7%	75.4%

SECTION VI. Predictive Ratios by Cost Groupings

Individuals are assigned to the cost categories based on their actual costs in either 2003 or 2004, depending on the scenario. For all of the prospective models, the predictive ratios are for 2004 predictions and 2004 actual costs. For all of the concurrent models, the predictive ratios are for 2003 outputs and 2003 actual costs.

Cost Groupings—Prospective & Concurrent

The following analysis shows how well the models predict average 2004 costs for members who had high, medium and low costs in 2004. For example, the 99–100 grouping represents the top 1 percent of the population in terms of future year PMPYs, while the 0–20 grouping contains the least expensive 20 percent of the population.

Table VI.1 highlights the fact that all risk-adjustment models underpredict high-cost individuals and overpredict low-cost individuals. Table VI.1 also shows that the predictive ratios increase as the cost percentiles decrease. The different models perform remarkably similarly, Clinical Risk Groups and Impact Pro performed relatively well at the 96th percentile and above (Ingenix PRG performed relatively well in the 96th–99th percentiles, but not as well at the 99th–100th percentiles). Impact Pro performed relatively well in all of the percentile ranges.

TABLE VI.1 Prospective without Prior Cost (Recalibrated, Nonlagged)
Predictive Ratios by Cost Percentile Groupings (Cost Groupings Defined for 2004)

	Percentile Ranges							
Risk Adjuster Tool	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
ACG	21.8%	42.5%	67.5%	100.0%	152.2%	265.0%	570.7%	8308.1%
CDPS	18.2%	38.4%	63.6%	96.8%	154.5%	275.1%	595.3%	9335.9%
Clinical Risk Groups*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	20.5%	41.7%	67.3%	100.1%	153.3%	263.6%	558.3%	7869.0%
DxCG RxGroups	18.2%	43.8%	72.8%	105.8%	155.0%	248.8%	516.9%	7914.0%
Ingenix PRG	19.2%	44.3%	72.6%	104.2%	152.9%	247.4%	523.9%	8301.4%
MedicaidRx	15.9%	40.1%	69.9%	107.0%	163.4%	261.9%	516.9%	7374.3%
Impact Pro	26.9%	48.3%	73.3%	103.9%	152.1%	241.4%	480.9%	6605.6%
Ingenix ERG	18.0%	41.5%	71.1%	108.7%	163.6%	261.4%	509.2%	6171.7%
ACG w/ Prior Cost**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Percentile Ranges							
Service Vendor	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.
** These models include prior cost as input.

Table VI.2, on the opposite page, shows predictive ratios for the concurrent models. When compared against Table VI.1, it is clear how much better the concurrent models stratify members by cost level, although the models still underpredict high-cost individuals and overpredict low-cost individuals.

Individuals with Low Costs in 2003 and High Costs in 2004

The following analysis measures how well the models predicted 2004 costs for "movers" (defined as individuals with low costs in 2003 and high costs in 2004). This is an important cohort to follow since part of the value of a risk adjuster, when compared against prior cost, is in its ability to predict changes in cost (i.e., low to high cost and high to low cost). The data used for the table is individuals with less than the median cost in 2003, and then with the percentile ranges in 2004 as indicated in the table.

As shown in Table VI.3, all of the models generally overpredict costs on average in 2004 for those with low costs in 2003 (see 0–100th percentile column). This is consistent with the prior tables, as risk adjusters generally overpredict costs for healthy people (and those who are relatively healthy in 2003 are more likely to be healthy in 2004). It is important not to interpret this finding as a deficiency in the models or methods. These results are due to the nature and variability of health care costs and the difficulty estimating costs for people who, by definition, have significant changes in their cost levels.

In addition, Table VI.3 shows how the different risk adjusters stratify their predictions for the highest-cost individuals who were low cost in the prior year. ERG has the best predictive ratios in each of the categories (excluding 0–100th percentile category, where Impact Pro had the best predictive ratio).

TABLE VI.2 Concurrent without Prior Cost (Offered, Nonlagged) Predictive Ratios by Cost Percentile Groupings (Cost Groupings Defined for 2003)

	Percentile Ranges							
Risk Adjuster Tool	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
ACG	57.0%	82.8%	94.8%	100.2%	107.6%	124.3%	137.9%	133.4%
CDPS	44.9%	60.9%	73.3%	86.4%	106.0%	142.9%	195.1%	283.1%
Clinical Risk Groups	62.8%	76.7%	83.8%	92.6%	105.8%	129.0%	158.9%	208.4%
DxCG DCG	75.2%	84.6%	89.0%	94.3%	102.9%	120.3%	133.4%	151.2%
DxCG RxGroups*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	43.2%	70.9%	88.1%	102.3%	116.6%	129.8%	136.3%	154.6%
Impact Pro*	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	54.4%	75.2%	88.4%	101.2%	114.0%	127.6%	134.9%	131.5%
ACG w/ Prior Cost**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	Percentile Ranges							
Service Vendor	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

TABLE VI.3 Predictive Ratios by 2004 Cost Percentile where <50th Percentile in 2003 (Prospective, Recalibrated, Nonlagged, without Prior Cost)

	2004 Cost Percentile Range						
Risk Adjuster Tool	**0-100th**	**70th-100th**	**75th-100th**	**80th-100th**	**85th-100th**	**90th-100th**	**95th-100th**
ACG	132.0%	16.9%	14.5%	12.3%	10.4%	8.3%	6.2%
CDPS	144.8%	14.8%	12.6%	10.7%	9.0%	7.3%	5.5%
Clinical Risk Groups*	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	126.8%	17.8%	15.2%	12.9%	10.8%	8.7%	6.6%
DxCG RxGroups	130.1%	16.2%	13.8%	11.6%	9.8%	7.9%	6.1%
Ingenix PRG	133.5%	15.7%	13.3%	11.2%	9.4%	7.6%	5.9%
MedicaidRx	126.5%	17.6%	15.0%	12.6%	10.5%	8.5%	6.7%
Impact Pro	110.6%	20.1%	17.2%	14.5%	12.2%	9.9%	7.6%
Ingenix ERG	112.1%	21.0%	18.0%	15.3%	12.8%	10.3%	7.7%
ACG w/ Prior Cost**	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	2004 Cost Percentile Range						
Service Vendor	**0-100th**	**70th-100th**	**75th-100th**	**80th-100th**	**85th-100th**	**90th-100th**	**95th-100th**
MEDai**	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table VI.2

* These models do not include a concurrent option.

** These models include prior cost as input.

Table VI.3

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input.

Note: The 0–100th percentile values were not adjusted, but all other values were normalized by 0–100th percentile values. Unadjusted predictive ratios can be calculated by multiplying shown values by 0–100th percentile values.

TABLE VI.4 Predictive Ratios by 2004 Cost Percentile where <50th Percentile in 2003 (Prospective, Recalibrated, Nonlagged, with Prior Cost)

	2004 Cost Percentile Range						
Risk Adjuster Tool	**0-100th**	**70th-100th**	**75th-100th**	**80th-100th**	**85th-100th**	**90th-100th**	**95th-100th**
ACG	127.4%	21.4%	18.3%	15.6%	13.1%	10.5%	7.9%
CDPS	126.8%	21.3%	18.2%	15.4%	13.0%	10.5%	8.0%
Clinical Risk Groups	102.1%	21.6%	18.5%	15.7%	13.1%	10.6%	7.9%
DxCG DCG	119.1%	21.6%	18.5%	15.7%	13.2%	10.6%	8.0%
DxCG RxGroups	110.8%	20.9%	17.7%	14.9%	12.4%	10.1%	7.8%
Ingenix PRG	113.9%	20.6%	17.5%	14.6%	12.2%	9.9%	7.7%
MedicaidRx	106.4%	21.6%	18.3%	15.3%	12.8%	10.3%	8.0%
Impact Pro	106.3%	21.6%	18.4%	15.5%	13.0%	10.6%	8.2%
Ingenix ERG	103.8%	22.6%	19.3%	16.3%	13.7%	11.0%	8.3%
ACG w/ Prior Cost	120.4%	20.7%	17.7%	15.1%	12.6%	10.2%	7.7%
DxCG UW Model	99.8%	21.8%	18.7%	15.8%	13.3%	10.7%	8.0%
	2004 Cost Percentile Range						
Service Vendor	**0-100th**	**70th-100th**	**75th-100th**	**80th-100th**	**85th-100th**	**90th-100th**	**95th-100th**
MEDai	93.5%	22.0%	18.8%	15.9%	13.4%	10.7%	8.1%

Note: The 0–100th percentile values were not adjusted, but all other values were normalized by 0–100th percentile values. Unadjusted predictive ratios can be calculated by multiplying shown values by 0–100th percentile values.

Table VI.4 shows results similar to Table VI.3, except that results for risk adjusters that include prior costs are shown, and the prior cost independent variable was added to all of the models that do not already include prior costs. The DxCG UW Model is a very good predictive ratio for the total cohort of low-cost individuals as shown in the 0–100th column. ERG has the best predictive ratios for all but the 0–100th percentile columns.

Like any predictive modeling tool, the performance of risk adjusters is affected by a host of factors including data and usage limitations. These and other factors are detailed below.

Population Specificity and Applicability

Models can be calibrated so that they perform reasonably well for populations for which they were not originally intended. For example, CDPS and MedicaidRx were originally created for Chronic Disabled and Medicaid populations, respectively, but performed well when calibrated and applied to a commercial data set.[11] However, the condition category groupings and information presented may not be specific enough for the analysis being performed. For example, risk adjusters intended for an over-age-65 population may not include adequate breakdowns of pregnancy-related and infant diseases.

It is important to consider all of the objectives for which the risk adjuster will be used and what information will be gathered. The age/gender and condition categories need to be meaningful for the population being measured and for the purpose for which the tool is being used. Customization of the tools by risk adjuster vendors, outside consultants or in-house staff can provide meaningful improvements. However, modifications and calibrations should be made carefully.

Turnover

The population to which a risk adjuster is applied may include persons who will not be enrolled during the prediction period, because of lapse (voluntary or involuntary) or death. Likewise, new participants may enter the risk pool, and there will be only limited or no claims data available for them during the experience period.

Milliman's "Optimal Renewal Guidelines" study measured the predictive performance of pure age/gender predictions, in addition to optimized risk adjuster predictions. The prospective R-squared value for the age/gender prediction was about 6 percent. The prospective R-squared value for the optimized risk adjuster prediction was about 25 percent. Therefore, a rough estimate of the R-squared once turnover within a population is considered would be as follows:

> [(0.06 x turnover rate + 0.25 x (1 - turnover rate)) / 0.25]
> x Pre-turnover R-squared.

For example, assume that there is turnover of 15 percent (that is, you do not or will not have diagnosis or drug use data for 15 percent of the participants in the prediction period) and the R-squared without considering turnover is 27 percent (prospective) for a particular analysis. The adjusted R-squared calculated using the formula above would be about 24 percent [(0.06 x 0.15 + 0.25 x 0.85) / 0.25 x 0.27]. This approach does not consider partial enrollment. Some vendors have added logic to develop risk scores for participants who enter during the experience period.

This equation simply assumes that new entrants will receive an age/gender prediction. Further, it assumes that the change in predictive power is equal to the continuous enrollment (pre-turnover) R-squared, multiplied by a portion of the proportional change in predictive power from optimized to age/gender as observed in the "Optimal Renewal Guidelines" study. In the above equation, the turnover rate is defined as the portion of the population that will be active during the rating period that was not available during the experience period. This is a simplified, illustrative formula as it assumes changes in R-squared are linear, and does not consider partial enrollment during either or both of the experience and prediction periods.

[11] Cumming et al., "A Comparative Analysis of Claims Based Methods."

It may be more appropriate to use the pre-turnover R-squared in place of the 0.25 value in the formula above, as the age/gender performance may not change materially with changes in the risk-adjustment methods (although modeling conditions are important and affect both values, which is why the equation above is presented). The equation for the post-turnover R-squared (assuming the age/gender R-squared does not vary for different analyses) would be simplified as follows:

(0.06 x turnover rate + Pre-turnover R-squared x (1 - turnover rate)

The formula would also work for MAPE, and might even be more appropriate since MAPE does not square error terms.

Lag Issues

When using a risk adjuster, the prediction period often begins several months in the future. For example, when developing small group renewal rates, the rate development typically takes place three to six months in advance of the rating period. This delay is referred to as prediction lag, and it affects model performance above and beyond turnover, which was previously discussed (prediction lag creates uncertainty because of the additional time for potential changes in the health status of members). For any prospective analysis, the fact that future costs are being predicted creates uncertainty because an individual's health status may change. However, for purposes of this study, prediction lag is defined as the period between the end of the data collection period and the beginning of the prediction period.

Many of the risk-adjustment models are calibrated on continuously enrolled populations for a time period that immediately follows the experience period. Any time the conditions differ between the calibration of offered weights and the application of the risk adjuster, it is important to consider adjusting the model. Several of the risk-adjustment vendors include models with prediction lag options in their suite of tools. A modest prediction lag should not have a strong influence on model performance, especially if the model is recalibrated for the specific situation. However, prediction lag will increase the effects of turnover since it expands the period for potential turnover.

Data lag is related to, but not the same as, prediction lag. Data ready for risk adjuster input must be actual paid claims. Incurred medical claims usually take two to four months to be paid (on average), with some claims potentially taking several years to be completely paid. Prescription drug claims are paid much more quickly, but still take a month or two to be considered completely paid. Therefore, potentially meaningful and timely claims data may not be available for use in a risk adjuster in many situations. While vendors have added models to minimize lag issues, data lag affects the performance of all models, especially those that rely primarily on medical data.

In this study the impact of data and prediction lag was analyzed. Table VII.1 shows the combined impact of data and prediction lag collectively on model performance:

This table shows that predictive performance is substantially impacted by data and prediction lag. The risk adjusters based on only pharmacy data are less affected. In this study claims that were incurred and paid during January to August 2003 were used to predict claim costs for calendar year 2004. Thus, a four-month data and prediction lag for the "lagged" analyses was modeled.

Data delays are an implementation problem for any risk-adjustment model. A continuous enrollment requirement can remove up to 40 percent to 50 percent of any currently enrolled Medicaid population from the clinical condition risk assessment (e.g., all new enrollees), thus dramatically reducing the predictive performance of the total capitation system. Therefore, it is important to understand the extent to which the delay has affected the performance of the model.[12]

TABLE VII.1 R-Squared Prospective Recalibrated (Without Prior Cost, 250K Truncation)

		R-squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**Lagged**	**Nonlagged**	**Change**	**Lagged**	**Nonlagged**	**Change**
ACG	Diag	15.2%	19.6%	4.4%	92.8%	88.8%	-4.0%
CDPS	Diag	14.5%	17.7%	3.2%	95.1%	91.9%	-3.2%
Clinical Risk Groups*	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	16.9%	21.3%	4.4%	91.2%	87.0%	-4.2%
DxCG RxGroups	Rx	18.2%	20.5%	2.3%	87.2%	85.3%	-1.9%
Ingenix PRG	Rx	18.9%	21.2%	2.3%	87.6%	85.6%	-2.0%
MedicaidRx	Rx	15.8%	17.7%	1.9%	90.1%	88.4%	-1.7%
Impact Pro	Med+Rx+Use	21.5%	25.6%	4.1%	84.9%	81.6%	-3.3%
Ingenix ERG	Med+Rx	17.4%	20.0%	2.6%	88.4%	86.1%	-2.3%
ACG - w/ Prior Cost**	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model**	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Lagged**	**Nonlagged**	**Change**	**Lagged**	**Nonlagged**	**Change**
MEDai**	All	N/A	N/A	N/A	N/A	N/A	N/A

* Model could not be recalibrated consistently with other models.

** These models include prior cost as input.

Data Issues

From the perspective of data used to assess risk, methods can be categorized by their reliance on demographic, prior expenditure and/or health data, including self-reported health status and lab results. This study examines methods that use claims-based health data. The risk-adjustment methods based on claims data can be further divided into methods that rely on diagnosis codes from claims or encounter data, methods that rely on prescription data as a proxy for diagnoses and methods that use prior costs (and various combinations of the three data sources).

Models using other health data, such as lab results or survey data on self-reported chronic disease or functional status are not included in this study. Use of this information represents the next exciting frontier for predictive modeling. The increasing adoption of standardized formats for electronic medical records (EMRs) will likely accelerate the development and utility of predictive models that use this information.

Methods that rely solely on demographic risk factors, such as age, gender and program eligibility status, are easy to administer. These methods are not measures of the care process and therefore do not produce the incentive to change treatment or coding to maximize risk scores. Unfortunately, these methods have relatively poor predictive value at an individual level or for risk-skewed groups.

[12] Cumming et al., "A Comparative Analysis of Claims Based Methods."

In contrast, an individual's total prior medical expenditure is a reasonably good predictor of future expenditure. These data are easier to manage than detailed encounter data. However, the incentives related to providing care in an efficient manner are very poor.

Health status measures, such as diagnoses and prescriptions, are good predictors and provide useful medical management information. Diagnostic data must be obtained by plans from providers. Often these data are difficult for some types of plans to obtain either because the plan has a capitation contract with providers that do not require data for payment or the plans are staff or group provider models that have little or no fee-for-service experience. Ambulatory diagnoses are also somewhat unreliably coded, but the diagnostic risk assessment software available generally has built-in safeguards to reduce the problems caused by incomplete data.

Changes in coding patterns over time are expected. For diagnosis-based methods, a major concern with coding changes is for ambulatory diagnoses. These codes have not been widely used as the basis for payment or rate setting, although this use is becoming more common. For example, it is an important component of the HCC model used for Medicare Advantage payment. Changes in coding practices may result in the identification of new cases with a primary condition, the improved refinement of coding for severity or the increase in the coding of all related conditions affecting treatment. These changes can create the appearance of a higher-risk population when compared with the population used to calibrate the prediction model. The results can, therefore, inflate the estimate of the total cost for a population.

Another significant data issue is accessibility. Some plans or purchasers may have better access than others to prescription drug data. Prescription drug data are timely and relatively clean and complete for major ambulatory drugs. In addition, these data do not need to be obtained from providers, eliminating a potentially burdensome administrative step. The incentives for efficiency may be poor if prescribing is increased in order to raise a plan or provider's risk score. Prescription-based risk assessment models generally rely on drugs believed to be nondiscretionary. However, with off-label prescribing, and to the extent that discretion remains in prescribing drugs for additional diseases or for less severe or marginal forms of the disease, caution should be exercised when prescription-based models are considered for provider payment applications. Also, it is generally more important to periodically update and calibrate pharmacy-based models because of the rapid introduction of new drugs and off-label uses.

TABLE VII.2 Comparison of Risk Measures

	Risk Measures			
Criteria	**Demographics**	**Prior Expenditures**	**Prescriptions**	**Health Diagnoses**
Data Quality	High	Medium	High	Medium
Prediction Accuracy	Low	High	High	High
Administrative Burden	Low	Medium	Medium	High
Utilization Incentive	Low/None	High	High*	Low
Diagnosis Coding Incentive	Low/None	Low	Low	High

* High for prescription drugs, low for all other services.

Table VII.2 qualitatively compares types of risk assessment methods based on risk measures/data sources.

The methods evaluated in this study differ to some extent in the number of conditions they incorporate. Some use almost all known diseases to assign risk scores. Others exclude minor, acute conditions under the assumption that these conditions are not relevant to risk selection. The models assume that they do not represent significant per capita costs and including them may produce a clinically needless proliferation of these codes. However, if the intent is to evaluate how primary care providers are managing these frequent acute minor problems, then a model that includes these conditions would be preferred.

Another difference is the assignment of disease measures to risk categories. The process may produce categories that are much too heterogeneous for a specific disease of interest. Some conditions are lumped with related, yet clinically quite distinct, diseases due to similar costs. In addition, more detailed coding to describe severity will not change the assignment to a risk category beyond the simple

identification of the disease. On the other hand, a disease such as diabetes has its own category in most products, and payment is affected by coding diabetes more specifically.

The approach to assigning individual risk scores also varies. Some methods are additive, with additional payment made for each additional identified disease category, and others are multiplicative (nearly all are hierarchal at some level). For payment applications, some of these categories may be arranged in hierarchies of related conditions—for example, pulmonary conditions, with payment made for only the highest cost category in the hierarchy, the assumption being that the categories with lower costs in the hierarchy indicate complications related to the more significant condition. This approach avoids "double" counting. Other methods address this relatedness of conditions by assigning individuals to mutually exclusive risk categories derived by interacting all of the individual's conditions or by identifying the individual's dominant condition.

The methods evaluated in this study have been designed to be as robust to data problems as possible while preserving predictive performance. The models typically require only one occurrence of the diagnosis or prescription in the assessment period to assign risk. The number of times the same code appears is typically irrelevant. Discretionary or ill-defined indicators are often excluded or assigned so as to minimize gaming incentives. This means that data need not be perfectly complete and detailed to be adequate for risk-adjustment. [13]

Group Size and State Regulation in Employer Group Renewal Rating

State regulation often greatly limits the rating action that small group carriers can take based on the risk adjuster predictions by limiting allowable rate changes due to medical risk factors (ranging from +/-10 percent to unlimited depending on the state). Group size also affects the predictive performance of risk adjuster models, because as groups become larger, variations in individuals' costs are less important, therefore prediction accuracy increases. Large groups also tend to have future costs that are more predictable based on their historic costs than smaller groups.

To understand the impact of rating regulations on predictive performance, suppose two methods for predicting a small group's health care costs are used. One method estimates the group's costs as 30 percent higher than average, while the other method estimates the group's costs as 35 percent higher. With the benefit of hindsight and actual claim data, the group's costs turn out to be 30 percent higher than average. Depending on which state the carrier was operating in, either method may have provided the carrier with all of the useful information they could use for purposes of setting the group's renewal rate. For example, Iowa allows only +/- 25 percent variation from the average rate due to the health status of the group. Therefore, if this was an Iowa renewal, both methods would have directed the carrier to rate the group up as high as possible and would have provided "perfect information" (depending on your perspective). However, in states with 35 percent or more allowable rating variation, the first method provides better information.

The "Optimal Renewal Guidelines" study concluded that state regulatory limits on small group rating significantly impacted the actionable predictive power of renewal methods, including those that used risk adjusters. In addition, meaningful differences between methods decreased as group size increased.

Table VII.3, on the next page, shows how group size and regulatory rating limits affect MAPE (excerpt from "Optimal Renewal Guidelines" study). The Risk Adjuster results represent optimized risk adjuster results, including prior costs.

As shown in this table, the MAPE results for both a manual rate and risk adjuster approach improve as group size increases and when rating limits are introduced and tighten (for the MAPE calculations with rating limits, actual costs were limited by allowable rate variation, decreasing the potential error). Historic loss ratio methods performed better than the manual rate approach, and showed less difference compared to the risk adjuster approach.

Uses of Health-Based Risk-Adjustment

There are many uses for health-based risk-adjustment by purchasers and plans. When selecting a health-based risk-adjustment method, two primary features differentiate the applications:

[13] Cumming et al., "A Comparative Analysis of Claims Based Methods." This subsection is substantially the same as the referenced report; the current report updates changes in current data issues.

TABLE VII.3	Impact of Group Size and Regulatory Rating Limits (MAPE with and without Cap)			
Group	Risk Adjuster +/- 25%		Manual Rate +/- 25%	
Size	Uncapped	With Cap	Uncapped	With Cap
1 Mbr	82.7%	16.4%	101.0%	19.7%
1 EE	70.2%	16.7%	85.8%	21.0%
3 EEs	50.8%	16.9%	59.9%	21.3%
10 EEs	32.0%	16.1%	36.8%	20.0%
25 EEs	21.3%	14.8%	24.1%	17.6%
50 EEs	15.1%	12.6%	17.2%	14.7%
150 EEs	9.1%	8.9%	10.3%	10.2%

- Does the application involve payment to providers or plans?
- Does the application's perspective focus on targeted subpopulations, or is it global?

Using the two distinguishing characteristics, specific applications can be categorized for the following four uses.

Provider or Plan Payment—Global Perspective

These uses include health plan premium rate setting and provider capitation. Under these conditions any of the diagnosis-based methods may be preferred because they are good predictors and may introduce less of a gaming incentive than the prescription-based models. Prior cost models should not be used. Risk selection at the provider level is usually more extreme than risk selection across health plans. When capitation or volume target incentives are used to pay providers, the concerns with diagnosis gaming and overtreatment become important. The use of actual utilization data, such as prescriptions, to indicate a disease and increase payment should be avoided or approached with caution. Diagnosis data are not immune from gaming, but criteria exist for diagnosing many, if not most, major conditions, and this helps provide a basis for validation. An additional benefit of using health-based risk-adjustment for capitation is that providers have a strong incentive to provide the data.

Provider or Plan Payment—Targeted Perspective

These uses include setting disease management payment levels, for example, carve-outs, high-cost case management or disease-specific payments. The selection should be limited to diagnosis-based models to avoid perverse incentives. One would need to explore which of the methods best captures the severity and complications associated with managing a specific disease on the one hand and high-cost complex cases with many co-morbidities on the other. It may also be true that, for the diseases of interest, one could become satisfied that the prescription indicating the presence of the condition or its severity is nondiscretionary, and then prescription-based systems or a combination of systems may be considered. Prior cost models should not be used, although some cost threshold (similar to a stop loss provision in some hospital diagnosis-related group (DRG) contracts) might be appropriate to include as an adjustment to payment.

No Provider or Plan Payment—Global Perspective

These uses include setting defined premium and contribution levels for employers and employees (i.e., small group underwriting), provider efficiency profiling, total medical cost forecasting and budgeting. Any of the methods could be applied for these uses because secondary incentives are weak when payment is not involved. Other factors, such as the cost of data and other uses for the risk assessment information, would dominate the selection. A prior cost variable should be included in the prediction for small group underwriting, as it increases the predictive power of the methods considerably.

A relatively new use of health-based risk-adjustment in rate setting is to adjust employee premiums in defined contribution products. The use of risk-adjustment within consumer-directed health plans will likely become important as these plans are more widely adopted.

No Provider or Plan Payment—Targeted Perspective

These uses include high-cost case identification, individual underwriting and disease management program planning and budgeting. In addition to the standard selection criteria, the selection would be based on which method provides the most meaningful clinical categorization of individuals. [14]

[14] Cumming et al., "A Comparative Analysis of Claims Based Methods." This subsection is substantially the same as the referenced report; the current report provides updated information.

Medical and Pharmacy Data Issues

Implementation will be more challenging if there is not some early testing and data handling in the planning phase. A simulation may be the first time the purchaser will be handling massive amounts of data, especially the encounter data.

Data should be examined for reasonableness. Examining the frequency distributions of various data elements will help identify incomplete encounter data. Although there are no norms, there is some information about what portion of members should be expected not to have any claims. Data may be missing because of subcapitation or because of carve-outs. A common problem is missing mental health provider data for a program that covers mental health services. Each person should have similar benefit plans or normalizing adjustments, and additional modeling will be necessary. Any differences between the populations and benefits and methods for addressing those differences should be noted in results.

Different types of plans have a variety of data problems. Staff model HMOs that have limited experience with fee-for-service billing could have problems providing data for encounters and the bundling of services. Plans whose systems truncate the number of diagnosis codes per record may potentially result in understated risk measures.

Data quality can be an issue at the plan level and at the provider level. Data concerns at the plan level revolve around completeness, while data issues at the provider level include both completeness and accuracy.

For diagnosis data, the concern at the plan level is to capture all diagnoses already recorded by the provider. Plans may be missing diagnoses for two reasons:

- Incomplete or unavailable encounter data from some providers
- Truncation of the number of diagnoses per encounter supplied by the provider.

Prescription data are almost always complete and accurate at the plan level for most significant conditions and do not involve data transfer from providers.

For diagnosis coding at the provider level, there are three possible activities that can change the number and distribution of diagnoses and can increase the measured risk for a population when, in fact, the underlying morbidity of the population may be stable:

- Diagnostic discovery: Increased number and severity of diagnoses are reported, all of which are appropriate. The correction of previous underreporting will reduce the problem of lack of persistence of diagnoses and will more fairly represent the illness burden of the population.
- Diagnostic creep: Increased number and severity of diagnoses for cases where the diagnosis is uncertain. This represents an upward bias in response to payment incentives. Many of the groupers underlying many risk-adjustment methods try to minimize this problem by bundling related diagnoses and by excluding ill-defined codes.

- Tentative diagnoses: Represents a potential source of error when a diagnosis is appropriately used to justify a diagnostic procedure (rule-out) or to signal the need to treat a person without confirmatory diagnostic tests as if the patient has the disease (presumptive), because delay in treatment is harmful. Here, too, the groupers underlying many risk-adjustment methods have rules for excluding codes that are highly likely to be tentative.

Purchasers have so far not detected significant changes in provider-level coding patterns, but it is important to be vigilant and to set up monitoring and auditing systems that examine coding practices.[15]

Eligibility Data Issues

It may require two months or more to receive updates of changes in eligibility status of plan members from the purchaser. For some large employers, the retroactive adjustment for new enrollment, enrollment status changes or terminations may take even longer.

To the extent eligibility information is out of date, the risk scoring will also be affected and can be materially biased. For example, if it takes several months for eligibility data to reflect the death of members, then those members will appear healthy for some period of time after their death. This may affect concurrent risk-adjustment applications most significantly.

The Time to Execute the Risk Scoring and the Frequency of Risk Scoring

Purchasers can control how often and how fast they compute and assign risk scores. Combined with the usual claims run-out lag, the range can be from a minimum of six months up to 24 months.

Data delays are an implementation problem for any risk-adjustment model. For individual-level prospective models, the enrollee often must be continuously eligible for 6–12 months in the assessment period, 6–18 months in the claims delay period, and 1–12 months in the payment period for a health plan to be paid for the risk of that enrollee. A continuous enrollment requirement can remove up to 40 percent to 50 percent of any currently enrolled Medicaid population from the clinical condition risk assessment (e.g., all new enrollees), thus dramatically reducing the predictive performance of the total capitation system. Therefore, it is important to know the extent to which the delay has reduced the performance of the model compared to its "laboratory" tested results that often included no delay. Section VII of this report includes a discussion of the impact of lag on model performance.

[15] Cumming et al., "A Comparative Analysis of Claims Based Methods." This subsection is substantially the same as the referenced report.

Follow-up Studies | SECTION IX.

The following list identifies beneficial studies recommended for follow-up analysis. These studies would build on the results presented in this report and the two preceding SOA risk adjuster research studies.

- Explicitly analyze the impact of turnover (i.e., a non-continuously enrolled population)
- Analyze Medicare's risk assessment tool, HCC
- Analyze predictive measures for different, homogeneous populations (Medicare, Medicaid, individual, small group, large group, HMO, PPO, etc.)
- Analyze impact of adding prior costs to risk adjuster predictions by group size (and how credibility of risk adjuster and prior cost components changes with group size)
- Analyze consistency of performance (robustness) across different data sets and over time.
- Explicitly analyze the impact of small group regulation for all of the models; the general impact of state regulation is expected to be similar for the different models
- Analyze the predictive improvement (or expected improvement) when more than 12 months of data are used
- Analyze potential predictive performance improvements with the inclusion of lab, HRA and other available data
- Analyze additional models more appropriate for disease management uses of risk adjusters, and use measures more meaningful for these uses (i.e., specificity).

References | SECTION X.

Cumming, R. B., D. Knutson, B. A. Cameron, and B. Derrick. 2002. A Comparative Analysis of Claims-Based Methods of Health Risk Assessment for Commercial Populations. A research study sponsored by the Society of Actuaries. May 24, 2002.

D. Dunn, A. Rosenblatt, D. Taira, E. Latimer, J. Bertko, T. Stoiber, P. Braun, S. Busch. 1996. A Comparative Analysis of Methods of Health Risk Assessment. SOA Monograph M-HB96-1. October 1996.

APPENDIX A-1. Offered, Prospective, Nonlagged, without Prior Costs

TABLE A-1.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	20.8%	19.2%	16.2%	87.7%	89.9%	90.4%
CDPS	Diag	17.6%	14.9%	12.4%	93.4%	95.3%	95.8%
Clinical Risk Groups	Diag	19.3%	17.5%	14.9%	88.7%	90.9%	91.4%
DxCG DCG	Diag	22.3%	20.6%	17.4%	85.3%	87.5%	88.0%
DxCG RxGroups	Rx	23.8%	20.4%	16.8%	82.9%	85.3%	85.9%
Ingenix PRG	Rx	25.0%	20.5%	17.2%	83.4%	85.8%	86.4%
MedicaidRx	Rx	19.3%	15.8%	12.9%	87.3%	89.6%	90.2%
Impact Pro	Med+Rx+Use	26.3%	24.4%	21.3%	79.3%	81.8%	82.4%
Ingenix ERG	Med+Rx	23.7%	19.7%	16.2%	84.1%	86.4%	87.0%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-1.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	88.4%	100.0%	96.7%	103.1%	99.6%	92.3%
CDPS	Diag	95.0%	73.4%	84.8%	76.4%	67.3%	92.5%
Clinical Risk Groups	Diag	85.1%	94.7%	99.7%	99.5%	91.5%	89.0%
DxCG DCG	Diag	93.3%	98.3%	98.6%	103.2%	86.4%	95.9%
DxCG RxGroups	Rx	95.5%	76.9%	97.9%	89.4%	89.2%	88.6%
Ingenix PRG	Rx	94.9%	93.9%	98.2%	89.7%	79.6%	87.1%
MedicaidRx	Rx	90.1%	94.9%	92.7%	79.1%	90.8%	94.0%
Impact Pro	Med+Rx+Use	97.6%	115.4%	96.4%	99.8%	95.1%	98.0%
Ingenix ERG	Med+Rx	90.0%	99.2%	94.8%	92.9%	80.0%	91.9%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-1.3 Predictive Ratios by 2004 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	22.1%	42.0%	66.0%	97.4%	147.7%	261.1%	597.2%	9690.4%
CDPS	Diag	14.6%	32.0%	55.4%	87.1%	144.7%	285.5%	763.0%	12765.0%
Clinical Risk Groups	Diag	22.0%	41.1%	64.0%	96.0%	149.5%	261.3%	606.0%	9781.5%
DxCG DCG	Diag	23.4%	43.0%	67.0%	98.3%	148.8%	257.3%	562.8%	8454.6%
DxCG RxGroups	Rx	19.9%	45.2%	73.3%	105.1%	152.3%	243.8%	516.0%	8096.4%
Ingenix PRG	Rx	21.1%	46.6%	74.6%	104.7%	149.5%	239.5%	512.9%	8226.8%
MedicaidRx	Rx	16.0%	41.2%	72.2%	109.7%	166.1%	260.9%	496.3%	6130.0%
Impact Pro	Med+Rx+Use	30.0%	49.4%	72.4%	100.7%	146.8%	237.0%	493.6%	7396.0%
Ingenix ERG	Med+Rx	17.7%	40.2%	68.9%	106.3%	161.6%	263.3%	533.5%	7162.8%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

APPENDIX A-2. Offered, Prospective, Nonlagged, with Prior Costs

TABLE A-2.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	N/A	N/A	N/A	N/A	N/A	N/A
CDPS	Diag	N/A	N/A	N/A	N/A	N/A	N/A
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	N/A	N/A	N/A	N/A	N/A	N/A
ACG w/ Prior Cost	Diag+$Rx	25.6%	22.4%	18.7%	82.8%	85.1%	85.6%
DxCG UW Model	Diag+$Total	31.3%	27.4%	23.6%	79.0%	80.1%	80.4%
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-2.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	N/A	N/A	N/A	N/A	N/A	N/A
CDPS	Diag	N/A	N/A	N/A	N/A	N/A	N/A
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	N/A	N/A	N/A	N/A	N/A	N/A
ACG w/ Prior Cost	Diag+$Rx	92.5%	109.0%	95.8%	97.5%	103.6%	91.0%
DxCG UW Model	Diag+$Total	93.2%	84.9%	91.1%	90.7%	103.6%	94.6%
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-2.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
CDPS	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ACG w/ Prior Cost	Diag+$Rx	22.3%	46.6%	71.9%	98.8%	142.1%	241.6%	570.6%	10010.0%
DxCG UW Model	Diag+$Total	22.2%	45.6%	71.4%	102.2%	150.4%	246.0%	524.8%	8377.8%
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

APPENDIX A-3. Offered, Prospective, Lagged, without Prior Costs

TABLE A-3.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	15.6%	14.5%	12.3%	91.6%	93.7%	94.1%
CDPS	Diag	13.9%	11.9%	9.8%	96.9%	98.8%	99.2%
Clinical Risk Groups	Diag	16.0%	14.1%	12.1%	91.8%	93.9%	94.4%
DxCG DCG	Diag	16.8%	15.1%	12.6%	89.4%	91.6%	92.1%
DxCG RxGroups	Rx	21.1%	18.0%	14.8%	85.1%	87.4%	88.0%
Ingenix PRG	Rx	22.5%	18.0%	15.2%	85.3%	87.8%	88.3%
MedicaidRx	Rx	16.5%	13.6%	11.1%	89.4%	91.7%	92.3%
Impact Pro	Med+Rx+Use	24.2%	21.4%	18.2%	83.1%	85.5%	86.1%
Ingenix ERG	Med+Rx	20.4%	16.9%	13.9%	86.5%	88.7%	89.3%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-3.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	81.7%	96.2%	91.1%	92.8%	105.5%	86.3%
CDPS	Diag	89.1%	70.5%	80.8%	70.4%	64.2%	87.2%
Clinical Risk Groups	Diag	79.2%	83.2%	94.7%	89.7%	98.2%	84.9%
DxCG DCG	Diag	88.9%	94.8%	94.9%	95.2%	88.5%	92.3%
DxCG RxGroups	Rx	90.3%	74.3%	98.0%	86.6%	93.0%	85.3%
Ingenix PRG	Rx	90.8%	89.4%	97.9%	86.9%	83.6%	84.6%
MedicaidRx	Rx	89.3%	96.6%	95.8%	79.7%	97.4%	93.3%
Impact Pro	Med+Rx+Use	90.7%	92.1%	95.2%	89.9%	96.3%	91.4%
Ingenix ERG	Med+Rx	86.8%	99.4%	96.1%	89.1%	83.4%	90.4%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-3.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	18.9%	38.0%	61.6%	93.2%	149.3%	272.7%	641.0%	11015.6%
CDPS	Diag	13.0%	29.9%	52.4%	83.5%	142.9%	290.5%	810.3%	14295.3%
Clinical Risk Groups	Diag	18.9%	37.7%	60.1%	92.4%	148.3%	269.8%	660.6%	11255.2%
DxCG DCG	Diag	20.3%	40.2%	64.4%	96.3%	150.7%	266.5%	597.7%	9589.7%
DxCG RxGroups	Rx	18.4%	43.7%	71.8%	104.5%	153.0%	246.5%	528.9%	8702.9%
Ingenix PRG	Rx	19.9%	45.2%	72.9%	103.6%	150.2%	243.2%	528.3%	8849.3%
MedicaidRx	Rx	15.7%	41.3%	72.8%	111.0%	167.2%	258.2%	481.2%	6226.6%
Impact Pro	Med+Rx+Use	22.2%	43.8%	68.9%	100.4%	152.7%	253.3%	540.0%	8691.8%
Ingenix ERG	Med+Rx	16.4%	39.1%	68.1%	106.1%	163.4%	265.1%	536.8%	7570.0%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

APPENDIX A-4. Offered, Prospective, Lagged, with Prior Costs

TABLE A-4.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	N/A	N/A	N/A	N/A	N/A	N/A
CDPS	Diag	N/A	N/A	N/A	N/A	N/A	N/A
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	N/A	N/A	N/A	N/A	N/A	N/A
ACG w/ Prior Cost	Diag+$Rx	21.7%	18.7%	15.6%	85.8%	88.1%	88.6%
DxCG UW Model	Diag+$Total	25.2%	21.3%	17.8%	84.3%	85.3%	85.6%
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-4.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	N/A	N/A	N/A	N/A	N/A	N/A
CDPS	Diag	N/A	N/A	N/A	N/A	N/A	N/A
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	N/A	N/A	N/A	N/A	N/A	N/A
ACG w/ Prior Cost	Diag+$Rx	86.5%	102.5%	90.5%	87.5%	108.1%	85.9%
DxCG UW Model	Diag+$Total	86.3%	78.0%	86.0%	82.6%	96.0%	88.1%
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-4.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
CDPS	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
ACG w/ Prior Cost	Diag+$Rx	19.1%	43.3%	68.7%	96.2%	143.9%	249.9%	604.0%	11078.8%
DxCG UW Model	Diag+$Total	18.0%	40.7%	66.4%	98.6%	151.7%	260.7%	584.9%	10058.2%
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

APPENDIX A-5. Offered, Concurrent, Nonlagged, without Prior Costs

TABLE A-5.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	29.4%	29.7%	27.4%	73.0%	75.0%	75.4%
CDPS	Diag	35.5%	32.9%	31.0%	79.0%	80.6%	81.0%
Clinical Risk Groups	Diag	47.1%	43.3%	39.9%	68.6%	70.5%	70.9%
DxCG DCG	Diag	57.2%	51.8%	49.8%	61.6%	65.0%	65.4%
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	32.1%	28.1%	24.6%	77.2%	79.1%	79.6%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	46.5%	42.4%	38.6%	65.8%	67.7%	68.2%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-5.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	109.0%	97.3%	90.6%	94.0%	44.4%	107.5%
CDPS	Diag	102.3%	73.4%	87.6%	74.4%	65.2%	89.9%
Clinical Risk Groups	Diag	92.0%	103.8%	92.8%	87.5%	80.9%	89.9%
DxCG DCG	Diag	93.8%	109.9%	96.3%	103.4%	80.9%	92.3%
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	83.2%	82.6%	93.3%	65.3%	68.8%	79.6%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	88.6%	108.7%	92.9%	89.9%	70.7%	86.4%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-5.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	57.0%	82.8%	94.8%	100.2%	107.6%	124.3%	137.9%	133.4%
CDPS	Diag	44.9%	60.9%	73.3%	86.4%	106.0%	142.9%	195.1%	283.1%
Clinical Risk Groups	Diag	62.8%	76.7%	83.8%	92.6%	105.8%	129.0%	158.9%	208.4%
DxCG DCG	Diag	75.2%	84.6%	89.0%	94.3%	102.9%	120.3%	133.4%	151.2%
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	43.2%	70.9%	88.1%	102.3%	116.6%	129.8%	136.3%	154.6%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	54.4%	75.2%	88.4%	101.2%	114.0%	127.6%	134.9%	131.5%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

APPENDIX A-6. Offered, Concurrent, Lagged, without Prior Costs

TABLE A-6.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	25.0%	24.4%	23.3%	77.9%	78.6%	78.7%
CDPS	Diag	29.5%	27.1%	26.2%	85.8%	86.4%	86.5%
Clinical Risk Groups	Diag	40.8%	37.3%	35.7%	76.5%	77.2%	77.3%
DxCG DCG	Diag	50.5%	43.0%	41.5%	68.3%	71.2%	71.3%
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	25.4%	22.5%	21.3%	83.0%	83.7%	83.8%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	39.1%	35.6%	33.9%	72.0%	72.7%	72.8%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-6.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	111.6%	101.1%	91.5%	94.5%	47.7%	111.7%
CDPS	Diag	98.6%	73.8%	88.3%	73.7%	63.5%	89.9%
Clinical Risk Groups	Diag	87.9%	100.8%	90.0%	85.1%	93.5%	88.2%
DxCG DCG	Diag	94.0%	116.2%	99.5%	105.7%	87.8%	96.2%
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	83.0%	84.0%	99.5%	69.2%	71.2%	81.9%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	86.9%	114.3%	96.9%	91.2%	73.6%	89.9%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-6.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	62.2%	83.8%	93.4%	98.4%	104.5%	120.4%	133.9%	126.6%
CDPS	Diag	49.1%	62.5%	71.2%	82.0%	100.9%	136.6%	194.9%	299.6%
Clinical Risk Groups	Diag	65.9%	75.2%	79.5%	87.4%	101.0%	126.2%	166.9%	235.5%
DxCG DCG	Diag	82.1%	88.4%	89.6%	93.3%	100.3%	114.8%	126.3%	145.1%
DxCG RxGroups	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	51.1%	76.3%	89.7%	101.6%	112.4%	120.1%	121.7%	136.8%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	60.6%	79.4%	89.7%	100.8%	111.4%	119.6%	121.8%	116.1%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

APPENDIX A-7. Recalibrated, Prospective, Nonlagged, without Prior Costs

TABLE A-7.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	21.8%	19.6%	16.6%	86.9%	88.8%	89.3%
CDPS	Diag	20.8%	17.7%	14.7%	89.9%	91.9%	92.4%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	24.9%	21.3%	17.8%	85.0%	87.0%	87.6%
DxCG RxGroups	Rx	25.1%	20.5%	16.8%	82.8%	85.3%	85.9%
Ingenix PRG	Rx	25.6%	21.2%	17.6%	83.3%	85.6%	86.2%
MedicaidRx	Rx	22.2%	17.7%	14.6%	86.1%	88.4%	89.0%
Impact Pro	Med+Rx+Use	28.3%	25.6%	22.0%	79.5%	81.6%	82.2%
Ingenix ERG	Med+Rx	24.4%	20.0%	16.4%	83.8%	86.1%	86.8%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-7.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	98.3%	90.9%	96.2%	100.8%	99.1%	98.0%
CDPS	Diag	97.1%	81.3%	97.7%	93.5%	94.9%	91.1%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	93.5%	91.1%	97.5%	96.0%	92.5%	98.5%
DxCG RxGroups	Rx	95.2%	72.4%	95.7%	86.7%	84.0%	89.2%
Ingenix PRG	Rx	93.0%	73.2%	96.0%	86.3%	85.6%	87.4%
MedicaidRx	Rx	91.9%	74.0%	95.2%	78.8%	84.7%	88.1%
Impact Pro	Med+Rx+Use	99.3%	97.5%	98.3%	97.0%	101.6%	97.8%
Ingenix ERG	Med+Rx	97.3%	92.6%	99.4%	94.5%	81.5%	92.3%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-7.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	21.8%	42.5%	67.5%	100.0%	152.2%	265.0%	570.7%	8308.1%
CDPS	Diag	18.2%	38.4%	63.6%	96.8%	154.5%	275.1%	595.3%	9335.9%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	20.5%	41.7%	67.3%	100.1%	153.3%	263.6%	558.3%	7869.0%
DxCG RxGroups	Rx	18.2%	43.8%	72.8%	105.8%	155.0%	248.8%	516.9%	7914.0%
Ingenix PRG	Rx	19.2%	44.3%	72.6%	104.2%	152.9%	247.4%	523.9%	8301.4%
MedicaidRx	Rx	15.9%	40.1%	69.9%	107.0%	163.4%	261.9%	516.9%	7374.3%
Impact Pro	Med+Rx+Use	26.9%	48.3%	73.3%	103.9%	152.1%	241.4%	480.9%	6605.6%
Ingenix ERG	Med+Rx	18.0%	41.5%	71.1%	108.7%	163.6%	261.4%	509.2%	6171.7%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

APPENDIX A-8. Recalibrated, Prospective, Nonlagged, with Prior Costs

TABLE A-8.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	24.2%	23.0%	20.2%	84.6%	86.2%	86.6%
CDPS	Diag	27.4%	24.6%	21.2%	83.7%	85.6%	86.3%
Clinical Risk Groups	Diag	21.5%	20.5%	18.4%	85.2%	86.6%	87.0%
DxCG DCG	Diag	29.7%	26.5%	22.9%	80.5%	82.5%	83.2%
DxCG RxGroups	Rx	30.6%	27.1%	23.4%	78.7%	80.7%	81.4%
Ingenix PRG	Rx	30.9%	27.4%	23.7%	78.9%	80.9%	81.5%
MedicaidRx	Rx	29.7%	26.3%	22.7%	79.9%	81.9%	82.6%
Impact Pro	Med+Rx+Use	29.3%	27.2%	24.0%	78.7%	80.6%	81.2%
Ingenix ERG	Med+Rx	30.0%	26.5%	22.8%	79.1%	81.2%	81.9%
ACG w/ Prior Cost	Diag+$Rx	27.7%	25.4%	22.1%	80.3%	82.1%	82.6%
DxCG UW Model	Diag+$Total	33.1%	29.1%	25.2%	76.1%	78.3%	78.9%
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	35.7%	32.1%	27.6%	73.0%	75.2%	75.6%

TABLE A-8.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	99.0%	91.0%	100.1%	105.6%	115.5%	99.2%
CDPS	Diag	93.2%	86.6%	99.7%	96.8%	96.4%	94.6%
Clinical Risk Groups	Diag	96.5%	110.2%	110.0%	115.8%	109.3%	101.8%
DxCG DCG	Diag	95.8%	90.2%	99.2%	96.2%	99.3%	100.3%
DxCG RxGroups	Rx	101.2%	79.6%	99.0%	97.0%	94.6%	96.8%
Ingenix PRG	Rx	97.9%	80.0%	98.4%	96.4%	93.5%	94.9%
MedicaidRx	Rx	97.9%	84.2%	98.7%	95.2%	96.3%	96.8%
Impact Pro	Med+Rx+Use	100.8%	99.9%	99.5%	98.6%	106.5%	100.0%
Ingenix ERG	Med+Rx	99.8%	92.6%	101.0%	97.8%	92.7%	97.2%
ACG w/ Prior Cost	Diag+$Rx	100.7%	101.0%	100.5%	102.5%	119.1%	100.1%
DxCG UW Model	Diag+$Total	99.1%	93.1%	100.7%	97.6%	107.3%	101.0%
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	104.4%	93.3%	102.6%	97.9%	96.1%	99.7%

TABLE A-8.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	27.1%	46.7%	69.6%	99.1%	146.5%	249.9%	544.2%	8433.1%
CDPS	Diag	24.2%	43.8%	67.8%	98.6%	150.4%	256.7%	546.1%	8537.4%
Clinical Risk Groups	Diag	28.4%	49.2%	73.0%	103.5%	150.4%	238.8%	488.7%	6808.8%
DxCG DCG	Diag	25.2%	45.6%	70.4%	101.1%	149.7%	248.5%	528.7%	7780.7%
DxCG RxGroups	Rx	24.9%	48.0%	75.0%	105.4%	151.3%	237.3%	482.6%	7177.5%
Ingenix PRG	Rx	25.0%	48.0%	74.5%	104.4%	150.6%	238.0%	489.1%	7426.9%
MedicaidRx	Rx	24.2%	46.4%	73.4%	106.2%	155.8%	243.8%	478.5%	6773.7%
Impact Pro	Med+Rx+Use	29.7%	50.6%	74.9%	103.6%	149.5%	235.0%	470.1%	6587.2%
Ingenix ERG	Med+Rx	24.3%	46.1%	73.6%	107.4%	156.4%	245.1%	482.0%	6226.3%
ACG w/ Prior Cost	Diag+$Rx	27.2%	51.7%	76.5%	102.1%	141.7%	230.3%	510.3%	8146.4%
DxCG UW Model	Diag+$Total	26.8%	50.9%	77.4%	107.6%	150.4%	229.0%	452.4%	6427.8%
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	29.5%	52.5%	78.0%	106.5%	145.4%	216.2%	411.9%	5592.5%

APPENDIX A-9. Recalibrated, Prospective, Lagged, without Prior Costs

TABLE A-9.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	16.8%	15.2%	12.8%	90.9%	92.8%	93.3%
CDPS	Diag	17.3%	14.5%	12.0%	93.1%	95.1%	95.7%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	20.3%	16.9%	13.9%	89.1%	91.2%	91.7%
DxCG RxGroups	Rx	22.7%	18.2%	14.9%	84.8%	87.2%	87.9%
Ingenix PRG	Rx	23.3%	18.9%	15.6%	85.3%	87.6%	88.2%
MedicaidRx	Rx	20.1%	15.8%	12.8%	87.8%	90.1%	90.7%
Impact Pro	Med+Rx+Use	24.9%	21.5%	18.2%	82.7%	84.9%	85.6%
Ingenix ERG	Med+Rx	21.6%	17.4%	14.3%	86.1%	88.4%	89.0%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-9.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	90.9%	82.3%	89.8%	89.7%	104.1%	89.5%
CDPS	Diag	88.9%	73.4%	94.7%	83.9%	95.1%	84.3%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	87.2%	81.0%	92.2%	84.7%	91.7%	90.8%
DxCG RxGroups	Rx	91.1%	68.1%	94.5%	81.5%	79.9%	85.2%
Ingenix PRG	Rx	89.6%	70.2%	94.5%	81.8%	81.6%	84.4%
MedicaidRx	Rx	89.0%	69.4%	93.2%	74.4%	79.3%	84.3%
Impact Pro	Med+Rx+Use	96.3%	85.0%	98.2%	90.3%	97.1%	92.6%
Ingenix ERG	Med+Rx	94.7%	82.4%	98.0%	87.6%	81.8%	88.6%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-9.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	18.4%	38.0%	62.2%	94.7%	152.7%	277.0%	625.3%	10186.8%
CDPS	Diag	15.8%	35.4%	59.9%	93.0%	154.5%	283.8%	639.3%	10974.9%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	16.9%	37.4%	62.6%	95.7%	153.5%	274.9%	619.0%	10000.0%
DxCG RxGroups	Rx	16.3%	41.6%	70.5%	104.5%	155.7%	253.7%	539.2%	8725.9%
Ingenix PRG	Rx	17.3%	42.1%	70.2%	102.9%	153.8%	253.1%	548.1%	9089.1%
MedicaidRx	Rx	14.3%	38.1%	67.3%	104.8%	162.7%	266.2%	546.3%	8485.7%
Impact Pro	Med+Rx+Use	22.0%	44.8%	71.1%	103.6%	155.3%	250.4%	515.6%	7683.3%
Ingenix ERG	Med+Rx	16.0%	39.1%	68.4%	106.3%	163.9%	266.3%	536.7%	7392.6%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

APPENDIX A-10. Recalibrated, Prospective, Lagged, with Prior Costs

TABLE A-10.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	18.0%	16.6%	14.3%	89.6%	91.2%	91.6%
CDPS	Diag	21.0%	17.9%	15.1%	89.3%	91.2%	91.7%
Clinical Risk Groups	Diag	17.3%	15.6%	13.6%	89.0%	90.6%	91.0%
DxCG DCG	Diag	23.0%	19.5%	16.4%	86.4%	88.3%	88.9%
DxCG RxGroups	Rx	25.3%	21.1%	17.7%	82.7%	84.9%	85.5%
Ingenix PRG	Rx	25.9%	21.7%	18.2%	82.9%	85.1%	85.6%
MedicaidRx	Rx	24.1%	19.9%	16.7%	84.5%	86.6%	87.1%
Impact Pro	Med+Rx+Use	25.4%	22.1%	18.9%	82.2%	84.2%	84.8%
Ingenix ERG	Med+Rx	24.5%	20.4%	17.1%	83.6%	85.8%	86.4%
ACG w/ Prior Cost	Diag+$Rx	23.0%	20.1%	17.0%	84.3%	86.2%	86.7%
DxCG UW Model	Diag+$Total	26.5%	22.0%	18.4%	82.0%	84.0%	84.6%
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	28.3%	24.1%	20.1%	79.7%	81.6%	81.5%

TABLE A-10.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	91.1%	81.8%	91.9%	92.2%	114.8%	90.2%
CDPS	Diag	86.7%	76.0%	95.8%	86.1%	100.3%	86.8%
Clinical Risk Groups	Diag	86.8%	95.5%	102.6%	100.9%	110.7%	93.2%
DxCG DCG	Diag	89.7%	80.2%	93.4%	85.4%	99.0%	92.8%
DxCG RxGroups	Rx	95.2%	72.8%	96.4%	87.7%	89.8%	89.8%
Ingenix PRG	Rx	93.2%	74.5%	96.3%	88.0%	86.6%	88.5%
MedicaidRx	Rx	92.5%	76.6%	95.6%	84.6%	89.9%	89.8%
Impact Pro	Med+Rx+Use	96.3%	85.2%	98.5%	91.3%	104.9%	94.4%
Ingenix ERG	Med+Rx	95.3%	81.5%	98.9%	89.7%	89.8%	91.5%
ACG w/ Prior Cost	Diag+$Rx	94.0%	91.4%	93.6%	90.7%	117.8%	92.1%
DxCG UW Model	Diag+$Total	92.7%	81.3%	95.6%	86.8%	104.4%	93.0%
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	95.6%	80.6%	97.9%	88.2%	96.0%	89.4%

TABLE A-10.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	20.9%	40.3%	63.6%	94.6%	149.5%	268.4%	611.4%	10234.1%
CDPS	Diag	18.9%	38.6%	62.9%	94.8%	152.6%	272.8%	607.6%	10339.0%
Clinical Risk Groups	Diag	22.5%	43.1%	66.9%	99.1%	152.0%	256.2%	560.1%	8816.9%
DxCG DCG	Diag	19.4%	40.0%	65.1%	97.2%	152.2%	265.6%	593.1%	9708.7%
DxCG RxGroups	Rx	19.5%	43.8%	72.0%	104.7%	154.2%	248.1%	522.0%	8314.8%
Ingenix PRG	Rx	20.0%	44.1%	71.7%	103.6%	153.0%	247.8%	526.5%	8542.2%
MedicaidRx	Rx	18.5%	41.6%	69.9%	105.1%	159.2%	256.3%	520.7%	8035.0%
Impact Pro	Med+Rx+Use	23.3%	46.1%	72.1%	103.4%	153.7%	246.8%	508.8%	7657.8%
Ingenix ERG	Med+Rx	19.0%	41.5%	70.0%	105.8%	160.0%	257.6%	523.0%	7385.3%
ACG w/ Prior Cost	Diag+$Rx	21.3%	46.3%	72.1%	99.4%	145.3%	244.3%	559.0%	9615.7%
DxCG UW Model	Diag+$Total	20.5%	44.3%	70.9%	102.9%	153.0%	249.8%	530.6%	8445.6%
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	21.5%	46.9%	73.6%	104.5%	148.7%	225.6%	443.2%	6853.3%

APPENDIX A-11. Recalibrated, Concurrent, Nonlagged, without Prior Costs

TABLE A-11.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	32.3%	31.5%	28.7%	75.2%	76.6%	77.0%
CDPS	Diag	38.3%	36.8%	35.2%	78.0%	79.6%	80.1%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	58.0%	54.5%	51.0%	61.3%	63.4%	64.1%
DxCG RxGroups	Rx	41.8%	36.9%	32.8%	70.0%	72.4%	73.0%
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	36.1%	31.0%	27.3%	75.7%	78.0%	78.5%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	48.1%	43.3%	39.5%	65.3%	68.0%	68.9%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-11.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	103.2%	102.5%	88.8%	91.3%	41.0%	100.6%
CDPS	Diag	104.7%	76.5%	87.1%	83.8%	80.1%	80.2%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	92.9%	98.4%	93.0%	95.8%	83.3%	94.7%
DxCG RxGroups	Rx	85.8%	79.7%	89.4%	75.2%	67.6%	79.6%
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	85.8%	75.9%	90.1%	65.0%	73.2%	79.9%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	92.5%	96.6%	93.7%	89.8%	74.8%	85.2%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-11.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	55.9%	80.6%	92.7%	99.3%	109.7%	128.3%	139.5%	129.2%
CDPS	Diag	51.6%	66.5%	76.4%	86.9%	104.6%	137.8%	185.1%	267.2%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	65.3%	79.4%	87.0%	94.6%	105.7%	125.4%	141.9%	157.6%
DxCG RxGroups	Rx	51.9%	77.8%	90.6%	98.9%	106.9%	120.3%	140.2%	197.6%
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	44.1%	71.7%	88.6%	102.1%	115.4%	128.3%	136.4%	161.3%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	54.7%	76.0%	88.6%	100.6%	113.1%	127.3%	136.1%	135.9%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

APPENDIX A-12. Recalibrated, Concurrent, Lagged, without Prior Costs

TABLE A-12.1 R-Squared and MAPE % by Truncation Level

		R-Squared			MAPE%		
Risk Adjuster Tool	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
ACG	Diag	24.6%	24.2%	23.1%	81.3%	81.7%	81.8%
CDPS	Diag	32.3%	30.2%	29.3%	84.8%	85.5%	85.6%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	51.5%	47.4%	45.5%	67.7%	68.8%	68.9%
DxCG RxGroups	Rx	35.0%	31.1%	29.5%	76.1%	77.1%	77.2%
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	29.5%	25.9%	24.5%	81.2%	82.0%	82.1%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	40.6%	36.5%	34.8%	70.9%	72.0%	72.2%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**100K**	**250K**	**None**	**100K**	**250K**	**None**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-12.2 Predictive Ratios by Medical Condition in 2003 (250K Truncation)

Risk Adjuster Tool	Inputs	Asthma	Breast Cancer	Diabetes	Heart Disease	HIV	Mental Illness
ACG	Diag	109.8%	111.8%	91.2%	93.8%	45.6%	104.3%
CDPS	Diag	105.1%	78.7%	87.1%	83.2%	70.8%	78.3%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	93.3%	103.2%	94.3%	96.8%	75.2%	96.1%
DxCG RxGroups	Rx	83.6%	76.7%	88.4%	73.8%	61.8%	78.3%
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	86.5%	74.3%	90.5%	65.8%	64.6%	80.3%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	92.9%	99.3%	94.6%	89.8%	67.6%	85.3%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**Asthma**	**Breast Cancer**	**Diabetes**	**Heart Disease**	**HIV**	**Mental Illness**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A

TABLE A-12.3 Predictive Ratios by 2003 Cost Quintile (250K Truncation)

Risk Adjuster Tool	Inputs	99-100	96-99	90-96	80-90	60-80	40-60	20-40	0-20
ACG	Diag	62.9%	83.3%	92.8%	98.3%	105.9%	121.7%	131.9%	119.5%
CDPS	Diag	53.8%	66.9%	74.2%	83.2%	100.7%	132.9%	185.5%	281.6%
Clinical Risk Groups	Diag	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG DCG	Diag	67.7%	80.6%	86.4%	93.2%	103.4%	121.6%	138.5%	155.1%
DxCG RxGroups	Rx	55.2%	78.8%	88.4%	95.7%	103.2%	116.8%	139.3%	198.5%
Ingenix PRG	Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
MedicaidRx	Rx	48.8%	74.5%	87.9%	99.9%	111.0%	122.2%	130.3%	154.7%
Impact Pro	Med+Rx+Use	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ingenix ERG	Med+Rx	58.4%	77.8%	88.0%	99.2%	110.8%	122.4%	129.0%	125.0%
ACG w/ Prior Cost	Diag+$Rx	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DxCG UW Model	Diag+$Total	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Service Vendor	**Inputs**	**99-100**	**96-99**	**90-96**	**80-90**	**60-80**	**40-60**	**20-40**	**0-20**
MEDai	All	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Notes

Notes

Notes

SOCIETY OF ACTUARIES

475 N. Martingale Road, Suite 600
Schaumburg, IL 60173