A Medical Utilization Review System for Ambulatory Care Based on Automated Claims Data

Sharron Silva, M.A.; Allen Berkowitz, Ph.D.; Suzanne Lizanich-Aro, R.N., M.P.H.; and Patricia Jenkins, Ph.D.
United Mine Workers Health and Retirement Funds
2021 K Street, N. W.
Washington, D. C. 20006

ABSTRACT
This article describes the use of an automated claims data base in reviewing the quality and medical necessity of ambulatory care provided to beneficiaries. The primary advantage to this approach is economy, since the data elements used in the review process are routinely collected as part of the claims payment process. This paper addresses issues relevant to the selection of physicians for review, strategies for improving the review process, and results of reviews conducted to date.

Over the course of the last four years, the United Mine Workers of America Health and Retirement Funds has developed and implemented a medical utilization review (UR) system based on the review of claims data. The system has two primary purposes: (1) to improve the quality of medical care provided to Funds' beneficiaries, and (2) to reduce Funds' expenditures for unnecessary medical services. The approach we have selected is unique in that it relies primarily on a computerized claims data base in order to select physicians for review and in order to reach a preliminary assessment of a physician's pattern of care.

To date, we have had the opportunity to examine statistically the practices of more than 600 primary care physicians who account for yearly billings in excess of $33 million. In addition, we have conducted peer review on more than sixty physicians and have intervened in nearly forty practices. Based on these experiences, we have gained valuable insights into effective strategies for using computerized claims data in the review of ambulatory care. In conducting such reviews, the Funds has one important advantage. The vast majority of Funds beneficiaries are elderly. Thus, the beneficiary population is relatively homogeneous when compared to typical beneficiary populations for many third-party payors.

This paper will describe the data elements used in this review process, the statistical procedures used to select physicians for review, problems associated with the peer review process, and results to date.

DATA ELEMENTS USED IN REVIEW
The claims data base maintained by the Funds consists primarily of those items needed to accurately process payment of claims. The information on this data base which is of particular relevance to the utilization review program is the following: (1) date of service; (2) primary diagnosis (coded using H-ICDA 2); (3) description of service (coded using CPT-4, modified); (4) quantity of service; (5) beneficiary identifier; (6) age of beneficiary; (7) sex of beneficiary; (8) physician identifier; and (9) specialty of physician.

SELECTION OF PHYSICIANS FOR REVIEW
In selecting physicians for peer review, we rely on a variety of measures of physician practice. Three examples are: (1) Visits per Beneficiary Seen; (2) Diagnostic Tests per Beneficiary Seen; and (3) Injections per Beneficiary Seen. These measures have been previously validated as part of a grant from the Health Care Finance Administration.* The comparability of these measures, when applied to individual physicians, is ensured by comparing physicians only to other physicians within the same specialty. In addition, the measures are computed by calendar quarter and are based on the number of beneficiaries seen by a particular physician.

Reports are generated based on these measures as part of the claims payment process. The resulting measures are transmitted to the Funds both as a hard copy report and on tape for secondary analysis. This secondary analysis is conducted by Funds staff using SAS computer programs to implement specific selection criteria.

When the UR system was first designed, the procedure proposed for selecting physicians for
review was to use the hard copy reports to examine the distribution of physicians on each of the UR practice measures. High values would then be compared to a regional mean, using a t-test. This strategy has since been modified in several ways. First, since the observed distributions departed significantly from normal distributions, the median was substituted for the mean as a comparison value. Second, a percentile cut-off was added to ensure that the physicians selected for review would be those with values that were high not only in relation to the median, but in relation to the total distribution as well.

In light of these and other changes in selection strategies, the use of SAS programs to implement selection procedures has proved quite effective. We have been able to produce a variety of analytic aids such as bar graphs showing the distribution of physicians on various measures. (See Figure 1.) We have also been able to conduct statistical tests accurately and quickly, even when it is necessary to examine the implications of using a variety of statistical procedures for selection.

<table>
<thead>
<tr>
<th>Percent of GP's in Each Category</th>
<th>0%</th>
<th>10%</th>
<th>20%</th>
<th>30%</th>
<th>40%</th>
<th>50%</th>
<th>60%</th>
<th>70%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td>Injections Per Beneficiary Seen</td>
<td>0.25</td>
<td>1.00</td>
<td>1.75</td>
<td>2.75</td>
<td>3.50</td>
<td>4.00+</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1

Distribution of General Practitioners by Average Number of Injections per Beneficiary Seen

It should be noted, however, that this approach does not resolve all potential problems in maintaining valid selection procedures. There are several issues related to the coding system and the interpretation of these physician practice measures that merit discussion. It is important to understand, for example, that coding systems are dynamic entities. Coding at the Funds is based on CPT-4, but adaptations have been made in the system to account for specific needs of the claims processing system. Over time, as costs and patterns of practice change, codes can change as well. The result is that, periodically, procedures that are counted under one code may be changed to another. Thus, it is important to monitor proposed coding changes to determine their effect on the UR physician practice measures in order to ensure that we are examining comparable measures across time.

A related issue is that occasionally the significance of a particular service may change, and associated computer programs may be adapted to the new need. One example we encountered relates to the fact that the Funds instituted copayments based on physician visits several years ago. This led to a change in the way that visits were counted. A marked change in the visit measure alerted us to a problem. As a result, we had to define a new algorithm for counting visits for the UR system, to ensure that we maintained a consistent measure.

Another factor that has had an impact on our selection measures relates to changing medical technology. Increasingly, we find that physicians are using automated panels in place of individual tests. This phenomenon tends to distort the meaning of our diagnostic test measure.

There are also situational factors that affect our ability to apply these selection measures uniformly to all physicians. For example, many physicians order diagnostic tests through outside laboratories. The bills for such tests are sometimes submitted directly by the laboratory rather than by the physician. This problem limits our ability to select physicians who make less-than-expected use of diagnostic tests, but it does not affect our ability to identify overutilizers.

THE PEER REVIEW PROCESS

As mentioned previously, computer-generated profiles are used in the initial peer review process. Figure 2 shows a sample page from such a profile as seen by the review physicians.

Initially, such profiles were generated manually from physician billing records. More recently, we have developed computer specifications and obtained computer-generated profiles.

One aspect of these profiles that should be noted is that they contain additional information not shown in Figure 2 that is important in supporting the review process. This includes (1) identifying information on the provider; (2) beneficiary information; and (3) information that permits location of the original claim. At various stages of the review and intervention process, such data may be required.

These profiles can show the medical care delivered to all Funds beneficiaries seen by a physician over a specified period of time, or a more limited profile may be ordered to show only...
the medical care delivered to a random sample of beneficiaries. In general, a random sample of fifteen or twenty cases is sufficient to assess patterns of care. The profiles can also include all care given to specific beneficiaries by any provider during the same time frame.

**Figure 2**
Sample Profile Page

<table>
<thead>
<tr>
<th>DATE OF SERVICE</th>
<th>DESCRIPTION</th>
<th>QTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/06</td>
<td>Chronic Ischemic Heart Disease</td>
<td></td>
</tr>
<tr>
<td>10/06</td>
<td>Intermediate Office Visit</td>
<td>001</td>
</tr>
<tr>
<td>10/06</td>
<td>Complete Blood Count</td>
<td>001</td>
</tr>
<tr>
<td>10/06</td>
<td>Glucose or Blood Sugars</td>
<td>001</td>
</tr>
<tr>
<td>10/06</td>
<td>Urinalysis</td>
<td>001</td>
</tr>
<tr>
<td>10/06</td>
<td>Electrocardiogram</td>
<td>001</td>
</tr>
<tr>
<td>11/16</td>
<td>Chronic Ischemic Heart Disease</td>
<td></td>
</tr>
<tr>
<td>11/16</td>
<td>Intermediate Office Visit</td>
<td>001</td>
</tr>
<tr>
<td>11/16</td>
<td>Urinalysis</td>
<td>001</td>
</tr>
<tr>
<td>11/16</td>
<td>Glucose or Blood Sugars</td>
<td>001</td>
</tr>
<tr>
<td>11/16</td>
<td>Multichannel Test</td>
<td>001</td>
</tr>
<tr>
<td>11/16</td>
<td>Chest X-Ray</td>
<td>001</td>
</tr>
</tbody>
</table>

**EVALUATION OF CARE:**

- Appears Appropriate
- Suggests Inappropriate Care

Comments (Specify if care was inappropriate, or review inconclusive)

The final pages of the profile contain a listing of all beneficiaries included in the profile (to be sent to the subject physician if the panel recommends that he or she be contacted). These pages also contain statistical information on the physician's practice, as shown in Figure 3.

Using this strategy in recent reviews, we have been able to secure more definitive assessments of the quality of care delivered by these physicians to Funds beneficiaries. Before the profiles were enhanced in this way, panel recommendations were restricted to decisions on whether or not to contact the physician involved to secure additional information to conduct a more thorough review. In several cases in recent meetings where we have prepared this supplementary information prior to the peer review, the nature of the medications and procedures indicated by the physician on the claim has suggested a need for more definitive intervention without the need to obtain additional information from the physician.

**RESULTS OF PEER REVIEW**

**IN RELATION TO SELECTION CRITERIA**

For our most recent reviews, we selected 90 percent as a cut-off point for each of our statistical
measures. Thus, physicians were identified for review if they were in the upper 10 percent on any measure for their specialty and if the difference between the physician's value and the Funds' median was significant using a t-test to compare the physician's practice to other physicians in the same specialty. This statistical approach was selected to ensure that the high observed value was the result of a persistent tendency to utilize a service in treating many patients and not the result of a few unusual cases in the physician's practice. The statistical test considers not only the number of beneficiaries seen by the physician in a calendar quarter but also considers the variability within each physician's practice profile.

Our comparison of peer review findings with initial selection values for the injection measure suggests that this 90 percent cut-off score may be too high. Virtually all physicians identified by this measure were judged to be practicing inappropriately, and many physicians selected for other measures were also questioned about their injection usage. In fact, our review panels felt that very few drugs administered by injection were indicated in ambulatory practice.

For diagnostic tests, the 90 percent criterion appeared to be more appropriate. However, when the usage of panel tests is considered in the computation of this variable, the measure will be even more effective. It is interesting to note that no physicians identified by other measures were cited as ordering too many diagnostic tests; however, in several cases, physicians identified by other measures were questioned by the review panels about their low usage of diagnostic tests.

For visits, the situation is more ambiguous. Distinguishing inappropriate visits is a more subtle medical issue and the type of profiles prepared for our most recent reviews may have contributed to some of the problems involved in making such judgments. The profiles we have presented to review panels in the past have included six months of care for approximately twenty beneficiaries. This time period leaves open the possibility that some of the physician's patients might be new patients. (This question also applies to the use of diagnostic tests.) In the future, we will be experimenting with profiles showing a longer period of care for fewer patients. We would expect that this would resolve some of the difficulties the review panels have had with respect to assessing appropriate visit levels.

**CONCLUSIONS**

In comparison to alternate approaches to the review of ambulatory care, this system is more cost-effective in two ways. First, it relies on a limited number of validated selection criteria, minimizing the amount of analyst time required to accomplish the selection of physicians for review. Second, initial reviews are conducted based on claims history data rather than on more detailed medical information which can be difficult to obtain and time-consuming to review. Thus, with only a marginal expenditure above what is required for the processing of claims, we have been able to develop and implement an approach to utilization review that has won the acceptance of a large number of well respected physicians. These physicians have not only agreed to participate in the review process by serving on review panels, but have consistently been willing to make recommendations for interventions in the majority of the cases they have reviewed.

The effectiveness of our approach to identifying physicians for review has been supported by peer review of more than sixty physicians. At present, we can only report generally on the results of contacting physicians recommended for interventions. The responses of these physicians have, in most cases, tended to confirm the initial peer review findings. Most have been cooperative in responding to our inquiries and supportive of our goals.

In conducting these reviews and contacting individual physicians for further information, a number of general practice problems have been identified that apply to the larger group of physicians serving Funds beneficiaries as well as to the physicians selected for review through the UR system. Thus, we anticipate that this review process will have an impact not only on the physicians reviewed, but on the overall quality of medical care delivered to Funds beneficiaries.