Overview of Predictive Modeling Tools for Medicaid Populations

David Knutson
Division of Health Policy and Management
University of Minnesota

Medicaid Best Buys 2008: Using Predictive Modeling to Pinpoint “High-Opportunity” Medicaid Beneficiaries

Center for Health Care Strategies
Webinar
August 13, 2008
Overview Outline

1. Introduction to predictive modeling
2. Variables used in predictive modeling
3. Criteria for choosing a predictive modeling approach
Predictive Modeling for Case Identification

- Predictive models (PM) combine risk factors to predict an individual’s future need for care, e.g., care management

- PM are decision support tools that provide one information source, among others, for identifying cases that need special services or programs

- PM can also be used for program planning and evaluation
Two types of prediction:
1) Predicting an individual’s future medical expenditures
2) Identifying individuals in a population who can benefit from a disease or care management intervention

All PM software in the market classifies individuals by future cost categories, with the focus being on high-cost individuals

Some PM tools add care gap logic (e.g., HEDIS, ACS admissions) to identify cases that can most benefit from care management interventions or produce savings
Typically use regression or data mining methods to develop multi-variate models.

Predictive accuracy for case identification is usually defined as the:

1) Proportion of individuals who were identified by PM, but were not actually a good fit for the intervention, i.e., false positives (specificity); and

2) Proportion of individuals in the population who were actually a good fit, but were not identified by PM (sensitivity).
Most PM tools rely on claims and enrollment data and include these variables as predictors:

- Age/Gender (always)
- Diagnosis codes (always)
- Prior cost (often)
- Utilization (often)
- Prescriptions claims as proxies for diagnoses (often)
- Care gap logic combines diagnoses with utilization (often)

Some can use additional data: HRA/Functional status (rarely)
# Assessing Data Variables for Predictive Models*

*Based on experience.

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>PROS</th>
<th>CONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age/Gender</td>
<td>• Easy to obtain&lt;br&gt;• Reliable and valid</td>
<td>• Poor predictive value, but useful when combined with other variables</td>
</tr>
<tr>
<td>Prescription Claims</td>
<td>• Predicts costs almost as well as diagnoses (diagnosis proxy)&lt;br&gt;• Short claims lag</td>
<td>• Perverse incentive if used for provider or plan payment&lt;br&gt;• Requires frequent model updates</td>
</tr>
<tr>
<td>Diagnoses (claims)</td>
<td>• Best predictor of cost&lt;br&gt;• Required for care gap analysis</td>
<td>• Ambulatory coding inconsistent&lt;br&gt;• Long claims lag</td>
</tr>
<tr>
<td>Functional Status*</td>
<td>• Useful for care gap analysis and intervention planning&lt;br&gt;• Can be collected upon enrollment</td>
<td>• Only modest predictive performance when used alone&lt;br&gt;• Not universally or reliably assessed.&lt;br&gt;• Adds little predictive performance combined with diagnoses</td>
</tr>
<tr>
<td>Utilization Data</td>
<td>• Important for identifying care gaps and potential cost savings&lt;br&gt;• Moderate improvement in predictive accuracy combined with diagnoses</td>
<td>• Perverse incentives if used for provider or plan payment</td>
</tr>
<tr>
<td>Prior Cost</td>
<td>• Good stand alone predictor of future cost&lt;br&gt;• Adds modest improvement in prediction when combined with diagnoses</td>
<td>• Not as good a predictor as diagnoses&lt;br&gt;• Adds little clinical information</td>
</tr>
</tbody>
</table>

*Based on experience.
The Medicaid Difference

- PM tools developed primarily for commercial health plan populations with a low prevalence of complexity and less intense interventions (e.g., screening reminders)

- Medicaid populations: More disability, serious mental health comorbidity, and social complexity

- PM tools for Medicaid not fully developed as yet to account for this complexity of need and intensity of care management
Criteria for choosing a predictive modeling approach: *Begin with the end in mind*

- How is the state planning to use PM?
- What is the target population/cohort that will be identified using PM?
- What information and program benefits are expected from PM?
- What data are needed and available?
- Has the state assembled an implementation workgroup with the required skills?
Summary

• PM is an important decision support tool to assist Medicaid programs in case identification and can also be used for program evaluation.

• Medicaid populations are more complex and the interventions more intense than the settings in which these tools have historically been applied.

• Medicaid programs must be clear when they design or buy a tool on how they plan to use it, the objectives they have for the tool, and how far the tool will get them without local adaptation.