



Affiliated Contractors (ACs) refers to Carriers, DMERCs, and FIs.

CONTENTS

Predictive Modeling and Data Display Special Study Update

Predictive Modeling and Data Display Special Study Update

The Centers for Medicare & Medicaid Services (CMS) awarded a Special Study: Predictive Modeling and Data Display to Livanta LLC (“Livanta”) on July 13, 2005. The goal of the Special Study was to explore the feasibility and identify possible applications to detect improper Medicare payments. The Special Study consists of 2 tasks: Task 1 is identifying tools used for Predictive Modeling, and Task 2 is identifying tools used for Data Display. Livanta broke each of those tasks into 2 evaluation phases: Phase 1 was conducted using PowerPoint slides submitted by the vendors, and Phase 2 was hosting a live demonstration of the products. The background, timeline, methodology and status were presented in the January Newsletter.

Task 1 - Predictive Modeling Results

Six vendors submitted predictive modeling results for the test file for Task 1 and were evaluated by the panel during Phase 1 using the summarized criteria presented below.

Summarized Task 1 Scoresheet

Priority	Task 1 Criterion	Weighted Score
1	Ability to predict anomalies; model validity	1000
2	Identify claim level anomalies and classify anomalies	100
3	Maintenance: models easily updated; output easy to retrieve and utilize	90

4	Cost of initial implementation and deployment	96
	Phase 1 TOTAL SCORE:	1286

Detailed criteria and scores for each vendor will not be presented, but the group’s results are summarized in the following paragraphs.

Out of a total possible weighted score of 1,286, the highest ranking model only achieved a weighted score of 559. In other words, the best model achieved a “grade” of 43% (559/1286), which was the direct result of low scores for the first and most important criterion – model predictive power.

Each claim type was analyzed separately. The following statistics were calculated for DME, Part A, and Part B:

- Confusion Matrix based from which the following statistics were calculated and compared: Accuracy, True positive rate (sensitivity), False positive rate, True negative rate (specificity), False negative rate.
- Based on probabilities provided by the vendors, the following additional statistics were calculated: Receiver Operating Characteristic (ROC) curve, Area under the ROC curve (C-statistic), Lift.

Numerous comparison charts and graphs for each claim type were created and evaluated by the panel. The key factor used for determining the score for the first evaluation criteria was the area under the ROC curve. Calculating the area under the ROC curve is important because it provides a measure for how much more effective the model’s results are beyond chance. In this way, the various models’ performance can legitimately be compared.

The following table indicates the effectiveness of all models combined in identifying the claims known to be in error. A large number of errors of each claim type (30%-41% of total) were not found by any model.

Number of Models Predicted Error Matrix

Claim Type	# of Models	# of Claims in Error That Were Accurately Detected	Claim Type	# of Models	# of Claims in Error That Were Accurately Detected	Claim Type	# of Models	# of Claims in Error That Were Accurately Detected
DME	0	203	Part A	0	1329	Part B	0	1786
	1	187		1	1321		1	2221
	2	73		2	1080		2	776
	3	30		3	441		3	217
	4	6		4	188		4	72
	5	0		5	61		5	3
	6	0		6	23		6	0
TOTAL		499	TOTAL		4,443	TOTAL		5,075

The next table shows predictive modeling results for the vendor with the best-performing model for each claim type using standardized cutoff points.

Best Performing Vendor Results by Claim Type

DME (Vendor 1)		
% of Population Selected	Error Rate in Review	% of All Claim Lines in Error Detected
1.0%	39.1%	6.2%
5.0%	25.8%	20.5%
10.0%	20.2%	32.1%
50.0%	9.0%	71.6%

Part A (Vendor 2)		
% of Population Selected	Error Rate in Selection	% of All Claims in Error Detected
1.0%	38.1%	4.3%
5.0%	29.7%	17.0%
10.0%	25.4%	29.0%
50.0%	13.9%	79.6%

Part B (Vendor 1)		
% of Population Selected	Error Rate in Selection	% of All Claim Lines in Error Detected
1.0%	46.9%	5.5%
5.0%	37.2%	21.7%
10.0%	30.2%	35.2%
50.0%	14.1%	82.2%

With the Vendor 1 model for DME, a selection and review of 1% of the total claim lines would be able to find 6.2% of total number of claim lines in error, an error rate (“hit rate”) of 39.1% within the selected claim lines. Similarly, with the Vendor 2 model for Part A, selecting and reviewing 1% of the all claims would identify 4.3% of the total number of claims in error, an error rate of 38.1% within the selected claims. Finally, with the Vendor 1 model for Part B, selecting and reviewing 1% of the all claim lines would be able to find 5.5% of total number of claim lines in error, representing an error rate of 46.9% within the selected claim lines.

Since most ACs can only review less than 1% of claims, the most effective model would have to have a hit rate substantially greater than 50% in order to select claims (or claim lines) for efficient and productive AC review. Based in the data shown, there is no single model or vendor that meets this standard.

As shown, no vendor submitted a model that performed efficiently enough under the current constraints to warrant proceeding to Phase 2, live demonstration. In addition, it does not appear that the claims sample available from CERT has sufficient depth to make predictive modeling valuable *within the CERT framework*. Four vendors, however, had models that could be included in subsequent activities. Panel recommendations for Task 1, Predictive Modeling, will be in the March Newsletter, along with the identification of the four vendors and their products.

Task 2 - Data Display Status

Four vendors participated in the live demonstrations in January, 2006. They were scored based on the following criteria for Phases 1 and 2.

Summarized Task 2 Scoresheet

Priority	Task 2 Criteria – Phase 1	Weighted Score	Task 2 Criteria – Phase 2	Weighted Score
1	Individual contractor data	900	Live demonstration	48
2	Graphical User Interface	702	Ease in changing display criteria, user friendly, graphic capability	36
3	Drill down reports	750	Display methodology	12
4	Comparative analysis	510	Techniques	18
5	Ease of use and training	405	Integration in business processes	18
6	Flexibility	150	Final cost	18
7	Data management	138		
8	Export and interface capability	102		
9	Cost	600		
	Phase 1 TOTAL SCORE:	4257	Phase 2 TOTAL SCORE:	150

The two top vendors scored a combined 'grade' 80.89% and 80.08%. Those two vendors are being asked to provide best and final pricing to Livanta, based on a Data Display Requirements document, in early March. Final recommendations will be made by the panel at that time, and the results will be published in the March Newsletter.

The purpose of the CERT Newsletter is to provide for an exchange of information among the Centers for Medicare and Medicaid Services (CMS), the CERT Review Contractor (CRC), the CERT Documentation Contractor (CDC), Affiliated Contractors (ACs) and Providers. The Newsletter is not intended to set CMS policy or replace CMS directives. The newsletter is published monthly by CDC. Archived copies are available on the CERT Website: <http://www.certprovider.org>

Send in questions, suggestions, and/or articles for inclusion in the newsletter to marylou@certcdc.com
 Deadline for March issue is March 10.
