Haemophilia
The Official Journal of the World Federation of Hemophilia,
European Association for Haemophilia and Allied Disorders and
the Hemostasis & Thrombosis Research Society

Validation of the prediction model for inhibitor development in PUPs with severe haemophilia A

S. M. HASHEMI,* K. FISCHER,*† K. G. M. MOONS* and H. M. VAN DEN BERG* ON BEHALF OF THE PedNet STUDY GROUP
*Julius Center for Health Sciences and Primary Care University Medical Center; and †Van Creveldkliniek University Medical Center, Utrecht, the Netherlands

The development of inhibitory alloantibodies against exogenous factor VIII (FVIII) is the most serious complication of haemophilia A treatment [1]. Inhibitor development usually occurs after 10–14 exposure days with FVIII in 25–30% of previously untreated patients (PUPs) with severe haemophilia A [2,3]. Patients with inhibitors suffer significant morbidity and reduced quality of life.

Genetic factors associated with the risk of inhibitor development in severe haemophilia A are *F8* gene mutation type, a positive family history for inhibitors, ethnicity and some polymorphisms of immune response genes [4]. Several studies have demonstrated that the risk for inhibitor development is also influenced by treatment-related or nongenetic factors. These include age at first exposure and presence and intensity of peak moments in treatment, such as surgery or treatment of major bleeding episodes and FVIII product type [2–4].

If prediction models can identify patients with a high-risk of inhibitor development at the start of treatment, alternative treatment strategies can be implemented to reduce the risk [5]. Ter Avest *et al.* [6] published a risk stratification score for inhibitor development based on a positive family history of inhibitors, high-risk *F8* gene mutation type and intensive treatment at first exposure. Before application in practice, prediction models have to be validated in another cohort to establish their predictive performance, since the performance estimated from the data from which the model was derived is generally overestimated [7]. The aim of this study was therefore to validate the previously developed prediction model in a large independent dataset of PUPs with severe haemophilia A.

The prediction model was developed in a retrospective, multicentre cohort of 332 PUPs with severe haemophilia A (residual FVIII activity <0.01 IU mL^{-1}), born between 1990 and 2000 with details of all initial 50 exposures to FVIII [3]. Our validation dataset contained 503 PUPs with severe haemophilia A born between 01 January 2000 and 31 December 2007, included from the PedNet Registry [8]. Data on all predictors and on treatment in the first 50 EDs were available. Both studies used the same outcome definition for clinically relevant inhibitor development, i.e. at least two positive inhibitor titres according to the local laboratory and a FVIII recovery of less than 66% of expected [8].

The three predictors were coded according to the original prediction model [6]. *Family history of inhibitors* was defined as being positive or negative. Patients with a negative family history of haemophilia A were included in the category of negative family history of inhibitors. *F8 gene mutation type* was divided into high- and low-risk gene mutation types. High-risk mutations included large deletions (>200 base pairs), nonsense mutations and intron 22 or 1 inversions. Low-risk mutations included all other mutations. *Intensive treatment at first treatment* was defined as at least five consecutive exposure days (ED: treatment with a FVIII product within 24 h) from the first ED onwards.

All patients had complete follow-up data until 50 EDs or until inhibitor development and limited missing data for *family history of inhibitors* (10.3%), *F8 gene mutation type* (12.1%) and *intensive treatment at first treatment* (1.8%). However, to reduce bias, we imputed the missing data of these predictors.

Predictive performance in the validation dataset was estimated by the model's discrimination (using the C-index) and calibration (using the calibration plot) [9–10]. Subsequently, predicted risks were compared to the observed risks, after using the predicted risk categories proposed in the development study (≤0.09 as low risk and ≥0.36 as high risk) [6]. Finally, for comparison purposes, we fitted the original model again in the validation dataset, using multivariable logistic regression.

Statistical analyses were performed using IBM SPSS Statistics 22 (SPSS, Inc., Chicago, IL, USA) and R, version 3.0.1 (http://cran.r-project.org/bin/windows/base).

Table 1 shows the patient and treatment characteristics in the derivation and the validation datasets. In the validation dataset, clinically relevant inhibitor development occurred in 151 (30%) out of 503 PUPs compared to 83 (25%) out of 332 in the derivation dataset. In the validation dataset, the C-index of the model was 0.65 (CI 0.60–0.70) compared to 0.74 (CI

Correspondence: H. Marijke van den Berg, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Huispost STR. 6.131, P.O. Box 85500, 3508 GA, Utrecht, the Netherlands.
Tel.: 0031 88 7568189; fax: +31 88 75593;
e-mail: H.M.vandeBerg@umcutrecht.nl

Accepted after revision 6 December 2015

DOI: 10.1111/hae.12895

0.67–0.80) in the derivation dataset. The calibration plot in Fig. 1 shows an underestimation of inhibitor development in the low-risk category and overestimation of inhibitor development in the high-risk categories. Identification of high- and low-risk patients was less efficient in the validation dataset: predicted inhibitor risks were 0.13 in the low-risk category (compared to 0.06 in the derivation dataset) and 0.45 in the high-risk category (compared to 0.57 in the derivation dataset) [6]. This reflects the results of the calibration plot (Fig. 1): in the validation dataset, the original prediction model predicted less patients in the low-risk category (underestimation) and more patients in the high-risk category (overestimation). Still the prediction model was able to predict 45% of the patients that will likely develop inhibitors at initial treatment, compared to the *a priori* risk of 30%.

When refitting the prediction model in the validation dataset considerable shifts in the ORs were observed: from OR 3.7 (CI 1.4–9.2) to 1.7 (CI 1.0–3.3) for *family history of inhibitors*, from OR 7.7 (CI 3.8–15.2) to 2.8 (CI 1.7–4.7) for *intensive treatment at first treatment* and from OR 3.5 (CI 1.8–6.9) to 2.9 (CI 1.9–4.6) for *F8 gene mutation type*.

The results in our validated prediction model confirm the importance of external validation. The original prediction model was validated in a small and highly selected cohort of only 64 PUPs [6]. These PUPs were recruited from different PUP registration studies and were the only patients that had data of gene defects available. It can be hypothesized that testing for gene defects was selected for children with a positive family history for inhibitors; 55% of the 64 PUPs had a positive family history for inhibitors, compared to 7.5% in the derivation dataset.

The original prediction model had good performance (C-index 0.74) and was developed to predict the risk at initial treatment [6]. However, it showed decreased performance (C-index 0.65) in the validation dataset; this observation is common and an important reason for external validation [7, 10].

We tried to find explanations for this decrease in predictive accuracy by exploring the contributions of individual predictors. The lower performance of the prediction model was likely caused by shifts in the odds ratios of the predictors' *family history of inhibitors* and *intensive treatment at first exposure*. For the predictor *intensive treatment* (OR 7.7 in the derivation dataset vs. 2.8 in the validation dataset) the shift might be explained by changes in treatment since 2000. Data collection in patients born after 2010 is ongoing and may confirm this hypothesis.

In conclusion, validation of the original prediction model for inhibitor development by ter Avest *et al.* in an external dataset of 503 PUPs with severe haemophilia A showed a worse discrimination and calibration. However, the independent association of the three predictors *high-risk F8 gene mutation type*, *positive family history for inhibitors* and *intensive treatment at first exposure* remained.

Table 1. Patient characteristics of PUPs in the validation and derivation datasets.

	Derivation dataset 1990–1999	Validation dataset 2000–2007
Entire cohort	*n* = 332	*n* = 503
No. of inhibitors (%)	83 (25)	151 (30)
Caucasian ethnicity (%)	89	88
Age at diagnosis, months		
Median (IQR)	6.7 (0.1–10.8)	6.3 (0.09–11.0)
Age at first exposure, months		
Median (IQR)	10.5 (6.2–14.7)	10.1 (6.2–13.8)
Family history of inhibitors		
Negative/NA (%)	308 (92.8)	455 (90.5)
Positive (%)	24 (7.2)	48 (9.5)
Gene mutation type		
Low risk (%)	118 (34.3)	187 (37.2)
High risk (%)	218 (65.7)	316 (62.8)
Intensive treatment (5 EDs)		
No (%)	281 (84.6)	427 (84.9)
Yes (%)	51 (15.4)	76 (15.1)
Inhibitor patients	*n* = 83	*n* = 151
Age at inhibitor development, months		
Median (IQR)	15.3 (10.4–21.9)	16.4 (11.7–20.2)
ED at inhibitor development, months		
Median (IQR)	14 (8–21)	14 (10–22)
Duration to inhibitor development, months		
Median (IQR)	6.1 (2.0–12.1)	4.6 (1.9–8.5)

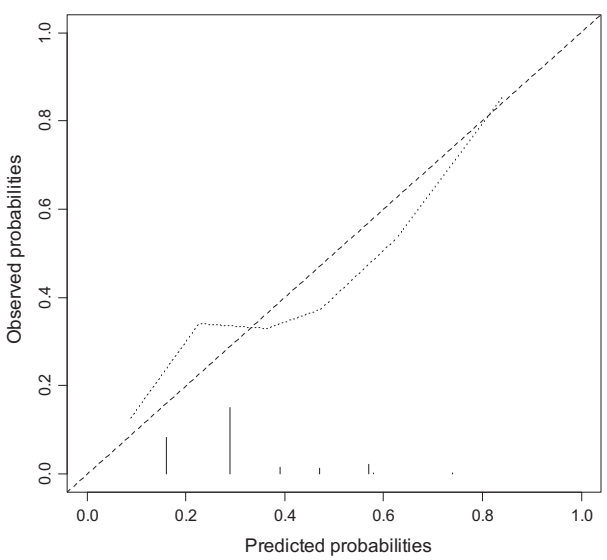

Fig. 1. Calibration plot of the prediction model in the validation dataset. The calibration plot of the prediction model when applied to the validation dataset. The ideal calibration plot is the diagonal line where predicted probabilities are equal to the observed probabilities. The plot shows an underestimation of inhibitor risk in the low-risk category (<0.25) by the original model; the observed probabilities are higher in the validation dataset. The plot shows an overestimation of inhibitor risk in the high-risk category (>0.50) by the original model; the observed probabilities are lower in the validation dataset.

Acknowledgements

The authors thank the study coordinator, E. Smink and L. Blijboom, for their support in data cleaning; the PedNet study group and the regional PedNet Study coordinators Kate Khair, Karin Lindvall and Yves Guillaume for their support of the study.

Author contributions

All members of the PedNet Registry and RODIN Study participated in designing the research and collected patient data. S.M. Hashemi analyzed and interpreted the data and wrote the first draft of the paper. K. Fischer, K.G.M. Moons conceived, designed and supervised the research and contributed to writing the paper. K.G.M. Moons and K. Fischer supervised the statistical analyses. All authors critically reviewed the final version of the paper.

Disclosure

S.M. Hashemi, K. Fischer and H.M. van den Berg have received unrestricted research funding for this project at the Van Creveldkliniek and the Julius Center for Health Sciences from Baxter Healthcare Cooperation. K. Fischer has received speaker's fees from Baxter, CSL Behring, Pfizer, NovoNordisk and Biotest; performed consultancy for Bayer, Baxter, Biogen, NovoNordisk and Pfizer; and has received research support from Bayer, Pfizer, Baxter and NovoNordisk. H.M. van den Berg has received unrestricted research support from CSL Behring, Novo Nordisk, Sobi Biogen, Wyeth, Baxter and Bayer. K.G.M. Moons receives funding from the Netherlands Organization for Scientific Research (project 9120.8004 and 918.10.615). The authors have no competing interests. The PedNet Study Group and the CANAL Study were funded by unrestricted research grants from Bayer and Baxter Healthcare Cooperation.

References

1 Mannucci PM, Tuddenham EG. The hemophilias: from royal genes to gene therapy. *N Engl J Med* 2001; **344**: 1773–9.
2 Gouw SC, van der Bom JG, Ljung R *et al.* PedNet and RODIN Study Group. Factor FVIII products and inhibitor development in severe hemophilia A. *N Engl J Med* 2013; **368**: 231–9.
3 Gouw SC, van der Bom JG, van den Berg H. Treatment-related risk factors of inhibitor development in previously untreated patients with hemophilia A: the CANAL cohort study. *Blood* 2007; **109**: 4648–54.
4 Astermark J. FVIII inhibitors: pathogenesis and avoidance. *Blood* 2015; **125**: 2045–51.
5 Astermark J. Prevention and prediction of inhibitor risk. *Haemophilia* 2012; **18**: 38–42.
6 Ter Avest PC, Fischer K, Mancuso ME *et al.* Risk stratification for inhibitor development at first treatment for severe hemophilia A: a tool for clinical practice. *J Thromb Haemost* 2008; **6**: 2048–54.
7 Moons KG, Kengne AP, Grobbee DE *et al.* Risk prediction models: II. External validation, model updating, and impact assessment. *Heart* 2012; **98**: 691–8.
8 Fischer K, Ljung R, Platokouki H *et al.* Prospective observational cohort studies for studying rare diseases: the European PedNet Haemophilia Registry. *Haemophilia* 2014; **20**: e280–6.
9 Harrell FE. Regression Model Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York: Springer; corrected edition, 2001.
10 Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009; **338**: b605.

