Regression modeling of longitudinal binary outcomes with outcome-dependent observation times

Benjamin French, PhD
Department of Biostatistics and Epidemiology
University of Pennsylvania
bcfrench@upenn.edu

Department of Biostatistics
Vanderbilt University
20 November 2013
Collaborators

Kay See Tan, MS  
University of Pennsylvania

Andrea Troxel, ScD  
University of Pennsylvania

Stephen Kimmel, MD, MSCE  
University of Pennsylvania

Kevin Volpp, MD, PhD  
University of Pennsylvania
• Anticoagulant used for prevention and treatment of thromboembolism
  ▶ Formation of a clot that obstructs blood flow in vein or artery
  ▶ Associated outcomes: Deep vein thrombosis, stroke

• Highly efficacious, but has a narrow therapeutic range
  ▶ Over-anticoagulation: Increased risk of bleeding complications
  ▶ Under-anticoagulation: Increased risk of thromboembolic events

• Requires frequent monitoring, which may lead to dose changes
  ▶ INR: International normalized ratio
  ▶ Ratio of patient’s prothrombin time to that of a normal sample

• Poor adherence contributes to poor anticoagulation control
  ▶ Missed doses → under-anticoagulation (Kimmel et al., 2007)
  ▶ Goal: Improve adherence to warfarin therapy (Kimmel et al., 2012)
INR monitoring

Below range    Within range    Above range

Months since enrollment

0 1 2 3 4 5 6

Patient

Below range    Within range    Above range

Months since enrollment

0 1 2 3 4 5 6

Patient

Longitudinal binary outcomes

Vanderbilt Biostatistics
Analysis issue

- Frequency and/or timing of data collection may depend on past outcome values, i.e. ‘informative’ observation times

- Selection bias: Participants with particular outcome values may have more observations; over-represented in the sample
- May result in biased estimates for exposure-outcome associations (Pepe and Anderson, 1994; French and Heagerty, 2009)
Warfarin trial

- Participants randomized to one of four interventions (A, B, C, D)
- Followed for a maximum of six months
- Protocol required an INR measured each month
- Primary outcome: Out-of-range INR (binary)

<table>
<thead>
<tr>
<th>Group</th>
<th>$P_0$</th>
<th>$P_{25}$</th>
<th>$P_{50}$</th>
<th>$P_{75}$</th>
<th>$P_{100}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>7</td>
<td>8</td>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>24</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>11</td>
<td>16</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td>10</td>
<td>14</td>
</tr>
</tbody>
</table>
Warfarin trial: Analysis plan

- Logistic regression model for log-odds of out-of-range INR
  \[
  \text{logit } P[Y_i(t) = 1] = \beta_0 + \beta_1[B] + \beta_2[C] + \beta_3[D]
  \]

- GEE with working independence (Liang and Zeger, 1986)
  - Adjusted for stratification factors
  - Included a linear term for study time

- How to account for different number of INRs for each participant, which may be ‘informative’ of outcome values?
  - Use only ‘monthly’ INRs, closest to scheduled follow-up visits
  - Use all available INRs in an ‘unweighted’ model
  - Use all available INRs in a ‘weighted’ model, in which participant-level weights are equal to the inverse of the number of INRs, i.e. $1/N_i$; cluster-weighted GEE (Williamson et al., 2003)
Warfarin trial: Analysis plan

- Logistic regression model for log-odds of out-of-range INR
  \[
  \text{logit } P[Y_i(t) = 1] = \beta_0 + \beta_1[B] + \beta_2[C] + \beta_3[D]
  \]

- GEE with working independence (Liang and Zeger, 1986)
  - Adjusted for stratification factors
  - Included a linear term for study time

- How to account for different number of INRs for each participant, which may be ‘informative’ of outcome values?
  - Use only ‘monthly’ INRs, closest to scheduled follow-up visits
  - Use all available INRs in an ‘unweighted’ model
  - Use all available INRs in a ‘weighted’ model, in which participant-level weights are equal to the inverse of the number of INRs, i.e. \(1/N_i\); cluster-weighted GEE (Williamson et al., 2003)
### Warfarin trial: Results

#### Odds ratios for out-of-range INR

<table>
<thead>
<tr>
<th>Group</th>
<th>Monthly</th>
<th>Unweighted</th>
<th>Weighted</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>1.14</td>
<td>0.95</td>
<td>0.98</td>
</tr>
<tr>
<td>C</td>
<td>0.76</td>
<td><strong>0.66</strong></td>
<td><strong>0.64</strong></td>
</tr>
<tr>
<td>D</td>
<td>0.90</td>
<td>0.78</td>
<td>0.77</td>
</tr>
</tbody>
</table>

- Clearly there are benefits to using all available data
  - Reduced bias
  - Increased efficiency
- Can we do better?
  - Consider joint models for longitudinal outcomes and observation times
Joint models

• Semi-parametric outcome model

\[
\text{logit } P[Y_i(t) = 1 \mid X_i(t)] = \mu(t) + \beta' X_i(t)
\] (1)

\begin{itemize}
  \item \(\mu(t)\): arbitrary function of time
  \item \(\beta\): parameters of interest
\end{itemize}

• Observation-time model

\[
E[dN_i(t) \mid Z_i(t)] = \exp\{\gamma' Z_i(t)\} d\Lambda(t)
\] (2)

\begin{itemize}
  \item \(d\Lambda(t)\): baseline intensity function
  \item \(\gamma\): intensity parameters
\end{itemize}

★ Dependence induced by shared covariates \(X_i(t) \subseteq Z_i(t)\)
★ Requires that observation-time model is correctly specified
Joint models

- Semi-parametric outcome model
  \[
  \text{logit } P[Y_i(t) = 1 | X_i(t)] = \mu(t) + \beta' X_i(t) \tag{1}
  \]
  - \(\mu(t)\): arbitrary function of time
  - \(\beta\): parameters of interest

- Observation-time model
  \[
  E[dN_i(t) | Z_i(t)] = \exp\{\gamma' Z_i(t)\}d\Lambda(t) \tag{2}
  \]
  - \(d\Lambda(t)\): baseline intensity function
  - \(\gamma\): intensity parameters

★ Dependence induced by shared covariates \(X_i(t) \subseteq Z_i(t)\)
★ Requires that observation-time model is correctly specified
Observation-time model (Lin et al., 2000)

\[ U_\gamma(\gamma) = \sum_{i=1}^{n} \int_{0}^{\tau} \{ Z_i(t) - \bar{Z}(t; \gamma) \} dN_i(t) \]  

(3)

with

\[ \bar{Z}(t; \gamma) = \frac{\sum_i 1[C_i > t] \exp\{\gamma' Z_i(t)\} Z_i(t)}{\sum_i 1[C_i > t] \exp\{\gamma' Z_i(t)\}} \]

- \( \gamma \) considered a nuisance w.r.t. parameters of interest \( \beta \)
- Correct observation-time model provides reliable estimates of \( \beta \) in the presence of outcome-dependent observation times
- Identifying an observation-time model could inform future predictions
Joint models: Estimation

For continuous outcomes it is straightforward to estimate $\mu(\cdot)$

$$M_i(t; \beta, \gamma) = \int_0^t Y_i(u)1[C_i > u]dN_i(u)$$

$$- \int_0^t \{\mu(u) + \beta'X_i(u)\}1[C_i > u] \exp{\gamma'X_i(u)}d\Lambda(u)$$

- $M_i(t; \beta, \gamma)$ is a mean-zero process
- $dM_i(t)$ and $\int_0^t X_i(u)dM_i(u)$ are also mean-zero processes
- Estimation is based a set of estimating equations of the form

$$\sum_{i=1}^n dM_i(t; \beta, A) = 0 \quad (4)$$

$$\sum_{i=1}^n \int_0^T X_i(t)dM_i(t; \beta, A) = 0 \quad (5)$$

- Solve for $\mu(\cdot)$ in (4) and plug into (5) to solve for $\beta$
Joint models: Estimation

Continuous outcome model (Bůžková and Lumley, 2009)

\[ U_{\beta}(\beta, \hat{\gamma}, \delta) = \sum_{i=1}^{n} \int_{0}^{T} \frac{1}{\omega_{i}(t; \hat{\gamma}, \delta)} [X_{i}(t) - \bar{X}(t; \delta)] \]
\[ \times \{ Y_{i}(t) - \bar{Y}^{*}(t; \delta) - \beta'[X_{i}(t) - \bar{X}(t; \delta)] \} dN_{i}(t) \]  

(6)

with

\[ \omega_{i}(t; \gamma, \delta) = \frac{\exp{\gamma'Z_{i}(t)}}{\exp{\delta'X_{i}(t)}} \]

\[ \bar{X}(t; \delta) = \frac{\sum_{i} 1[C_{i} > t] \exp{\delta'X_{i}(t)}X_{i}(t)}{\sum_{i} 1[C_{i} > t] \exp{\delta'X_{i}(t)}} \]

\[ \bar{Y}^{*}(t; \delta) = \frac{\sum_{i} 1[C_{i} > t] \exp{\delta'X_{i}(t)}Y_{i}^{*}(t)}{\sum_{i} 1[C_{i} > t] \exp{\delta'X_{i}(t)}} \]

* Observations are weighted by inverse intensity rate ratio (IIIRR)
Joint models: Estimation

For binary outcomes it is not straightforward to estimate $\mu(\cdot)$

$$M_i(t; \beta, \gamma) = \int_0^t Y_i(u)1[C_i > u]dN_i(u)$$

$$\quad - \int_0^t \left( \frac{\exp\{\mu(u) + \beta'X_i(u)\}}{1 + \exp\{\mu(u) + \beta'X_i(u)\}} \right) 1[C_i > u] \exp\{\gamma'X_i(u)\}d\Lambda(u)$$

- For continuous outcomes used a set of estimating equations

$$\sum_{i=1}^{n} dM_i(t; \beta, A) = 0 \quad (4)$$

$$\sum_{i=1}^{n} \int_0^\tau X_i(t)dM_i(t; \beta, A) = 0 \quad (5)$$

- For binary outcomes cannot directly solve for $\mu(\cdot)$ in (4) and plug into (5)
Joint models: Estimation

Alternatives for binary outcomes

1. Solve for $\mu(\cdot)$ and $\beta$ simultaneously using an iterative procedure

$$\sum_{i=1}^{n} dM_i(t; \beta, A) = 0 \quad (4)$$

$$\sum_{i=1}^{n} \int_{0}^{\tau} X_i(t) dM_i(t; \beta, A) = 0 \quad (5)$$

- $\mu(\cdot)$ is infinite dimensional so computational issues may arise
- Sparse data resulting from few subjects at unique observation times

2. Impose some (still flexible) structure on $\mu(\cdot)$ using B-splines

- Standard GEE estimating equations with B-splines and weights
- ‘Iterative’ estimating equations with B-splines and weights
Warfarin trial: Joint models

- Semi-parametric outcome model
  - Logistic regression model for log-odds of out-of-range INR
    \[
    \text{logit } P[Y_i(t) = 1] = \mu(t) + \beta' X_i(t)
    \]

- Observation-time models
  1. Outcome-model covariates, including treatment group
    \[
    \lambda_i(t) = \lambda(t) \exp\{\gamma' X_i(t)\}
    \]
  2. Outcome-model covariates and out-of-range INR at previous visit
    \[
    \lambda_i(t) = \lambda(t) \exp\{\gamma'_X X_i(t) + \gamma'_Y Y_i(t - 1)\}
    \]

used to calculate observation-level IIRR weights for outcome model
Warfarin trial: Joint models

- Semi-parametric outcome model
  - Logistic regression model for log-odds of out-of-range INR
    \[
    \text{logit } P[Y_i(t) = 1] = \mu(t) + \beta'X_i(t)
    \]

- Observation-time models
  1. Outcome-model covariates, including treatment group
    \[
    \lambda_i(t) = \lambda(t) \exp\{\gamma'X_i(t)\}
    \]
  2. Outcome-model covariates and out-of-range INR at previous visit
    \[
    \lambda_i(t) = \lambda(t) \exp\{\gamma'X_i(t) + \gamma_Y Y_i(t - 1)\}
    \]

  used to calculate observation-level IIRR weights for outcome model
### Warfarin trial: Results

<table>
<thead>
<tr>
<th>Group</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>0.032</td>
<td>0.026</td>
</tr>
<tr>
<td>C</td>
<td>-0.036</td>
<td>-0.014</td>
</tr>
<tr>
<td>D</td>
<td>0.004</td>
<td>0.022</td>
</tr>
<tr>
<td>Out-of-range</td>
<td>—</td>
<td>0.367</td>
</tr>
</tbody>
</table>

- No clear differences in visit intensity between randomized groups, but agrees with differences in number of INRs per participant
- Out-of-range INR associated with an increase in visit intensity
## Warfarin trial: Results

### IIRR weights for outcome model

<table>
<thead>
<tr>
<th></th>
<th>$P_0$</th>
<th>$P_{25}$</th>
<th>$P_{50}$</th>
<th>$P_{75}$</th>
<th>$P_{100}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>0.86</td>
<td>1.05</td>
<td>1.11</td>
<td>1.16</td>
<td>1.36</td>
</tr>
<tr>
<td>Model 2</td>
<td>0.55</td>
<td>0.70</td>
<td>0.95</td>
<td>1.00</td>
<td>1.16</td>
</tr>
</tbody>
</table>

- Modest variability in weights within models
- Spread of weights similar between models
## Warfarin trial: Results

<table>
<thead>
<tr>
<th>Group</th>
<th>Monthly</th>
<th>Unweighted</th>
<th>Weighted</th>
<th>Joint models</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GEE models</td>
<td></td>
<td></td>
<td>Model 1</td>
</tr>
<tr>
<td>B</td>
<td>1.14</td>
<td>0.95</td>
<td>0.98</td>
<td>0.93</td>
</tr>
<tr>
<td>C</td>
<td>0.76</td>
<td><strong>0.66</strong></td>
<td><strong>0.64</strong></td>
<td><strong>0.65</strong></td>
</tr>
<tr>
<td>D</td>
<td>0.90</td>
<td>0.78</td>
<td>0.77</td>
<td>0.78</td>
</tr>
</tbody>
</table>

- Similar results between models that use all available data
• Longitudinal data with outcome-dependent observation times motivate consideration of joint models that specify a dependence structure between outcome and observation-time processes

• Outcome-observation dependence may be induced by shared covariates in outcome and observation-time processes

• Considered a semi-parametric outcome model that did not specify the functional form for temporal trends
  ▶ Outcome model weighted by inverse intensity rate ratio
  ▶ Required correctly specified observation-time model
  ▶ Model verification and/or sensitivity analyses may be required

• Extended methods for continuous outcomes to binary outcomes, which required some flexible specification for temporal trends
Future directions

• Consider latent variables, in addition to covariates, to induce dependence between outcome and observation-time processes
• Consider clustering of study participants within physicians
• Evaluate impact of informative censoring on estimation and inference
• Exploit joint models to estimate prediction rules
  ▶ Patient perspective
    ★ Risk of future poor outcomes given outcome history
    ★ Optimal time to next visit given visit and outcome history
  ▶ Provider perspective
    ★ Evaluate patient prognosis, inform personalized treatment strategies
    ★ Quantify demand for care infrastructure, e.g., beds, staff
A Pharmacogenetic versus a Clinical Algorithm for Warfarin Dosing

Stephen E. Kimmel, M.D., Benjamin French, Ph.D., Scott E. Kasner, M.D., Julie A. Johnson, Pharm.D., Jeffrey L. Anderson, M.D., Brian F. Gage, M.D., Yves D. Rosenberg, M.D., Charles S. Eby, M.D., Ph.D., Rosemary A. Madigan, R.N., M.P.H., Robert B. McBane, M.D., Sherif Z. Abdel-Rahman, Ph.D., Scott M. Stevens, M.D., Steven Yale, M.D., Emile R. Mohler III, M.D., Margaret C. Fang, M.D., Vinay Shah, M.D., Richard B. Horenstein, M.D., Nita A. Limdi, Pharm.D., Ph.D., James A.S. Muldowney III, M.D., Jaspal Gujral, M.B., B.S., Patrice Delafontaine, M.D., Robert J. Desnick, M.D., Ph.D., Thomas L. Ortel, M.D., Ph.D., Henny H. Billett, M.D., Robert C. Pendleton, M.D., Nancy L. Geller, Ph.D., Jonathan L. Halperin, M.D., Samuel Z. Goldhaber, M.D., Michael D. Caldwell, M.D., Ph.D., Robert M. Califf, M.D., and Jonas H. Ellenberg, Ph.D., for the COAG Investigators*

The authors’ affiliations are listed in the Appendix. Address reprint requests to Dr. Kimmel at the Center for Therapeutic Effectiveness Research, 923 Blockley Hall, 423 Guardian Dr., Philadelphia, PA 19104-6021, or at stevek@mail.med.upenn.edu.

* A complete list of investigators and committees in the Clarification of Optimal Anticoagulation through Genetics (COAG) trial is provided in the Supplementary Appendix, available at NEJM.org.

This article was published on November 19, 2013, at NEJM.org.
DOI: 10.1056/NEJMoa1310669

Copyright © 2013 Massachusetts Medical Society

ABSTRACT

Background

The clinical utility of genotype-guided (pharmacogenetically based) dosing of warfarin has been tested only in small clinical trials or observational studies, with equivocal results.

Methods

We randomly assigned 1015 patients to receive doses of warfarin during the first 5 days of therapy that were determined according to a dosing algorithm that included both clinical variables and genotype data or to one that included clinical variables only. All patients and clinicians were unaware of the dose of warfarin during the first 4 weeks of therapy. The primary outcome was the percentage of time that the international normalized ratio (INR) was in the therapeutic range from day 4 or 5 through day 28 of therapy.

Results

At 4 weeks, the mean percentage of time in the therapeutic range was 45.2% in the genotype-guided group and 45.4% in the clinically guided group (adjusted mean difference, [genotype-guided group minus clinically guided group], −0.2; 95% confidence interval, −3.4 to 3.1; P = 0.91). There also was no significant between-group difference among patients with a predicted dose difference between the two algorithms of 1 mg per day or more. There was, however, a significant interaction between dosing strategy and race (P = 0.003). Among black patients, the mean percentage of time in the therapeutic range was less in the genotype-guided group than in the clinically guided group. The rates of the combined outcome of any INR of 4 or more, major bleeding, or thromboembolism did not differ significantly according to dosing strategy.

Conclusions

Genotype-guided dosing of warfarin did not improve anticoagulation control during the first 4 weeks of therapy. (Funded by the National Heart, Lung, and Blood Institute and others; COAG ClinicalTrials.gov number, NCT00839657.)
Genotype-guided initiation of warfarin

In the genotype-guided group, the mean percentage of time in the therapeutic range was significantly lower than in the clinically guided group (35.2% vs. 43.5%; adjusted mean difference, −8.3%; P = 0.01). Among nonblack patients, the mean percentage of time in the therapeutic range was slightly higher in the genotype-guided group than in the clinically guided group (48.8% vs. 46.1%; adjusted mean difference, 2.8%; P = 0.15). There were no significant differences in the percentage of time in the therapeutic range according to sex or the total number of genetic variants (Table 2).

Anticoagulation Control and Dose Prediction

There were no significant between-group differences in the mean percentage of time above the therapeutic range (INR, >3) or below the therapeutic range (INR, <2) (Fig. 2, and Table S3 in the Supplementary Appendix). However, black patients in the genotype-guided group were more likely to have INRs above the therapeutic range than were those in the clinically guided group (Fig. S2 and Table S3 in the Supplementary Appendix).

There was no overall between-group difference in the time to the first INR in the therapeutic range (Table S4 in the Supplementary Appendix). However, black patients in the genotype-guided group took longer on average to reach the first therapeutic INR than did those in the clinically guided group (Table S4 and Fig. S3 in the Supplementary Appendix). The time to the determination of the maintenance dose did not differ significantly between the two groups overall or among patients stratified by sex, total number of genetic variants, or race (Table 2).

Figure 1. Distribution of Time in the Therapeutic Range.

Side-by-side density plots show the distribution of the percentage of time in the therapeutic range of the international normalized ratio (INR) from the completion of the intervention period (day 4 or 5) to day 28 of therapy for the two study groups among all patients (at left), among patients stratified according to the absolute difference in the predicted initial daily dose of warfarin between the two algorithms (≥1 mg [primary subgroup] vs. <1 mg) (at top right), and among patients stratified according to race (at bottom right). The horizontal lines indicate the mean percentage of time in the therapeutic range.