Development and implementation of a predictive model for preemptive pharmacogenetic testing

Jonathan S. Schildcrout
Department of Biostatistics
Vanderbilt University

In collaboration with:
Yaping Shi, Erica Bowton, Jill Pulley, Melissa Basford, William Gregg,
Julie Field, Ioana Danciu, James Cowan, Frank Harrell, Dan Roden,
Josh Peterson, Joshua Denny

Seattle Symposium on Healthcare Data Analytics
September 29, 2014
Outline

Introduction and motivation

Model construction and training cohort analysis

Model portability: Implementation into a CDS tool and beyond

Model performance on validation cohorts

Summary, limitations and next steps
Motivation

- A growing body of literature/knowledge relates genetic variation to medication responsiveness
  - >100 drugs with pharmacogenetic (PGx) information on FDA labels
- National interest in translating research findings into medical practice.
- PREDICT: Pharmacogenomic Resource for Enhanced Decisions in Care & Treatment (PREDICT) program at VUMC
  - A QI initiative to use preemptive, panel-based genotyping to deliver genotype-tailored prescribing guidance at the point of care
Motivation (cont.)

- Panel-based, multiplexed genotyping might be effective because...
  - Many people are exposed to meds with PGx effects: In 53000 'medical home' patients at Vanderbilt, over a 5 year timeframe
    - 65% prescribed $\geq 1$ med with FDA label.
    - 40% prescribed $\geq 2$ meds with FDA label.
    - 23% prescribed $\geq 3$ meds with FDA label.
    - At the time of that analysis, 57 such labels existed.
  - A large percentage of people also have variant alleles that confer excess risk in at least some medication.
Motivation (cont)

- PREDICT: have genetic data in the EHR prior to prescribing.
- Our task was to identify a subset of patients at high risk for a PGx script where alternative treatment options were thought to be available (required for genomic information to be useful)
Today

- Describe an effort to identify high risk patients for a statin, clopidogrel, or warfarin script, who are naive to the medications at baseline.
  - **Construction**: Built a model to predict who is going to be prescribed a PG× med (2005 - 2010).
  - **Implementation**: Considerations for a clinical decision support (CDS) tool that alerts physicians if a patient is at high risk.
    - Other work will examine whether people followed the recommendation.
  - **Validation**: Examine model performance between 2010 and 2013.
    - My goal: to get the most bang for the buck (i.e., enrich the genotyped pool with those eventually prescribed a PG× med).
The 2005-2010 Training (T) cohort

- We use patient data from Vanderbilt’s ’synthetic derivative’
  - A de-identified version of the EMR
- Inclusion criteria:
  - Vanderbilt was identified as their ’medical home’ (MH) between 2005 and 2010
    - ≥ 3 outpatient visits within 2 years in internal medicine (primary care), cancer, hematology, hypertension, rheumatology, nephrology, cardiology, diabetes, neurology, nutrition or pulmonary clinics (24 clinics).
    - Age, height, and weight available prior to the MH date
    - Naive to statin, warfarin, and clopidogrel as of the MH date
- Patients followed through June 30, 2010.
Model construction

- Predict (first of) clopidogrel, warfarin, or statin prescription
- Considerations and constraints:
  - Data must be readily available in the EMR (no obscure labs)
  - CDS has to be run on a daily basis without having to maintain it
    - Limits the number of variables to be included
  - Models should be interpretable and portable
- Independent variables.
  - Age, gender, BMI, race
  - ICD-9 and medication history identified medical history variables
    - Type II diabetes, coronary artery disease, atrial fibrillation, hypertension, atherosclerosis, congestive heart failure, previous blood clot, and dialysis
Modeling

- Variable follow-up times → Use Cox model for the time to prescription of Warfarin, Statin, or Clopidogrel from the MH date

\[
\lambda(t; X) = \lambda(t; 0) \cdot \exp(X\beta)
\]

modeled on the log scale

\[
\log\{\lambda(t; X)\} = \log\{\lambda(t; 0)\} + X\beta
\]

letting \( T \) be the time to the prescription, then

\[
S(t; X) = \text{pr}(T > t \mid X) = S(t; 0)^{\exp(X\beta)}
\]

is the probability of being medication free at time \( t \) for those with covariates \( X \) where

\( S(t; 0) = \exp\{-\int_0^t \lambda(u; 0)\,du\} \) is the survivor function for those with all \( X \) set to reference values (i.e., the baseline survivor function)
For $t = 1095$, we provided the implementation team estimates of:

- Risk: $R(t; \mathbf{X}) = 1 - S(t; \mathbf{X})$
- 95\% CI: $1 - \exp\{\log S(t; \mathbf{X}) \pm z_{0.975} \times SE[\log S(t; \mathbf{X})]\}$

Decision was made to recommend genotyping if $\hat{R}(1095; \mathbf{X}) > 0.4$.

Based on a discussion that considered the cost of genotyping and the fraction of people leaving vascular clinics on Clopidogrel.
Demographics and Baseline Characteristics of the T cohort

<table>
<thead>
<tr>
<th>Variable</th>
<th>Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>16020</td>
</tr>
<tr>
<td>Age (years)</td>
<td>51 [29, 70]</td>
</tr>
<tr>
<td>Male</td>
<td>37.6</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>13.8</td>
</tr>
<tr>
<td>Other</td>
<td>3.1</td>
</tr>
<tr>
<td>White</td>
<td>83.2</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>28 [21, 39]</td>
</tr>
<tr>
<td>Follow-up (days)</td>
<td>1182 [148, 1720]</td>
</tr>
</tbody>
</table>
## Demographics and Baseline Characteristics of the T cohort

<table>
<thead>
<tr>
<th>Variable</th>
<th>Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medical Hx</td>
<td></td>
</tr>
<tr>
<td>Type II Diabetes</td>
<td>17.8</td>
</tr>
<tr>
<td>CAD</td>
<td>4.3</td>
</tr>
<tr>
<td>Atrial Fibrillation</td>
<td>0.1</td>
</tr>
<tr>
<td>Hypertension</td>
<td>32.7</td>
</tr>
<tr>
<td>Atherosclerosis</td>
<td>8.1</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
<td>3.4</td>
</tr>
<tr>
<td>Previous Clot</td>
<td>1.0</td>
</tr>
<tr>
<td>Dialysis</td>
<td>0.7</td>
</tr>
<tr>
<td>Statin</td>
<td>19.4</td>
</tr>
<tr>
<td>Warfarin</td>
<td>5.0</td>
</tr>
<tr>
<td>Clopidogrel</td>
<td>2.8</td>
</tr>
<tr>
<td>Any (S, W, C) Med</td>
<td>23.4</td>
</tr>
</tbody>
</table>
T cohort: Medication-free survival

Training Data

Model for pre-emptive pharmacogenetic testing
T cohort: Effect estimates for the Cox model

Model construction and training cohort analysis
Bootstrap-based estimates of calibration (AUROC ≈ 0.67)
Model Portability

- At participating clinics, each time a patient came into the office, we need risk and uncertainty estimates pre-calculated to recommend genotyping (if necessary).
- Results needed to be produced almost in real time.
- Part of this was easy.
  - $\hat{S}(1095; 0)$ was provided directly to the implementation team
  - The estimated linear predictor, $\mathbf{X}\hat{\beta}$ is

```r
> Function(CoxModel)
function(age = 48,bmi = 27.750891,male = 0,race3 = "White",t2d = 0,cad = afib.icd = 0,htn = 0,ath = 0,chf = 0,clot = 0,dial = 0) {
-4.7305438+0.065459713*age-3.0194841e-05*pmax(age-24,0)^3+0.0001037356*pmax(age-41,0)^3-0.00015032806*pmax(age-51,0)^3+0.0001060992*pmax(age-60,0)^3-2.9311903e-05*pmax(age-76,0)^3+0.065024817*bmi-0.00010665707*pmax(bmi-20.07749,0)^3+0.00012222854*pmax(bmi-24.578615,0)^3-0.00010824019*pmax(bmi-28.051752,0)^3+0.00016694303*pmax(bmi-32.435083,0)^3-7.4274303e-05*pmax(bmi-43.639397,0)^3+0.23279775*female+0.0860601*(race3="Other")+0.049906138*(race3="White")+0.58213821*t2d+0.21707034*cad+0.4003414*afib.icd-0.0013109944*htn+0.036763009*ath+0.19367406*chf+0.11672828*clot+0.67819823*dial }
```
Model Portability (cont.)

- Uncertainty estimate are complicated
- \( \hat{SE}[\log\{\hat{S}(t; X)\}] \): difficult to calculate in real time and w/o original data due to parametric and non-parametric elements of the Cox model
- We built a model for standard errors
  - Using the fitted Cox model, calculate \( \log(\hat{SE}) \) for all subjects
  - OLS regression of \( \log(\hat{SE}) \) on \( X \) to obtain the model for \( \log(\hat{SE}) \).

- Examine internal validity / accuracy with the bootstrap.
- For each bootstrap sample, b:
  1. Fit the Cox model and calculate \( \log(\hat{SE})^b \)
  2. Using OLS, fit \( \log(\hat{SE})^b \sim X^b \)
  3. Apply this model to the original dataset and compare the bootstrap based fitted values to each subjects \( \log(\hat{SE}) \)
SE calibration plots

Standard Error Calibration

Original dataset: $\log(\text{SE}(\log(\hat{S}(t, X))))$

Bootstrap-based: $\log(\text{SE}(\log(\hat{S}(t, X))))$

$R^2 = 0.991$
Model Portability (cont)

- Uncertainty estimates (i.e. CIs) can now be included into CDS

```r
> Function(LogSELogSurvMod)
function(age = 51,bmi = 28.051752,male = 0,race3 = "White",t2d = 0,cad = 0,
afib.icd = 0, htn = 0,ath = 0, chf = 0, clot = 0, dial = 0) {
-5.8958463+0.022684794*age+7.7786132e-06*pmax(age-24,0)^3-
4.1132366e-05*pmax(age-41,0)^3+4.9258159e-05*pmax(age-51,0)^3-
1.2269317e-05*pmax(age-60,0)^3-3.6350898e-06*pmax(age-76,0)^3-
0.020381286*bmi+0.00094882724*pmax(bmi-20.07749,0)^3-
0.0029280907*pmax(bmi-24.578615,0)^3+0.0026136747*pmax(bmi-28.051752,0)^3-
0.0065024166*pmax(bmi-32.435083,0)^3+1.583047e-05*pmax(bmi-43.639397,0)^3+
0.24621685*male+0.52703602*(race3=="Other")-0.10316402*(race3=="White")+
0.68012785*t2d+0.15662518*cad+2.273217*afib.icd+0.060475234*htn+
0.39928921*ath+0.50325011*chf+0.85058168*clot+1.4264086*dial }
```
Validation and Implementation cohorts (2010-2013)

- Validation (V) and Implementation (I) cohorts
  - MH date had to have occurred July 2010 - March 2013.
  - Did not include those in the T cohort.
- V cohort
  - Same inclusion as T cohort (same clinics)
- I cohort: The cohort in whom the model was deployed
  - Patients sought care at: internal medicine (primary care), cardiology, hypertension, diabetes, anticoagulation, ophthalmology, nephrology, renal transplant or urology clinics (120 clinics)
Validation and Implementation cohorts (2010-2013)

- Four datasets in which to validate and evaluate model performance
  1. Validation: Baseline and longitudinal data
  2. Implementation: Baseline and longitudinal data
- For longitudinal data use a derived time scale:
  - Each time the a patient comes to a clinic, it is a ‘new’ opportunity to recommend genotyping.
  - Subject $i$: observed at $t = \{0, s_1, s_2\}$ prior to her event time $T_i$.
  - The derived time scale use for predictions is $\{T_i, T_i - s_1, T_i - s_2\}$ with covariates $\{X_{i0}, X_{is1}, X_{is2}\}$.
### Demographics and Baseline Characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Training</th>
<th>Validation</th>
<th>Implementation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
<td>16020</td>
<td>12794 [55344]</td>
<td>18950 [61847]</td>
</tr>
<tr>
<td>Age (years)</td>
<td>51 [29, 70]</td>
<td>48 [26, 68]</td>
<td>46 [26, 69]</td>
</tr>
<tr>
<td>Male</td>
<td>37.6</td>
<td>38.5</td>
<td>36.9</td>
</tr>
<tr>
<td>Black</td>
<td>13.8</td>
<td>11</td>
<td>10.5</td>
</tr>
<tr>
<td>Other</td>
<td>3.1</td>
<td>4.5</td>
<td>4.8</td>
</tr>
<tr>
<td>White</td>
<td>83.2</td>
<td>84.5</td>
<td>84.7</td>
</tr>
<tr>
<td>BMI kg/m2</td>
<td>28 [21, 39]</td>
<td>27 [21, 38]</td>
<td>27 [21, 38]</td>
</tr>
<tr>
<td>Follow-up (days)</td>
<td>1182 [148, 1720]</td>
<td>361 [47, 812]</td>
<td>316 [23, 774]</td>
</tr>
<tr>
<td>Type II Diabetes</td>
<td>17.8</td>
<td>10.6 [12.8]</td>
<td>8.1 [9.8]</td>
</tr>
<tr>
<td>CAD</td>
<td>4.3</td>
<td>1.3 [2.0]</td>
<td>2.9 [3.9]</td>
</tr>
<tr>
<td>Atrial Fibrillation</td>
<td>0.1</td>
<td>1.4 [2.2]</td>
<td>3.6 [4.5]</td>
</tr>
<tr>
<td>Hypertension</td>
<td>32.7</td>
<td>26.2 [31.9]</td>
<td>32.1 [37.3]</td>
</tr>
<tr>
<td>Atherosclerosis</td>
<td>8.1</td>
<td>2.9 [4.2]</td>
<td>4.5 [6.1]</td>
</tr>
<tr>
<td>Congestive Heart Failure</td>
<td>3.4</td>
<td>2.0 [2.5]</td>
<td>2.4 [3.1]</td>
</tr>
<tr>
<td>Previous Clot</td>
<td>1.0</td>
<td>1.2 [2.2]</td>
<td>0.8 [1.2]</td>
</tr>
<tr>
<td>Dialysis</td>
<td>0.7</td>
<td>0.5 [0.7]</td>
<td>1.3 [1.5]</td>
</tr>
<tr>
<td>Statin</td>
<td>19.4</td>
<td>6.5</td>
<td>8.5</td>
</tr>
<tr>
<td>Warfarin</td>
<td>5.0</td>
<td>3.6</td>
<td>3.7</td>
</tr>
<tr>
<td>Clopidogrel</td>
<td>2.8</td>
<td>2.0</td>
<td>2.4</td>
</tr>
<tr>
<td>Any Med</td>
<td>23.4</td>
<td>9.6</td>
<td>11.5</td>
</tr>
</tbody>
</table>

Model for pre-emptive pharmacogenetic testing
Medication-free survival

Baseline Data

Longitudinal Data

Model for pre-emptive pharmacogenetic testing
Calibration Plots

Baseline data: One-year risk

Longitudinal data: One-year risk

Model for pre-emptive pharmacogenetic testing
AUROC over time

Model performance on validation cohorts

Baseline data

Longitudinal data

Days since medical home

Days since last visit

Model for pre-emptive pharmacogenetic testing
Enrichment analysis

- Overall goal: considering limited resources, create a genotyped cohort enriched with those who need it.
Enrichment (over time) using baseline data

Model performance on validation cohorts

Implementation Cohort: Baseline data
(N=1343 genotyped)

Validation Cohort: Baseline data
(N=771 genotyped)

Cumulative incidence of PGx med script

Days since medical home

Model-based
High BMI and age>50
Randomly sampled

Model for pre-emptive pharmacogenetic testing
Model performance on validation cohorts

Enrichment (over time) using longitudinal data

Implementation Cohort: Longitudinal data
(N=1673 genotyped)

Validation Cohort: Longitudinal data
(N=1047 genotyped)
Summary and limitations

- Discussed a real-world CDS implementation of a model to identify those who should participate in a pre-emptive genotyping program.
- The model:
  - preserved risk rankings and V and I cohorts
  - underestimated risk slightly (by a lot!) in the V (I) cohort
  - enriched the pool of genotyped people with those who may well benefit from it.
Limitations

• Generalizability
  ▶ Incomplete data in the EHR (med hx, etc)
  ▶ To other institutions / clinics?
  ▶ MH requirement: 3 visits/2 yrs and weight / height availability
  ▶ Patients had to be naive to the meds.

• Built one (statin dominated) model for three outcomes.

• Considerations for genotyping should be more thoughtful than I’ve made them. For each medication, consider:
  1. Cost associated with an event ’caused by” variant alleles.
  2. Risk decrease / increase of all AEs with alternative treatment.
  3. Variant allele prevalence among those on med
Steps to improving predictivity

- Separate outcomes when building models.
  - Will also allow us to make PREDICT more flexible and weigh each of the PGx medication script risks against one another.
- Include clinic type as an independent variable.
- Allow past exposure to one med to predict future exposure to another.
  - T cohort: Among those exposed to $\geq 1$ med, 15% exposed to $\geq 2$.
  - I cohort: Among those exposed to $\geq 1$ med, 22% exposed to $\geq 2$.
- Explore relatively sophisticated prediction techniques (random forests, etc.)