Using the Lasso to simultaneously identify the covariate and variance-covariance structures of nonlinear mixed-effects models

Erik Olofsen
Department of Anesthesiology, Leiden University Medical Center, The Netherlands
email: e.olofsen@lumc.nl

Objectives

Remifentanil is a synthetic opioid with ultrafast elimination. Its pharmacokinetics after a one-minute infusion in children (age range 5 days - 17 years) have been described by Kinder Ross *et al.* using model-independent analysis [1]. The first objective of the present study was to obtain a compartmental model of remifentanil pharmacokinetics with the possible inclusion of covariates age and weight.

The Lasso has been suggested as a method to find the covariate coefficients [2, 3, 4]. Interindividual variabilities of the structural pharmacokinetic parameters with possibly a smaller number of random effects can be described by a (not necessarily square) strength matrix transforming the random effects [5]. The covariate and strength matrix coefficients are likely to be interdependent. Therefore, the second objective of this study was to explore the ability of the Lasso to simultaneously find optimal covariate and variance-covariance coefficients.

Methods

The remifentanil concentration data of study [1] were kindly provided by GlaxoSmithKline. Because of the fast elimination, concentrations were below the lower limit of quantitation within 30 minutes; 196 samples were available in 36 children. Two-compartment models were fit to the data using NONMEM (FOCEI, SIGDIGITS=4); each structural parameter θ_i in $[V_1\ k_{10}\ k_{12}\ k_{21}]^T$ was written as

$$\theta_{ij} = \theta_{\text{pop},i} \cdot \exp(\beta_i^T C + \eta_{ij}), \qquad (1)$$

where j, θ_{pop}, η, β, and C denote individuals, population values, random effects, and vectors of covariate coefficients and covariates, respectively. Age (plus 280 days) and weight were log-transformed and normalized (zero mean and unit variance), denoted by $\tilde{A}$ and $\tilde{W}$, respectively. Because age and weight are highly correlated ($\tilde{W}$ was well predicted by $\tilde{A}$), $\tilde{W}$ and $\tilde{W} - \tilde{A}$ were incorporated as covariates.

A standard model was identified by first finding a diagonal (co)variance matrix Ω of the random effects and subsequently finding covariate coefficients significantly different from zero *via* backward elimination. Akaike's information-theoretic criterion (AIC) was used for model discrimination [6].

The Lasso was implemented by the minimization of a penalized likelihood (PL), given by

$$\text{PL} = \text{-2LL}(\theta_{\text{pop}}, \beta, \gamma, \Lambda, \sigma^2) + \lambda \sum_k |\alpha_k|, \qquad (2)$$

where α denotes the parameters that are subjected to model selection: the larger λ, the more α_k will be zero. LL denotes log-likelihood, σ^2 interindividual variability, and strength matrix Λ was incorporated by substituting (in eq. (1)):

$$\eta = \Lambda(\eta' + \gamma C) \ \text{ with } \ \text{var}(\eta'_l) = 1 \ \ (l = 1, \cdots, 4). \qquad (3)$$

Coefficients γ allow for the incorporation of covariates *via* Λ, which could be efficient. The elements of Λ are not identifiable without the penalty term. There were a total of 32 coefficients α: 16 for covariates and 16 for Λ. At a certain λ, the coefficients different from zero (threshold 0.0001) were identified; subsequently, maximum likelihood estimates of those coefficients and the structural parameters were obtained (the "hybrid" method [7]) and AIC computed. Using a grid search (λ was varied between 0.1 and 100 in 300 increments) the value of λ was located that identified the best model in terms of AIC.

Methods (continued)

Simulations were performed to assess the influence of the covariates and random effects (posterior predictive check [8]) on the pharmacokinetic profile.

Results

The standard model consisted of a diagonal matrix Ω with non-zero elements for V_1 and k_{10}; four covariate coefficients were different from zero. The AIC of this model was 395.2.

The optimal model obtained *via* the Lasso consisted of six coefficients of the strength matrix using two random effects and six covariate coefficients (all $\hat{\gamma} = 0$). The AIC of this model was 390.2. This model was found at 10 of the 300 evaluated values of λ. 161 NONMEM runs were successful; 139 ended with rounding errors. AIC as a function of λ displayed multiple local minima, and the number of coefficients as a function of λ was not monotonically decreasing (see figure 1 below).

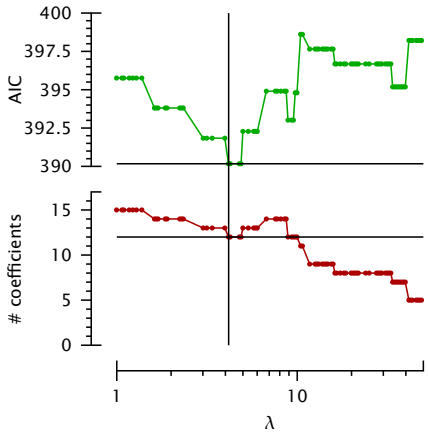

Figure 1: AIC (upper panel) and number of coefficients (lower panel) as a function of λ (see eq. (2)). The cross hairs indicate the optimal value of λ, the corresponding AIC and number of coefficients.

Parameter estimates (SE) of the optimal model were: $\hat{V}_1$ = 1.59 (0.13), $\hat{k}_{10}$ = 0.574 (0.055), $\hat{k}_{12}$ = 0.375 (0.045), and $\hat{k}_{21}$ = 0.187 (0.014). Covariate coefficients estimates $\hat{\beta}_{\tilde{W}}$ were, for the standard and optimal model respectively:

$$\begin{pmatrix} 0.827\ (0.050) \\ 0 \\ -0.160\ (0.067) \\ 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0.828\ (0.046) \\ 0 \\ -0.260\ (0.057) \\ -0.097\ (0.055) \end{pmatrix}.$$

Simulated concentrations from the optimal model, after administration of a dose relative to weight, were monotonically increasing with age; this was not so for simulations from the standard model: it (most probably erroneously) displayed lowest concentrations at about 1 year of age.

A full (co)variance matrix of interindividual variabilities η of the structural parameters can be constructed by calculating $\Omega = \Lambda\Lambda^T$, yielding

$$\begin{pmatrix} 0.114 & & & \\ -0.121 & 0.203 & & \\ -0.0852 & 0.210 & 0.257 & \\ -0.0315 & 0.0778 & 0.0952 & 0.0352 \end{pmatrix}.$$

The first two elements of the diagonal Ω of the standard model were 0.0477 and 0.0462; the remaining were zero.

Results (continued)

Prediction intervals obtained from the optimal model were smaller than those from the standard model, except during a few minutes after infusion, and when extrapolating after the study period of 30 minutes (see figure 2 below). This can be explained by the relatively large covariance between k_{10} and k_{12} incorporated in the optimal model. At the end of the one-minute infusion, the prediction interval from the optimal model was 14.0 – 55.6 ng/ml; with covariates incorporated it reduced to 17.3 – 46.5 ng/ml.

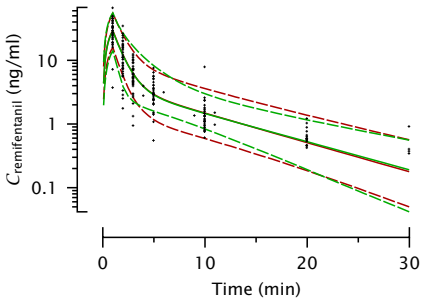

Figure 2: Medians and 95% prediction intervals from the standard model (red) and optimal model (green). The plus signs denote the measured concentrations (LLOQ ≈ 0.5 ng/ml).

Conclusions

With respect to the standard approach, the Lasso provided a better (in terms of AIC) description of the present data; it identified a full variance-covariance matrix Ω with less random effects than structural parameters, and the better characterization of the "unexplained" interindividual variability resulted in improved estimation of the covariate coefficients.

While an exhaustive search of all combinations of 0-fixed and free coefficients is often not feasible, both the methods of forward/backward selection and the Lasso are not guaranteed to find the best model. The (non-continuous) function AIC(λ) displayed multiple local minima, which requires the model to be evaluated at many values of λ. Alternative decompositions of Ω, such as an eigenvalue-eigenvector decomposition (see *e.g.* [9]) might reduce the complexity of AIC(λ).

In the presence of a large number of parameters, the penalty term in eq. (2) has a stabilizing effect on NONMEM's minimization procedure, but it often fails to converge, because at present no recipe exists for eliminating coefficients α_k close to zero.

References

[1] A Kinder Ross, P J Davis, G deL. Dear, B Ginsberg, F X McGowan, R D Stiller, L Graham Henson, C Huffman, and K T Muir. Pharmacokinetics of remifentanil in anesthetized pediatric patients undergoing elective surgery or diagnostic procedures. *Anesth Analg*, 93:1393–401, 2001.

[2] R Tibshirani. Regression shrinkage and selection via the Lasso. *J R Statist Soc*, B 58:267–88, 1996.

[3] J Ribbing, J Nyberg, and E N Jonsson. The LASSO: A novel method for predictive covariate modelling of population pharmacokinetics/pharmacodynamics. Abstract 713 of PAGE meeting (http://www.page-meeting.org), 2005.

[4] J Ribbing, J Nyberg, E O Caster, and E N Jonsson. The LASSO: A novel method for predictive covariate model building in nonlinear mixed effects models. Abstract 936 of PAGE meeting (http://www.page-meeting.org), 2006.

[5] E Olofsen and A Dahan. Population pharmacokinetics/pharmacodynamics of anesthetics. *AAPS Journal*, 7, 2005. Article 39 (http://www.aapsj.org).

[6] K P Burnham and D R Anderson. *Model selection and multimodel inference, 2nd ed.* Springer, New York, 2002.

[7] B Efron, T Hastie, I Johnstone, and R Tibshirani. Least angle regression. *Ann Statist*, 32:407–99, 2004.

[8] Y Yano, S L Beal, and L B Sheiner. Evaluating pharmacokinetic/pharmacodynamic models using the posterior predictive check. *J Pharmacokin Pharmacodyn*, 28:171–192, 2001.

[9] J C Pinheiro and D M Bates. *Mixed-effects models in S and S-PLUS.* Springer, New York, 2000.