Predicting hemoglobin levels in whole blood donors using transition models and mixed effects models

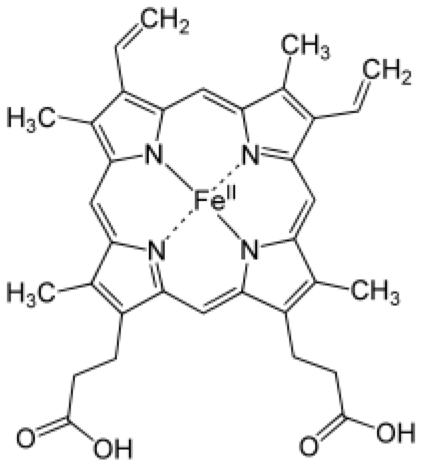
公衛三 梁嫚芳 b07801003

Contents

Background

Blood transfusion

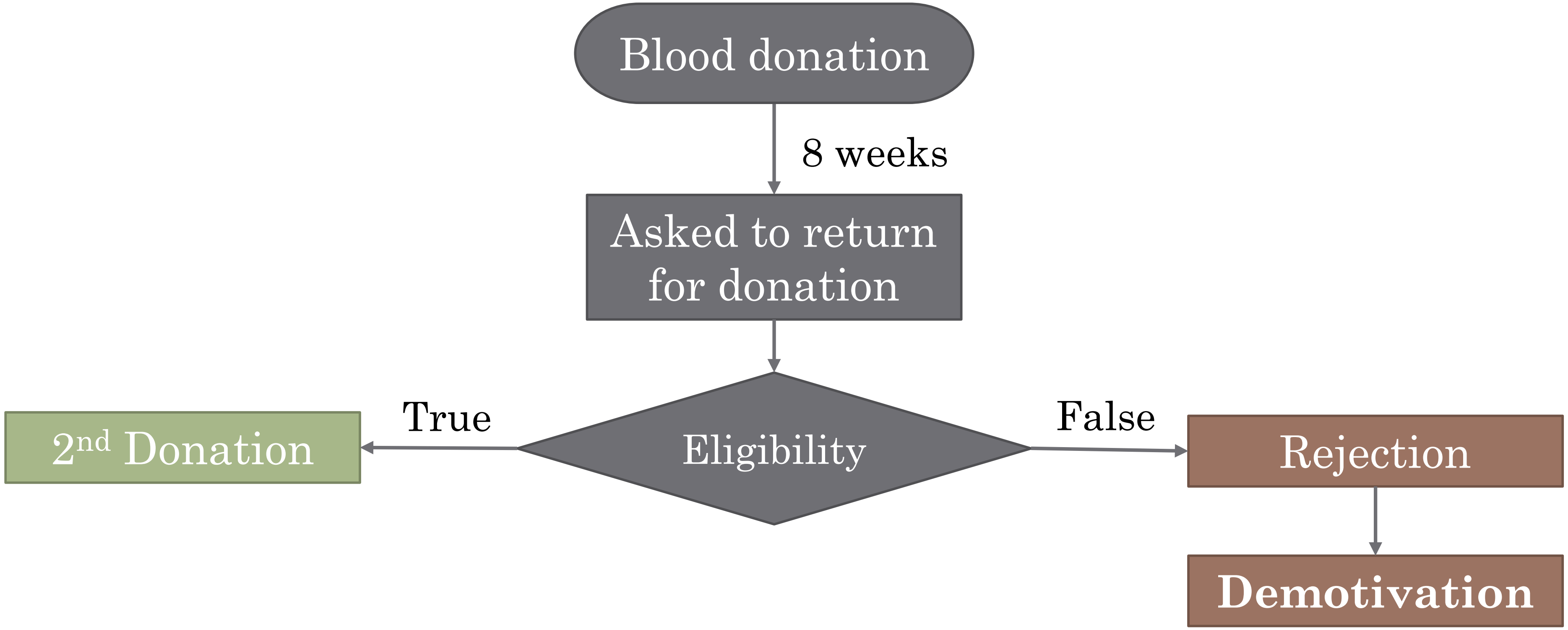
- Essential part of modern healthcare
  - ∵ Helps save millions of lives each year
- Blood donations are in great need!
  - ∵ Artificial substitute of blood has yet to be found



- Ineligibility: **hemoglobin (Hb) level**
  - Netherland
    - Male: < 8.4 mmol/l (135 g/l)
    - Female: < 7.8 mmol/l (125 g/l)
  - Taiwan
    - Male: < 13 g/dl
    - Female: < 12 g/dl

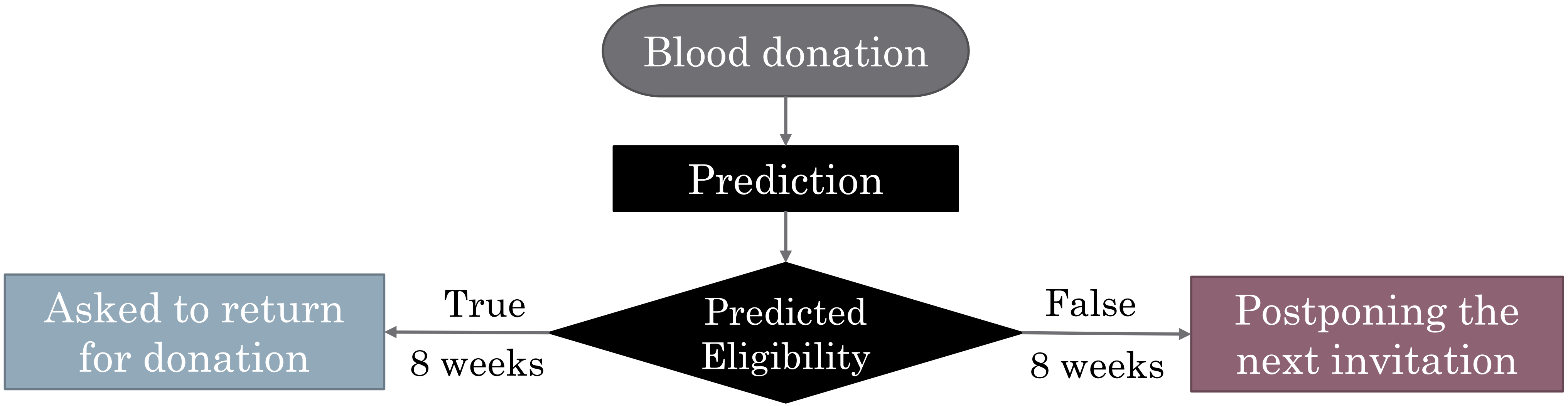
Difficulties in Blood Donation

- **Ineligibility → Demotivation of donor**



How to improve?

- **Predict donors' Hb level → Apply appropriate interventions**
  - (e.g. planning of donors' visits)



Methods

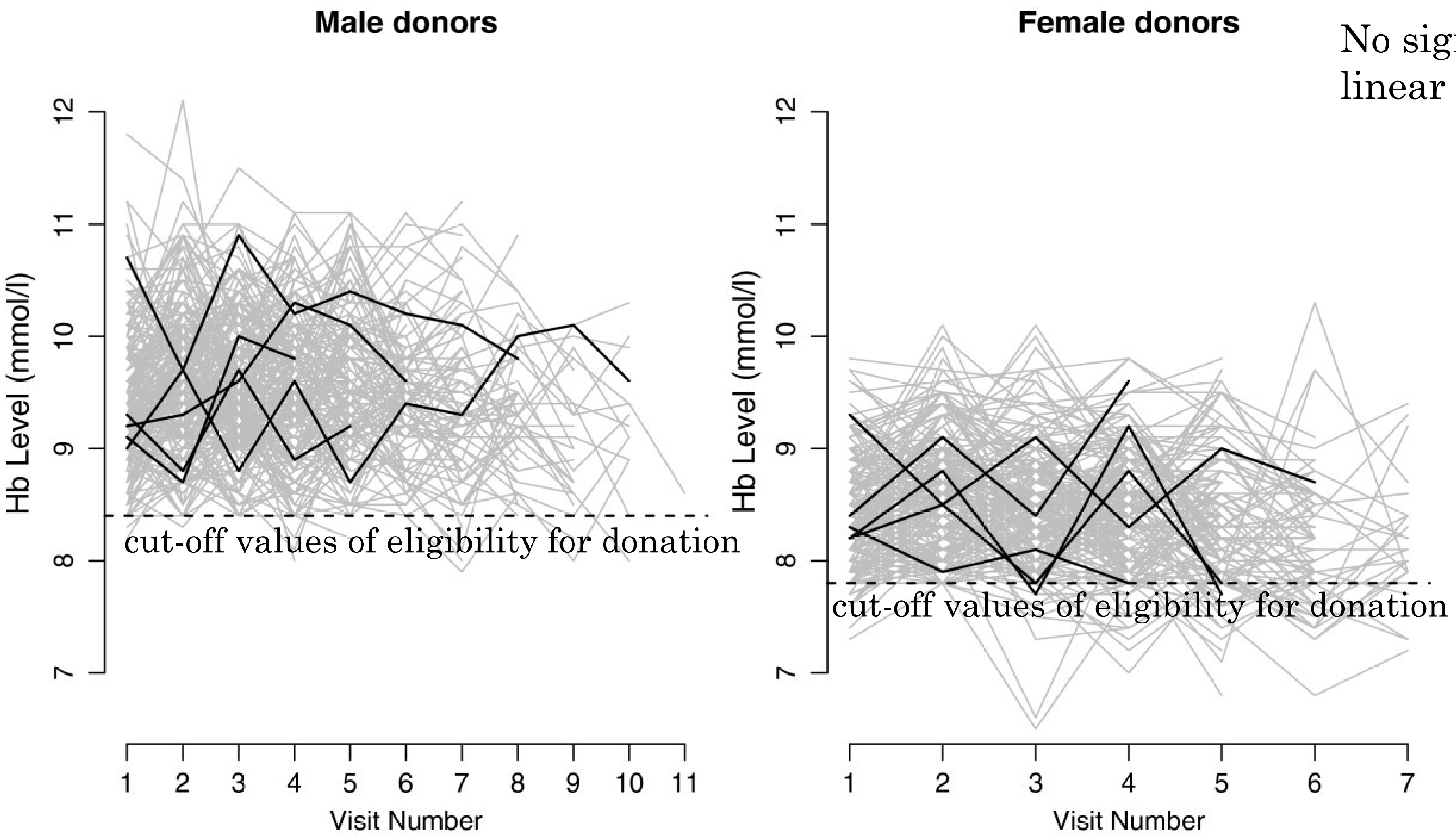
- Data
- Training and Validation (Prediction performance)
- Statistical models
  - Multiple linear regression model
  - Transition model
  - Linear mixed effects model

Data

- Source: Sanquin Blood Supply, Netherlands
- Sample
  - During the time period, first-visit Donors the blood bank and donated at least twice
  - Time period: January 1, 2007 ~ December 31, 2009
  - Donation Type: Whole blood donation
- Sample size: 15,625 (54.6% women)
- Type of data: Repeated measurement

Variables in the data

- Sequential number of the visit
- **Outcomes**
  - Hb level (Continuous, mmol/l)
- **Predictors**
  - Age (Continuous)
    - Hb level may decrease when aging
  - Season (cold - fall & winter = 0, warm - spring & summer = 1)
    - Hb level is lower on average in warm season
  - DPV: status of the previous visit (donation = 1, deferral = 0)
    - The first visit is defined to be ‘no donation’
- **Stratification Variable**
  - Sex (Female, Male)



No significant linear correlation

- **Hemoglobin levels profile.** Profile of hemoglobin levels for successive visits to the blood bank of a random sample of male and female donors. The profiles of 5 randomly selected donors are highlighted.

Statistical models

- Multiple linear regression model
  - As a benchmark to show the capability of transition and mixed effects models
- Transition model
- Linear mixed effects model

Workflow of the research

Inputs

- Age
- Season
- DPV

Statistical models

- Multiple linear regression model
- Transition model
- Linear mixed effects model

Outputs

- Hb level

Compare prediction performance

Workflow of the research

Inputs

- Age
- Season
- DPV

Statistical models

- Multiple linear regression model
- Transition model
- Linear mixed effects model

Outputs

- Hb level

Compare prediction performance

- AIC
- BIC
- MSPE
- AUC

Tools

R version 2.15.2

- stats package for the multiple linear regression models
- nlme package for the mixed effects models
- KalmanLike and the mle functions in the stats4 package for the transition models
- mixAK and pROC packages to draw profile and ROC curve plots.

Significance level of $\alpha = 0.05$ and no correction for multiple testing was implemented.

Training and Validation

Model selection criteria

Validation criteria

Descriptive Statistics

Training and Validation

- Randomly divide samples (n=15625) into
  - Training set (n=7709)
  - Validation set (n=7916)
- Model
  - Multiple Linear regression model
  - Transition model : AR(1)~AR(5)
  - Mixed effects model
- **Model selection criteria**
  - **AIC**
  - **BIC**
  - **MSPE** (Mean Squared Prediction Error)

- Linear regression and transition model
  - $AIC = -2logL + 2p$
  - $BIC = -2logL + 2p \times log(n)$

  (n : sample size)
- Mixed effect model
  - $AIC = -2logL_R + 2p$
  - $BIC = -2logL_R + 2p \times log(n)$

  (n : cluster number)
  (L_R =Restricted (Residual) log likelihood)

  Used to select R or G matrix
- MSPE
  - Use MSE to be the error metric :

$$MSPE_t = \sum_{i=1}^{N_t} \frac{(\hat{y}_{it} - y_{it})^2}{N_t}$$

Training and Validation

- Randomly divide samples (n=15625) into
  - Training set (n=7709)
  - Validation set (n=7916)
- Model
  - Multiple Linear regression model
  - Transition model : AR(1)~AR(5)
  - Mixed effects model
- **Validation criteria**
  - **AUC** (Area under the ROC curve)
  - **MSPE** (Mean Squared Prediction Error)

- AUC
  - Ability to discrimination
  - Using bootstrap technique to test the difference
- MSPE
  - Use MSE to be the error metric：

$$MSPE_t = \sum_{i=1}^{N_t} \frac{(\hat{y}_{it} - y_{it})^2}{N_t}$$

Descriptive Statistics

- The distribution of covariates is similar
- Median of visit number is 5

Data set	Gender	#Donor	#Deferral	#Cold Season	Age: Mean (SD)	Visit: Med (IQR)
Training data set	Male	3610	769 (4.58%)	10213 (50.05%)	34.57 (12.9)	5 (3)
	Female	4306	1596 (9.62%)	10387 (49.71%)	32.66 (12.8)	5 (1)
	Total	**7916**	**2365 (7.08)**%	**20600 (49.88%)**	**33.53 (12.9)**	**5 (2)**
Validation data set	Male	3449	688 (4.27%)	9781 (49.95%)	34.28 (12.6)	5 (3)
	Female	4260	1729 (10.41%)	10341 (49.54%)	32.77 (12.8)	5 (2)
	Total	**7709**	**2417 (7.38%)**	**20122 (49.74%)**	**33.45 (12.7)**	**5 (2)**

Note: SD= Standard deviation, IQR= Interquartile range.

Statistical models

Multiple linear regression model

Transition model

Linear mixed effects model

Multiple linear regression model

$$y_{it} = \alpha + \beta_1 Age_{it} + \beta_2 Season_{it} + \beta_3 DPV_{it} + \epsilon_{it}, \quad (1)$$

where y_{it} is the tth observation of the ith individual,
α is an unknown constant (intercept),
β's are unknown regression coefficients.
Assumed that the residuals ε_{it} are normally distributed and mutua
independent with mean zero and constant variance, i.e.,

$$\varepsilon_{it} \overset{\text{iid}}{\sim} N\left(0, \sigma_\epsilon^2\right)$$

Due to the fact that this model cannot take into account the intra-subject correlations and the previous Hb levels, it is only presented as a benchmark model to show the capability of transition and mixed effects models.

Transition model (with order q)

$$
\begin{aligned}
y_{it} = \alpha &+ \beta_1 Age_{it} + \beta_2 Season_{it} + \beta_3 DPV_{it} \\
&+ \sum_{r=1}^{q} \gamma_r (y_{it-r} - (\beta_1 Age_{it-r} \\
&\qquad + \beta_2 Season_{it-r} + \beta_3 DPV_{it-r})) \\
&+ \epsilon_{it},
\end{aligned} \tag{2}
$$

where Age_{it-r}, $Season_{it-r}$, DPV_{it-r} are rth lagged response & covariates, respectively

γ_r is the corresponding coefficient of the rth lag.

- Linear quadratic estimation (Kalman filter)
  - Calculate the exact likelihood function
- Including the information of donors who have made fewer visits than the order of the transition model

Linear mixed effects model

- Apply the empirical Bayes method (EB, not MME) to predict random effect
- Use a likelihood ratio test to choose random intercept model or random intercept and slope model

- MME : Mixed model quation

$$\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix}^{-1} \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

$$\operatorname{var}\begin{bmatrix} \hat{\beta} \\ \hat{\gamma} \end{bmatrix} = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix}^{-1}$$

Linear mixed effects model

- Random intercept model - Male

$$y_{it} = \alpha + b_{0i} + \beta_1 Age_{it} + \beta_2 Season_{it} + \beta_3 DPV_{it} + \epsilon_{it}, \quad (3)$$

where β's are regression coefficients (fixed effects),
b_{io} is the random intercept, the deviation of the ith subject-specific mean from the population mean of Hb levels.
Assumed that b_{oi} and ε_{it} are normally distributed and mutually independent with mean zero and different constant variances, i.e.,

Linear mixed effects model

- Random intercept and slope model - Female

$$y_{it} = \alpha + b_{0i} + (b_{1i} + \beta_1)Age_{it} + \beta_2 Season_{it} + \beta_3 DPV_{it} + \epsilon_{it}, \tag{4}$$

where β's contains population-specific parameters.
$b_i = (b_{i0}, b_{i1})$ contains subject-specific parameters (intercept and the effects of age)
b_{i1} = deviation of the ith subject-specific slope from the population mean of Hb levels,
εi=(εi1,…,εini)′ is a vector containing the common error components, with $\varepsilon_i \sim N(0, \Sigma_i)$.

Results

Variables

Model selection

Variables

Age

Season

DPV

Table 2 Parameter estimates (standard errors) of the models estimated using the training data set for male donors

Parameter	*Model LR*	*AR(1)*	*AR(2)*	*AR(3)*	*AR(4)*	*AR(5)*	*Model LME*
intercept	9.6448	9.6309	9.6441	9.6560	9.6617	9.6633	9.6719
	(0.0142)	(0.0206)	(0.0231)	(0.0243)	(0.0246)	(0.0247)	(0.0243)
Age	-0.0045	-0.0043	-0.0044	-0.0045	-0.0047	-0.0047	-0.0049
	(0.0003)	(0.0005)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0006)
Season(Warm)	-0.0627	-0.0615	-0.0681	-0.0699	-0.0693	-0.0694	-0.0698
	(0.0089)	(0.0074)	(0.0066)	(0.0066)	(0.0067)	(0.0067)	(0.0067)
DPV	-0.0610	-0.0469	-0.0350	-0.0385	-0.0440	-0.0474	-0.0636
(Donation)	(0.0092)	(0.0089)	(0.0079)	(0.0074)	(0.0072)	(0.0072)	(0.0068)
γ_1	—	0.5158	0.3685	0.3053	0.2746	0.2630	—
	—	(0.0061)	(0.0068)	(0.0076)	(0.0082)	(0.0087)	—
γ_2	—	—	0.2888	0.2080	0.1766	0.1621	—
	—	—	(0.0078)	(0.0087)	(0.0084)	(0.0091)	—
γ_3	—	—	—	0.2207	0.1730	0.1581	—
	—	—	—	(0.0095)	(0.0104)	(0.0109)	—
γ_4	—	—	—	—	0.1488	0.1257	—
	—	—	—	—	(0.0123)	(0.0129)	—
γ_5	—	—	—	—	—	0.0829	—
	—	—	—	—	—	(0.0167)	—

Table 2 Parameter estimates (standard errors) of the models estimated using the training data set for male donors

Parameter	*Model LR*	*AR(1)*	*AR(2)*	*AR(3)*	*AR(4)*	*AR(5)*	*Model LME*
intercept	9.6448	9.6309	9.6441	9.6560	9.6617	9.6633	9.6719
	(0.0142)	(0.0206)	(0.0231)	(0.0243)	(0.0246)	(0.0247)	(0.0243)
Age	-0.0045	-0.0043	-0.0044	-0.0045	-0.0047	-0.0047	-0.0049
	(0.0003)	(0.0005)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0006)
Season(Warm)	-0.0627	-0.0615	-0.0681	-0.0699	-0.0693	-0.0694	-0.0698
	(0.0089)	(0.0074)	(0.0066)	(0.0066)	(0.0067)	(0.0067)	(0.0067)
DPV	-0.0610	-0.0469	-0.0350	-0.0385	-0.0440	-0.0474	-0.0636
(Donation)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	(0.0068)
γ_1	—	0.5158	0.3685	0.3053	0.2746	0.2630	—
	—	(0.0061)	(0.0068)	(0.0076)	(0.0082)	(0.0087)	—
γ_2	—	—	0.2888	0.2080	0.1766	0.1621	—
	—	—	(0.0078)	(0.0087)	(0.0084)	(0.0091)	—
γ_3	—	—	—	0.2207	0.1730	0.1581	—
	—	—	—	(0.0095)	(0.0104)	(0.0109)	—
γ_4	—	—	—	—	0.1488	0.1257	—
	—	—	—	—	(0.0123)	(0.0129)	—
γ_5	—	—	—	—	—	0.0829	—
	—	—	—	—	—	(0.0167)	—

Hb level statistically significantly decreases when aging

Table 2 Parameter estimates (standard errors) of the models estimated using the training data set for male donors

Parameter	*Model LR*	*AR(1)*	*AR(2)*	*AR(3)*	*AR(4)*	*AR(5)*	*Model LME*
intercept	9.6448	9.6309	9.6441	9.6560	9.6617	9.6633	9.6719
	(0.0142)	(0.0206)	(0.0231)	(0.0243)	(0.0246)	(0.0247)	(0.0243)
Age	-0.0045	-0.0043	-0.0044	-0.0045	-0.0047	-0.0047	-0.0049
	(0.0003)	(0.0005)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0006)
Season(Warm)	-0.0627	-0.0615	-0.0681	-0.0699	-0.0693	-0.0694	-0.0698
	(0.0089)	(0.0074)	(0.0066)	(0.0066)	(0.0067)	(0.0067)	(0.0067)
DPV	-0.0610	-0.0469	-0.0350	-0.0385	-0.0440	-0.0474	-0.0636
(Donation)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	(0.0068)
γ_1	—	0.5158	0.3685	0.3053	0.2746	0.2630	—
	—	(0.0061)	(0.0068)	(0.0076)	(0.0082)	(0.0087)	—
γ_2	—	—	0.2888	0.2080	0.1766	0.1621	—
	—	—	(0.0078)	(0.0087)	(0.0084)	(0.0091)	—
γ_3	—	—	—	0.2207	0.1730	0.1581	—
	—	—	—	(0.0095)	(0.0104)	(0.0109)	—
γ_4	—	—	—	—	0.1488	0.1257	—
	—	—	—	—	(0.0123)	(0.0129)	—
γ_5	—	—	—	—	—	0.0829	—
	—	—	—	—	—	(0.0167)	—

Hb level is lower on average during warm season

Table 2 Parameter estimates (standard errors) of the models estimated using the training data set for male donors

Parameter	*Model LR*	*AR(1)*	*AR(2)*	*AR(3)*	*AR(4)*	*AR(5)*	*Model LME*
intercept	9.6448	9.6309	9.6441	9.6560	9.6617	9.6633	9.6719
	(0.0142)	(0.0206)	(0.0231)	(0.0243)	(0.0246)	(0.0247)	(0.0243)
Age	-0.0045	-0.0043	-0.0044	-0.0045	-0.0047	-0.0047	-0.0049
	(0.0003)	(0.0005)	(0.0006)	(0.0006)	(0.0006)	(0.0007)	(0.0006)
Season(Warm)	-0.0627	-0.0615	-0.0681	-0.0699	-0.0693	-0.0694	-0.0698
	(0.0089)	(0.0074)	(0.0066)	(0.0066)	(0.0067)	(0.0067)	(0.0067)
DPV	-0.0610	-0.0469	-0.0350	-0.0385	-0.0440	-0.0474	-0.0636
(Donation)	(0.0092)	(0.0089)	(0.0079)	(0.0074)	(0.0072)	(0.0072)	(0.0068)
γ_1	—	0.5158	0.3685	0.3053	0.2746	0.2630	—
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	—
γ_2	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	—
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	—
γ_3	—	—	—	0.2207	0.1730	0.1581	—
	—	—	—	(0.0095)	(0.0104)	(0.0109)	—
γ_4	—	—	—	—	0.1488	0.1257	—
	—	—	—	—	(0.0123)	(0.0129)	—
γ_5	—	—	—	—	—	0.0829	—
	—	—	—	—	—	(0.0167)	—

Donation in the previous visit has a negative effect on the current Hb level

Table 2 Parameter estimates (standard errors) of the models estimated using the training data set for male donors

Parameter	*Model LR*	*AR(1)*	*AR(2)*	*AR(3)*	*AR(4)*	*AR(5)*	*Model LME*
intercept	9.6448	9.6309	9.6441	9.6560	9.6617	9.6633	9.6719
	(0.0142)	(0.0206)	(0.0231)	(0.0243)	(0.0246)	(0.0247)	(0.0243)
Age	-0.0045	-0.0043	-0.0044	-0.0045	-0.0047	-0.0047	-0.0049
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	(0.0006)
Season(Warm)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0698
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	(0.0067)
DPV	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	-0.0636
(Donation)	(0.0092)	(0.0089)	(0.0079)	(0.0074)	(0.0072)	(0.0072)	(0.0068)
γ_1	—	0.5158	0.3685	0.3053	0.2746	0.2630	—
	—	(0.0061)	(0.0068)	(0.0076)	(0.0082)	(0.0087)	—
γ_2	—	—	0.2888	0.2080	0.1766	0.1621	—
	—	—	(0.0078)	(0.0087)	(0.0084)	(0.0091)	—
γ_3	—	—	—	0.2207	0.1730	0.1581	—
	—	—	—	(0.0095)	(0.0104)	(0.0109)	—
γ_4	—	—	—	—	0.1488	0.1257	—
	—	—	—	—	(0.0123)	(0.0129)	—
γ_5	—	—	—	—	—	0.0829	—
	—	—	—	—	—	(0.0167)	—

Transition effects (regression coefficients of past Hb values) are significant, although the effect of previous Hb level decreases with the lag

Table 3 Parameter estimates (standard errors) of the models estimated using the training data set for female donors

Parameter	*Model LR*	*AR(1)*	*AR(2)*	*AR(3)*	*AR(4)*	*AR(5)*	*Model LME*
intercept	8.2737	8.2394	8.2555	8.2678	8.2698	8.2702	8.2832
	(0.0123)	(0.0164)	(0.0180)	(0.0186)	(0.0187)	(0.0187)	(0.0181)
Age	0.0042	0.0044	0.0042	0.0040	0.0040	0.0040	0.0037
	(0.0003)	(0.0004)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Season(Warm)	-0.0347	-0.0405	-0.0415	-0.0413	-0.0415	-0.0415	-0.0411
	(0.0078)	(0.0062)	(0.0060)	(0.0062)	(0.0061)	(0.0061)	(0.0062)
DPV	-0.1106	-0.1411	-0.1273	-0.1307	-0.1335	-0.1346	-0.1387
(Donation)	(0.0079)	(0.0075)	(0.0067)	(0.0064)	(0.0063)	(0.0063)	(0.0060)
γ_1	—	0.4669	0.3457	0.3012	0.2878	0.2830	—
	—	(0.0062)	(0.0067)	(0.0074)	(0.0080)	(0.0084)	—
γ_2	—	—	0.2573	0.1963	0.1793	0.1693	—
	—	—	(0.0080)	(0.0088)	(0.0089)	(0.0099)	—
γ_3	—	—	—	0.1742	0.1486	0.1360	—
	—	—	—	(0.0100)	(0.0112)	(0.0121)	—
γ_4	—	—	—	—	0.0831	0.0623	—
	—	—	—	—	(0.0157)	(0.0182)	—
γ_5	—	—	—	—	—	0.0681	—
	—	—	—	—	—	(0.0264)	—

Table 3 Parameter estimates (standard errors) of the models estimated using the training data set for female donors

Parameter	*Model LR*	*AR(1)*	*AR(2)*	*AR(3)*	*AR(4)*	*AR(5)*	*Model LME*
intercept	8.2737	8.2394	8.2555	8.2678	8.2698	8.2702	8.2832
	(0.0123)	(0.0164)	(0.0180)	(0.0186)	(0.0187)	(0.0187)	(0.0181)
Age	0.0042	0.0044	0.0042	0.0040	0.0040	0.0040	0.0037
	(0.0003)	(0.0004)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Season(Warm)	-0.0347	-0.0405	-0.0415	-0.0413	-0.0415	-0.0415	-0.0411
	(0.0078)	(0.0062)	(0.0060)	(0.0062)	(0.0061)	(0.0061)	(0.0062)
DPV	-0.1106	-0.1411	-0.1273	-0.1307	-0.1335	-0.1346	-0.1387
(Donation)	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	(0.0060)
γ_1	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	—
	—	(0.0062)	(0.0067)	(0.0074)	(0.0080)	(0.0084)	—
γ_2	—	—	0.2573	0.1963	0.1793	0.1693	—
	—	—	(0.0080)	(0.0088)	(0.0089)	(0.0099)	—
γ_3	—	—	—	0.1742	0.1486	0.1360	—
	—	—	—	(0.0100)	(0.0112)	(0.0121)	—
γ_4	—	—	—	—	0.0831	0.0623	—
	—	—	—	—	(0.0157)	(0.0182)	—
γ_5	—	—	—	—	—	0.0681	—
	—	—	—	—	—	(0.0264)	—

Contrary to male donors, Hb level of female donors statistically significantly increases when aging

Table 3 Parameter estimates (standard errors) of the models estimated using the training data set for female donors

The directions of other variables' effect on Hb level in female donors are the same as those in male donors.

Parameter	*Model LR*	*AR(1)*	*AR(2)*	*AR(3)*	*AR(4)*	*AR(5)*	*Model LME*
intercept	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	8.2832
	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	(0.0181)
Age	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	0.0037
	(0.0003)	(0.0004)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)
Season(Warm)	-0.0347	-0.0405	-0.0415	-0.0413	-0.0415	-0.0415	-0.0411
	(0.0078)	(0.0062)	(0.0060)	(0.0062)	(0.0061)	(0.0061)	(0.0062)
DPV	-0.1106	-0.1411	-0.1273	-0.1307	-0.1335	-0.1346	-0.1387
(Donation)	(0.0079)	(0.0075)	(0.0067)	(0.0064)	(0.0063)	(0.0063)	(0.0060)
γ_1	—	0.4669	0.3457	0.3012	0.2878	0.2830	—
	—	(0.0062)	(0.0067)	(0.0074)	(0.0080)	(0.0084)	—
γ_2	—	—	0.2573	0.1963	0.1793	0.1693	—
	—	—	(0.0080)	(0.0088)	(0.0089)	(0.0099)	—
γ_3	—	—	—	0.1742	0.1486	0.1360	—
	—	—	—	(0.0100)	(0.0112)	(0.0121)	—
γ_4	—	—	—	—	0.0831	0.0623	—
	—	—	—	—	(0.0157)	(0.0182)	—
γ_5	—	—	—	—	—	0.0681	—
	—	—	—	—	—	(0.0264)	—

Model selection

AIC

BIC

MSPE

AUC

On the training data (AIC, BIC)

- **Male : Transition model AR(5)**
- **Female : Random slope and intercept model**

Table 4 AIC, BIC, and MSEP values for different models for both genders based on the training data set

	Male donors			Female donors		
Model	**AIC**	**BIC**	**MSPE**	**AIC**	**BIC**	**MSPE**
Linear Regression	37087.8	37127.2	4.14	35968.9	36008.6	2.29
Mixed Effects	30524.3	30571.6	2.90	30058.0	30113.6	1.75
AR(1)	32051.0	32098.3	3.07	31559.1	31606.7	1.81
AR(2)	30936.4	30991.6	2.85	30664.7	30720.3	1.73
AR(3)	30471.9	30535.0	2.78	30375.1	30438.7	1.71
AR(4)	30342.5	30413.4	2.78	30341.7	30413.2	1.72
AR(5)	30321.4	30400.2	2.79	30325.1	30404.5	1.72

Note: Lower values of AIC, BIC, and MSEP indicate better model fit.

On the validation data (MSPE)

- **Transition model provides a better prediction than the mixed effects model, especially at high visit numbers**
- At first visit number, all data are independent, so linear regression model's MSPE closes to others.
- As visit number increases, data become correlated, so linear regression model's MSPE also increases. (poor accuracy)

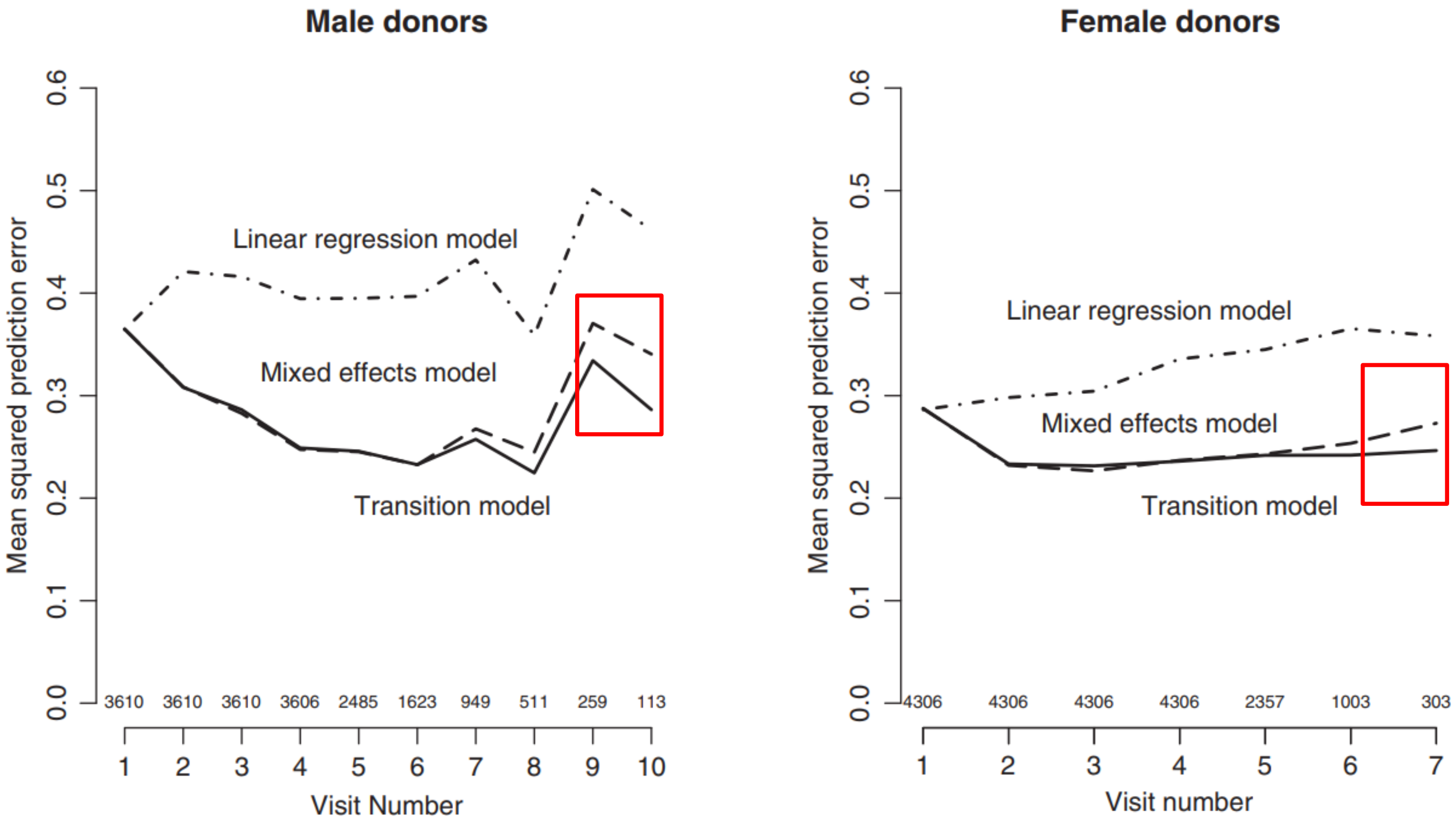


Figure 2 Mean squared prediction error. Mean squared prediction error of the linear regression model, the linear mixed effects model, and the 5th order transition model, as a function of the visit number. The included numbers of individuals are displayed above the horizontal axis.

On the validation data (AUC)

- **Transition model has a larger AUC than mixed effects model (p <.001)**, and thus offers a better trade-off between sensitivity and specificity.
- Clinical cut-off value is not the best one in both ROC curve

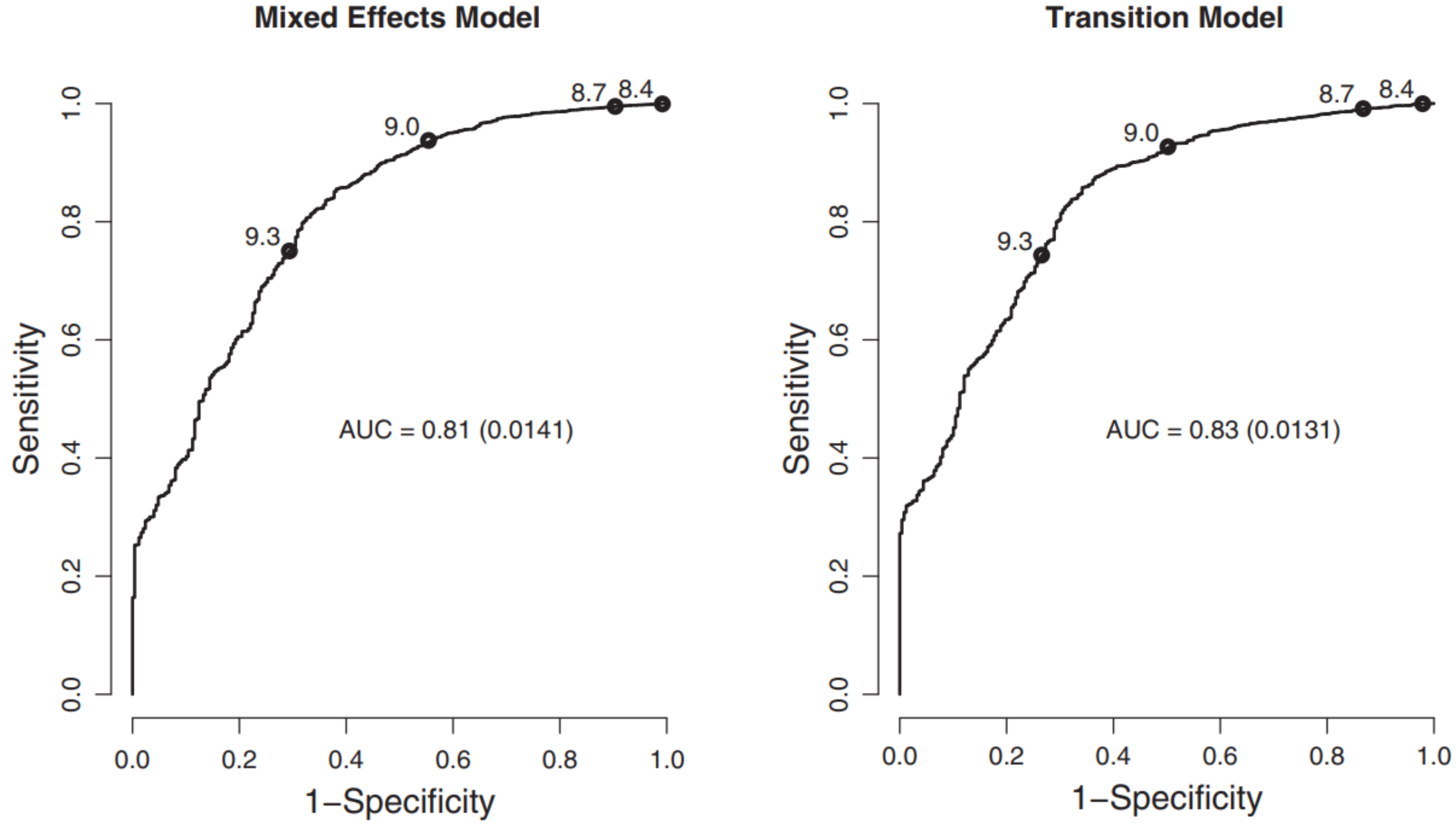


Figure 3 ROC curves for male donors. ROC curves of the prediction of eligibility for donation in male donors, for two different models. The standard errors of the AUCs are shown in parentheses. Different cut-off points for the predicted value are displayed on the curves.

Discussion

Advantage of transition model

Limitations

Advantage of transition model

- Transition model offers better predictions for longer time series
- Transition model is convenient in practice and needs less historical information compared to the mixed effects model

Limitations

- Data set used is unbalanced in the sense that the time intervals between visits vary considerably
- More factors that are possibly associated with Hb level
  - such as physical activity, race, nutrition and smoking status
- The ultimate purpose of is not the prediction of the future Hb value, but rather to determine the best time for the donor to return for donation
  - Future work: focus on the optimal timing of future donations

Conclusion

Conclusion

- Transition model provides a better prediction than the mixed effects model, especially at high visit numbers.
- The paper shows the capabilities of using longitudinal models for prediction and that our findings may help reduce the number of deferred candidate in the blood banks.