Screening for Impaired Fasting Glucose or Diabetes in Managed Care

Laura McEwen, PhD
and Bill Herman, MD, MPH
University of Michigan
Ann Arbor, Michigan
Introduction

• Persons with impaired fasting glucose (IFG), defined as fasting glucose between 110 and 125 mg/dL, are at a greater risk for diabetes and cardiovascular disease

• Studies have shown that both lifestyle and medication interventions can delay or prevent the development of diabetes
Introduction

• Translation of proven-effective diabetes prevention strategies has not occurred in routine clinical practice

• One potential barrier has been the difficulty in identifying patients with IFG

• Applying mathematical models to existing managed care data may be a practical and inexpensive way to identify patients with IFG
Specific Aim

• To develop a set of mathematical models that can be applied by managed care health plans to routinely collected patient data to predict a patient’s risk of having IFG or diabetes
Population

• 700,000+ member managed care organization in southeast Michigan

• Eligibility
  – At least 18 years old
  – Not pregnant
  – No history of diabetes (i.e. not on the health plan’s diabetes registry)
  – Members of “Healthy Blue Living” (HBL)
  – Continuously enrolled in the health plan for 12 consecutive months either before or after their HBL enrollment
“Healthy Blue Living” (HBL)

• A new and relatively small (70,000 member) commercial health insurance product
• Requires annual assessment of height, weight, blood pressure, lipids, and fasting glucose by a primary care physician and a commitment to develop an action plan
• Responses to the qualifying form are recorded in a database maintained by the health plan
• Employers get a 10% reduction in premiums and patients get increased benefits and reduced out-of-pocket costs
Health Plan Data

- **Demographic data from HBL enrollment forms**
  - Age
  - Sex
  - Race

- **Administrative data (flagged if at least 1 claim in time period)**
  - History of gestational diabetes
  - History of polycystic ovarian syndrome
  - Obesity
  - Hypertension
  - Dyslipidemia
  - Cardiovascular disease
  - Antihypertensive medication
  - Lipid lowering medication

- **Lab values from laboratory database from date closest to HBL enrollment**
  - High density lipoprotein
  - Low density lipoprotein
  - Total cholesterol
  - Triglycerides

- **Clinical data from HBL enrollment form**
  - Body mass index
  - Systolic and diastolic blood pressure
  - High density lipoprotein
  - Low density lipoprotein
  - Total cholesterol
  - Fasting blood glucose
Methods

- Study population N=27,941
- Randomly divided dataset into two equal parts, a development set and a validation set
- Dichotomized the outcome as undiagnosed IFG or diabetes based on a FBG recorded from the HBL enrollment form $\geq 110$ mg/dL
- Developed 3 models using available data to account for different levels of data available to the health plan
  - Only data available through administrative claims
  - Data available through administrative claims and lab values obtained from health plan laboratory database
  - Data available through administrative claims, lab values, and the HBL enrollment form
Methods

• To obtain the most parsimonious models from the panels of risk factors, we estimated the number of variables in the model that would result in the minimum Schwarz information criterion (SC) and Akaike information criterion (AIC)

• Once the optimal number of variables was determined, a combination of approaches including stepwise regression, regression choosing the best model based on the score statistic given a set number of variables, and clinical experience guided our development of the final three models

Shtatland ES, Cain E, Barton MB. The perils of stepwise logistic regression and how to escape them using information criteria and the output delivery system. 26th Annual SAS Users Group International Conference, Long Beach, CA, 2001
## Characteristics of the Study Population

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total Population N=27941</th>
<th>No IFG N=26545</th>
<th>IFG N=1396</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>45 10</td>
<td>45 10</td>
<td>50 10</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Sex (% female)</td>
<td>54</td>
<td>55</td>
<td>39</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Race (% white)</td>
<td>85</td>
<td>85</td>
<td>86</td>
<td>0.1616</td>
</tr>
<tr>
<td>Smoking status (% nonsmoker)</td>
<td>94</td>
<td>94</td>
<td>93</td>
<td>0.0349</td>
</tr>
<tr>
<td>History of GDM (females only) (%)</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>0.8344</td>
</tr>
<tr>
<td>History of PCOS (females only) (%)</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0.2767</td>
</tr>
<tr>
<td>Obesity (%)</td>
<td>10</td>
<td>10</td>
<td>18</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Hypertension (%)</td>
<td>31</td>
<td>30</td>
<td>54</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Dyslipidemia (%)</td>
<td>60</td>
<td>59</td>
<td>75</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Cardiovascular disease (%)</td>
<td>18</td>
<td>17</td>
<td>25</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Antihypertensive medication (%)</td>
<td>18</td>
<td>17</td>
<td>34</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Lipid lowering medication (%)</td>
<td>20</td>
<td>19</td>
<td>38</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Data are mean std deviation or percentage as appropriate.
Characteristics of the Study Population

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Total Population  N=27941</th>
<th>No IFG  N=26545</th>
<th>IFG  N=1396</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI</td>
<td>29.6</td>
<td>28.6</td>
<td>32.7</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>120.13</td>
<td>119.13</td>
<td>126.14</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>76.9</td>
<td>75.9</td>
<td>79.9</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>HDL (from lab or HBL form)</td>
<td>51.15</td>
<td>51.15</td>
<td>46.14</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>LDL (from lab or HBL form)</td>
<td>115.30</td>
<td>115.30</td>
<td>118.33</td>
<td>0.0078</td>
</tr>
<tr>
<td>Total cholesterol (from lab or HBL from)</td>
<td>190.35</td>
<td>190.35</td>
<td>194.38</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Triglycerides</td>
<td>120.35</td>
<td>118.88</td>
<td>160.104</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Data are mean ± std deviation or percentage as appropriate
Prevalence of IFG or DM by age

5 Year Age Groups

Prevalence of IFG or DM

5 Year Age Groups

LT 25
25- 29
30- 34
35- 39
40- 44
45- 49
50- 54
55- 59
60- 64
65- 69
70- 74
GE 75

14
12
10
8
6
4
2
0
Models 1, 2, and 3

• Model 1: administrative data (N=9445 total, 566 (6%) with IFG)
  – Age
  – Sex
  – Obesity
  – Hypertension
  – Lipid lowering medication

• Model 2: administrative claims, and lab data (N=3229 total, 236 (8%) with IFG)
  – Age
  – Sex
  – Lipid lowering medication
  – HDL from lab data

• Model 3: administrative claims, lab data, and patient encounter (N=6424 total, 389 (6%) with IFG)
  – Age
  – Sex
  – Hypertension
  – BMI
  – Systolic blood pressure
  – HDL cholesterol from lab data or HBL enrollment form
  – Triglycerides
## Model 1 – administrative data only

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimated regression coefficient</th>
<th>Estimated standard error</th>
<th>p-value</th>
<th>Estimated OR</th>
<th>95% CI for OR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>-5.0069</td>
<td>0.3233</td>
<td>&lt;0.001</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Age (years)</td>
<td>0.0403</td>
<td>0.00627</td>
<td>&lt;0.001</td>
<td>1.041</td>
<td>1.028 – 1.054</td>
</tr>
<tr>
<td>Female</td>
<td>-0.6163</td>
<td>0.0920</td>
<td>&lt;0.001</td>
<td>0.540</td>
<td>0.451 – 0.647</td>
</tr>
<tr>
<td>Obesity</td>
<td>0.6024</td>
<td>0.1164</td>
<td>&lt;0.001</td>
<td>1.827</td>
<td>1.454 – 2.295</td>
</tr>
<tr>
<td>Hypertension</td>
<td>0.5300</td>
<td>0.0962</td>
<td>&lt;0.001</td>
<td>1.669</td>
<td>1.407 – 2.051</td>
</tr>
<tr>
<td>Lipid lowering medication</td>
<td>0.2894</td>
<td>0.0956</td>
<td>0.0025</td>
<td>1.336</td>
<td>1.107 – 1.611</td>
</tr>
</tbody>
</table>

\[ R^2 = 0.0681 \]

Area under the curve = 0.689

Hosmer and Lemeshow Goodness-of-fit p=0.0621
To produce ROC curves for each model:
Optimal cutpoint

- Chosen where sum of sensitivity and specificity is maximized
- At probability of 0.050
  - Sensitivity of 73%
  - Specificity of 56%
  - Positive predictive value (PPV) of 9%
- Our population has a low prevalence of the outcome (6%), if everything stayed the same, except the prevalence increased, our PPV would increase accordingly
  - Prevalence = 15%  PPV = 30%
  - Prevalence = 30%  PPV = 53%
## Verification Population

### Probability = 0.050

<table>
<thead>
<tr>
<th></th>
<th>True D</th>
<th>True non D</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +</td>
<td>432</td>
<td>3961</td>
<td>4393</td>
</tr>
<tr>
<td>Test -</td>
<td>185</td>
<td>4848</td>
<td>5033</td>
</tr>
<tr>
<td>Total</td>
<td>617</td>
<td>8809</td>
<td>9426</td>
</tr>
</tbody>
</table>

- **Sensitivity** = 70%
- **Specificity** = 55%
- **PPV** = 10%

### Probability = 0.100

<table>
<thead>
<tr>
<th></th>
<th>True D</th>
<th>True non D</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test +</td>
<td>172</td>
<td>1058</td>
<td>1230</td>
</tr>
<tr>
<td>Test -</td>
<td>445</td>
<td>7751</td>
<td>8196</td>
</tr>
<tr>
<td>Total</td>
<td>617</td>
<td>8809</td>
<td>9426</td>
</tr>
</tbody>
</table>

- **Sensitivity** = 30%
- **Specificity** = 88%
- **PPV** = 14%
Conclusion

• These multivariate logistic regression equations can be applied in managed care plans using different levels of available data (administrative claims, laboratory, and pharmacy) to identify persons at high risk for IFG or diabetes who might benefit from interventions.

• Since the PPV is only 10-14%, health plans would face a high number of false positives and need to follow up.

• Whether or not it is cost-effective to use these equations to identify high-risk groups would depend on the prevalence of IFG and diabetes in the health plan’s population.