Latent Class Mixed-Effects Transition Model: A model to predict hemoglobin in blood donors

Kazem Nasserinejad[1], Joost van Rosmalen[2], Wim de Kort[3], Dimitris Rizopoulos[4], Emmanuel Lesaffre[5]

[1]k.nasserinejad@erasmusmc.nl, [2]j.vanrosmalen@erasmusmc.nl, [4]d.rizopoulos@erasmusmc.nl@erasmusmc.nl, Department of Biostatistics, Erasmus MC.

[3] w.dekort@sanquin.nl , Department of Donor Studies, Sanquin Research & Department of Public Health, Academic Medical Center in Amsterdam.

[5]emmanuel.lesaffre@med.kuleuven.be, Department of L-Biostat, KU Leuven.

Abstract

Due to a) heterogeneity of the initial hemoglobin (Hb) level, b) state dependence of a donor's Hb values, c) varying time intervals between donations, d) the temporary reduction in Hb after blood donation, and e) the fact that the recovery process may change with the number of donations and may differ between donors, longitudinal data of Hb values of blood donors provide unique statistical challenges. To estimate the shape and duration of the recovery process, and to predict future Hb values, we proposed a latent class mixed-effects transition model.

Keywords: latent class mixed-effects model; transition model; change point model.

1. Introduction

Blood donors experience a temporary reduction in their hemoglobin (Hb) value after donation. At each visit the Hb value is measured, and a too low Hb value leads to a deferral for donation. Prediction of Hb for the subsequent visit for a donor is both important and not straightforward. It is important because deferred candidates rarely come back for donation. It is not straightforward because of a) heterogeneity of the initial Hb level, b) state dependence of a donor's Hb values, c) varying time intervals between donations, d) the temporary reduction in Hb after blood donation, and e) the fact that the recovery process may change with the number of donations and may differ between donors.

This requires a complex statistical model to capture these sources of variation for proper inference. To build this appropriate statistical model to estimate the shape and duration of the recovery process, and to predict future Hb values, we proposed a latent class mixed-effects transition model for the Hb values. In this model, a flexible function was used to model the recovery process after donation. All models were estimated in a Bayesian way, using data of new entrant blood donors during 2005 to 2012 in the Donor InSight study in the Netherlands. Informative priors were used for parameters of the recovery process that were not identified using the observed data, based on results from the previous literatures. Parameter estimates are

obtained using a MCMC algorithm by the JAGS. We also show how to forecast a future Hb value based on the available history of donations in a fully Bayesian approach.

2.   Hb recovery after blood donation

The Hb recovery process after blood donation is illustrated in Figure 1. In this figure θ is the amount of Hb reduction after a blood donation, δ indicates the time that Hb reaches its minimum value after donation. Indeed, after donation of 500 ml (8% of the blood volume) of blood, on average a male donor looses 242 mg and a female donor 217 mg of iron. This will cause Hb to decrease and to reach its minimum value in a few days after donation. Then the Hb value starts to recover gradually to its pre-donation level. The time that is needed for a full recovery is given by RT, and TSPD represents the time since previous donation.

Figure 1. Hb recovery process after blood donation.
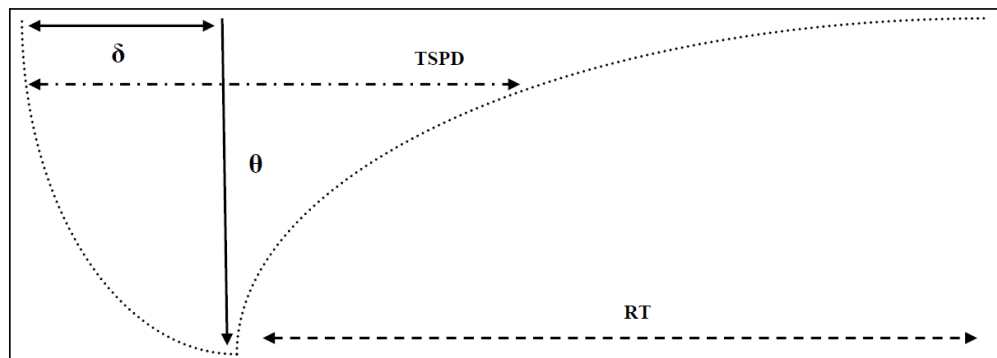
Unfortunately in this study we have only the Hb value prior to donation at each visit, and there is no information on the Hb level between two invitations. The observed inter-donation interval is at least 56 days in our data set. Therefore, we cannot accurately estimate the trajectory of Hb value during the first 56 days after donation. We incorporate in our model that the reduction of Hb after giving blood takes around three days based on previous studies results, and that the amount of the reduction is approximately 8% of the Hb value. This information was incorporated in our proposed model via informative priors, since without this extra information fitting such a complex model was not possible. We use the following Hb recovery function (HRF):

$$ HRF_{it} = \theta \left[\max \left(\frac{RT - (TSPD_{it} - \delta)}{RT}, 0 \right) \right]^{\psi}, \qquad (1) $$

$TSPD_{it}$ represents the time since previous donation for the ith donor at time t. The parameter ψ indicates the shape of the recovery function of the Hb value after donation. Values of ψ smaller than one indicate a convex trajectory for Hb recovery, while values greater than

one indicate a concave trajectory for Hb recovery. For $\psi = 1$, Hb recovery is linear. The recovery time may differ between donors due to diet, genetic factors, and other unobserved characteristics. In addition, some donors experience a reduction of their iron reserve after a few donations, which may cause the recovery time to increase with the number of donations. Therefore, the postulated function based on (1) may oversimplify the recovery process. So, we use finite mixture modeling to capture the heterogeneity in recovery time by assuming that donors belong to two different latent classes. Namely, one class of donors is assumed to have a constant recovery time and the other class is assumed to exhibit a non-constant recovery time. The function could be:

$$HRF_{it,g(i),\kappa} = \theta \left[\max \left(\frac{RT_{g(i),\kappa} - (TSPD_{it} - \delta)}{RT_{g(i),\kappa}}, 0 \right) \right]^{\psi},$$

where g(i) is the latent class membership of donor i, which has to be estimated from the data. It is assumed that donors who belong to (latent) class I (g(i) = 1) have a constant recovery time. For donors in class II, it is assumed that the recovery time changes after a certain number of donations (k). This change point k is assumed equal for all people in class II and must be estimated from the data.

3.   Statistical model

We propose a latent class mixed-effect transition model for these data. Formally, let Hb_{it} denote Hb recorded for the ith individual (i=1,2,…,N) at t different times (t=1,2,…,T_i), together with a set of P strictly exogenous covariates. Since the donation intervals are not equal, the data set is unbalanced. Furthermore, there is a reduction in Hb value after donation.

Since the current Hb value is associated with the Hb value observed at the previous visit and the time since previous donation, we need to take into account the time interval since previous donation and the reduction in Hb due to donation. All of these aspects must be incorporated in the statistical model for predicting Hb. Our proposed model is as follows:

$$Hb_{it} = \beta_0 + b_{i1} + \beta_1 Age_{i1} + \beta_2 Season_{it} + \gamma Hb.pd_{it} + HRF_{it,g(i),\kappa} + \varepsilon_{it}, \quad t = 2, ..., T_i$$

where b_{i1} controls the heterogeneity partly explaining the intra-subject correlation, γ is the lagged impact of Hb at the previous donation (Hb.pdit).

One of the assumptions in classical mixed-effects models is that the covariates in the model are exogenous. But this assumption is violated in mixed-effects transition models where one of the covariates is the lag variable, which is endogenous. This issue relates to the initial conditions problem (ICP). Ignoring the ICP results in inconsistent estimates in the model and standard estimation techniques cannot be applied. To handle the ICP in our model, we used a reduced form equation for the initial period similar to the dynamic equation, but excluding the lag effect from the model. We let the model take into account the correction between the unobserved individual effects and the initial state.

4.   Results and conclusion

The latent class mixed-effects transition model was preferred over the simpler models according to the DIC and based on an evaluation of model fit. This finding shows that it is important to account for both unobserved individual heterogeneity and state dependence among the Hb values for an individual. The model results show that there is heterogeneity in recovery time, i.e. 55% and 52% of male and female donors have a constant recovery time during successive donations. The remaining donors have a longer recovery time and their recovery time increases after a number of donations (see Figure 2). This increase in recovery time might be attributed to a reduction of the iron reserves in these donors. In addition, our findings point to a concave Hb recovery process. That is, the recovery process is fastest at the beginning and becomes slower over time, to our knowledge, this fact not been found in previous studies.

Our results support a donation interval longer than 56 days for both sexes. This finding is also supported by previous studies. The U.S. Food and Drug Administration (FDA) is currently considering revising this interval to better protect donors.

In conclusion, we developed a statistical modeling approach that allows classifying donors in two subgroups based on their Hb recovery time. This is of high practical importance because identification of the class for a donor could improve the planning of donors' visits to the blood banks and help to tailor donation intervals and prevent iron deficiency and donor deferrals.

Figure 2. Hb profiles separately for each class obtained from proposed model for both male (upper panels) and female (lower panels) donors. The donors assigned to the highest class probability. The horizontal lines show the threshold for donation eligibility. The arrows show the recovery time change point.