Prediction models for hemoglobin deferral in whole blood donors

A.M. Baart

ISBN 978-94-6191-763-8

Design cover: Promotie In Zicht, Arnhem
Lay-out: Promotie In Zicht, Arnhem
Print: Ipskamp Drukkers, Nijmegen

The research described in this thesis was funded by
Sanquin Blood Supply under grant number 06-007.

Financial support for the publication of this thesis
was provided by HemoCue Diagnostics B.V.

Prediction models for hemoglobin deferral in whole blood donors

Predictiemodellen voor afkeuring vanwege een te laag hemoglobinegehalte bij volbloeddonors

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op dinsdag 25 juni 2013 des ochtends te 10.30 uur

door

Alexandra Mireille Baart
geboren op 8 september 1981
te Rotterdam

Promotor:

Prof.dr. K.G.M. Moons

Co-promotoren:

Dr. Y. Vergouwe

Dr. W.L.A.M. de Kort

Contents

Chapter 1

General introduction

Blood donation

Blood donors help to save millions of lives each year by donating either whole blood or plasma. Although blood donations and subsequent transfusions are meant to help other people, they are associated with several risks, for both the recipients and for the donors. For example, recipients may contract an infectious disease and donors may develop iron deficiency. To secure blood safety and donor health, several safety measures exist and donors have to fulfil various eligibility criteria for donation. In the Netherlands, donors must be between 18 and 70 years old, and prior to each blood donation donors have to fill out an eligibility questionnaire to identify known medical conditions and perilous behaviour. In a physical examination, body weight, pulse rate, blood pressure and the hemoglobin (Hb) level are measured. According to European guidelines,[1] donors should have a body weight of at least 50 kg, a regular pulse, a systolic blood pressure between 90 and 180 mmHg, a diastolic blood pressure between 50 and 100 mmHg, and an Hb level of at least 8.4 mmol/l (13.5 g/dl) for men and 7.8 mmol/l (12.5 g/dl) for women. Furthermore, the minimum time interval between two whole blood donations must be 56 days, and the maximum number of whole blood donations allowed per year is five for men and three for women.
In the Netherlands, blood donation is organized by the Sanquin Blood Supply Foundation. In 2011, A total of 398,397 blood donors were registered at Sanquin. These donors provided 538,282 whole blood donations and 347,554 plasma donations in 2011.[2]

Effect of whole blood donation on the iron status

Blood donation poses a risk of iron deficiency to blood donors. Iron is an important element of the Hb protein, which is found in red blood cells. In humans, most iron is found in red blood cells, incorporated in Hb. Red blood cells contain about 60-75% of the total body iron. In adults, the total body iron content is normally 3-5 g with typically higher values in men than in women.[3] With a blood donation, a substantial amount of iron is lost. A donation of 500 ml whole blood contains about 200-250 mg iron,[4] which is 4-8% of total body iron. When the iron intake is not sufficient to replenish the iron loss due to donation, a negative iron balance occurs. Subsequent blood donations may then gradually lead to iron deficiency.
Three stages of iron deficiency can be distinguished: iron depletion, iron deficient erythropoiesis and iron deficiency anemia.[5] Iron depletion is marked by running out iron stores. In the stage of iron deficient erythropoiesis, the iron supply to the erythropoietic bone marrow is becoming insufficient for erythropoiesis. However, Hb levels are still normal. When finally the iron supply becomes insufficient to produce a normal amount of Hb, iron deficiency anemia becomes apparent. Iron depletion and iron deficient erythropoiesis are thus sub-clinical stages; with anemia clinical symptoms appear.

The primary function of Hb is oxygen transport through the bloodstream from the lungs to all other tissues in the body, and one of the first symptoms of anemia is decreased fitness through a diminished oxygen supply to the body tissues. Furthermore, as iron is also an important element of several other proteins, iron deficiency also affect DNA synthesis,[6,7] the immune system,[8] and energy metabolism through impaired mitochondrial electron transport.[9,10]

Blood donors on average need several weeks to replenish the lost iron after a blood donation.[11] However, there are wide variations in the duration of the recovery period among individual donors. The European guidelines with relation to a minimum donation interval and a maximum number of donations per year may therefore not be safe for each individual donor. Indeed, depleted iron stores in blood donors are not uncommon[4,5,12] and also iron deficient erythropoiesis occurs.[13,14]

Hb deferral

To protect donors from developing iron deficiency anemia after blood donations, the iron status of blood donors is assessed prior to donation. Most commonly, this is done by measuring Hb levels. Donors with low Hb levels are deferred from donation to prevent anemia afterwards. Furthermore, deferral of donors with low Hb levels also ensures that blood units for transfusion meet the required standards for Hb content.[15] Hb cutoff levels for donation described in the European Commission Directive[1] are 8.4 mmol/l (13.5 g/dl) for men and 7.8 mmol/l (12.5 g/dl) for women.

A substantial number of donors is deferred from donation because of low Hb levels. In 2011, Dutch male donors were 12,583 (2.2%) times deferred because of a low Hb level and for Dutch female donors this number was 23,768 (5.5%). Although deferrals are meant to protect donors, they are also demoralizing for donors. As a consequence, the risk of donor lapse is increased due to deferral,[16-19] although the donor's Hb level could actually be sufficient at the time of the next invitation to donate.

Timely estimating the risk of Hb deferral in blood donors, i.e. before being invited, could be helpful in the management of the donation program and the retention of donors. At the individual level, such predictions of Hb deferral risk may guide the decision whether a donor can be invited for the next donation, or whether it is better to postpone the invitation. From a management perspective, these predictions may decrease the number of donor deferrals for low Hb levels.

Factors associated with Hb deferral

Several factors are known to be associated with low Hb levels or Hb deferral. Demographic characteristics such as sex and age are associated with Hb levels. Hb levels rise substantially during childhood. In men, there is a small decrease in Hb levels with increasing age,

whereas in women, Hb levels rise by the effect of menopause due to hormonal changes and the cessation of iron loss through menstruation.[20] Despite lower Hb cutoff levels for donation for women in most countries, Hb deferral occurs more frequently in women than in men.

Body mass index (BMI) is also associated with Hb levels: a greater BMI is associated with higher Hb levels.[21] Likewise, blood volume might be associated with Hb deferral in blood donors. The amount of blood given with a whole blood donation is around 500 ml. Donors with a large blood volume lose relatively less blood with a blood donation compared to donors with a small blood volume. Therefore, it is likely that donors with a large blood volume need less time to recover after a blood donation and have a smaller risk of Hb deferral at the next visit to the blood collection center.

Another factor that is associated with Hb deferral is seasonality. Hb levels decrease with increasing daily temperature and are thus lower in warmer seasons. Consequently, in summer months deferral rates are higher.[22,23]

Furthermore, specific characteristics of the donation history might be associated with Hb deferral. Hb levels measured at previous visits to the blood collection center are obviously associated with current Hb levels and previous Hb deferral is likely associated with Hb deferral at a next visit. The longer the time interval between two donations, the more time for the donor to recover from the previous donation; thus time interval between two subsequent donations or visits is also associated with Hb deferral.

Studies have shown that an increased donation frequency is associated with lower iron stores.[4,12,24] With advancing iron deficiency low Hb levels follow iron depletion. Consequently, an increased donation frequency might also be associated with Hb deferral. Donors may switch from donating whole blood to donating plasma only and vice versa. With a plasma donation, most of the red blood cells are returned to the donor and thus the effect of such a donation on the donor's iron status is small. A history of plasma donation rather than whole blood donation might therefore be inversely associated with Hb deferral.

Finally, values of other iron parameters in blood might be associated with Hb deferral. Hb levels are only low in an advanced stage of iron deficiency. As a consequence, it may occur that donors pass the Hb screening test while they have already depleted iron stores or even iron deficient erythropoiesis,[4,5,12-14] as these conditions remain undetected with Hb screening. Especially these donors are at high risk of developing iron deficiency anemia after a blood donation and are more likely to be deferred at their next visit to the blood collection center. Iron parameters that respond in an early stage to a low iron status may therefore be predictive for Hb deferral. There are several iron parameters available to assess iron depletion or iron deficient erythropoiesis. Iron depletion can be assessed by measuring serum ferritin levels.[5,25] Tests for the diagnosis of iron deficient erythropoiesis include measurement of plasma iron,[5] total iron binding capacity or transferrin,[5] transferrin saturation, soluble transferrin receptor (sTfR) concentration,[26] the sTfR index (sTfR divided by log-transformed ferritin values)[27,28] and zinc protoporphyrin (ZPP).[29] Another iron

parameter with which sub-clinical iron deficiency can be assessed is the recently discovered iron regulatory protein hepcidin.[30,31] Each of the above mentioned tests has its own advantages and disadvantages. ZPP is measured by an automated technology and its attractive features are its ability to perform immediate point-of-care assays and its relative low price. It may therefore especially be useful for donor screening.
All these factors might be useful to develop so-called multivariable prediction models for Hb deferral in blood donors.

Prediction modeling

In general, prediction models are useful tools to estimate the risk that a certain disease or outcome is present (diagnosis) or will occur in the future (prognosis). The model predictions can be used to identify individuals that are at risk for a certain outcome, to assist in clinical decisions such as applying treatment or an intervention, or to inform patients and their relatives about the course of their disease.[32] In blood bank practice, prediction models for Hb levels or Hb deferral might be used to identify donors at high risk of Hb deferral. The model predictions may be helpful to decide for each individual donor at the moment of possible invitation for a donation whether they can indeed be invited. Donors with a low predicted risk can be invited with preference, whereas for those with a high risk of Hb deferral the invitation may be postponed. Moreover, high risk individuals may even be adviced to undergo interventions such as an iron rich diet. Accordingly, prediction of Hb deferral may decrease the number of donor deferrals for low Hb levels, and subsequently increase donor satisfaction.

Prediction models are usually developed with regression analysis techniques. From a set of candidate predictors, the strongest predictors are selected in a so-called final prediction model.
Once such prediction model is developed, the performance of the model needs to be examined (validation). Overall performance measures are related to goodness-of-fit, which relates the ability of a model to fit a given set of data. Two key aspects that are often examined in the validation procedure are calibration and discrimination. Calibration is the agreement between predicted probabilities and observed frequencies. Discrimination refers to the ability of the model to differentiate between patients with the outcome (Hb deferral) and patients without the outcome. Calibration and discrimination measures are mere statistical performance measures. Ideally, the performance of a prediction model is validated externally, i.e. tested on its predictive ability in some other population than (external to) the study population that was used for model development.[33-37]
If the model performance is poor in external validation, the model needs to be updated by adjusting it with the new data of the validation population. Also, the added value of

one or more new predictors to the existing model can immediately be examined in such external validation study. Updating the existing model is preferred over fitting a new model, because with model updating the information captured in the original model development study and that of the validation study are combined, resulting in a more robust and thus generalisable model.[38,39]
After external validation and, if necessary, updating of a prediction model, the clinical impact of using the model should be assessed. This can be done prospectively in a (cluster) randomized trial, but also using decision analytical or decision modeling techniques. In such impact studies, the effect of using the model on decision making by health professionals, patient outcome, and costs can be compared with a setting in which the prediction model is not used. If such studies show adequate results, the prediction model may be ready to be implemented in practice.[40,41]

Outline of this thesis

This thesis presents various studies on prediction models for Hb deferral in whole blood donors. Several studies on predictive factors for Hb deferral have been published before.[14,42,43] However, in earlier studies no advanced statistical methods were used for model development, nor were developed prediction models formally validated externally. In the studies presented in this thesis, prediction models are formally developed, validated and updated with sophisticated statistical techniques.
Chapter 2 describes the development of an initial prediction model for Hb deferral. This model was developed in a sample of Dutch - male and female - whole blood donors. Thereafter, we developed and internally validated sex-specific prediction models in a large cohort consisting of all Dutch whole blood donors. This study is described in chapter 3. Subsequently, the sex-specific prediction models were externally validated and updated in a cohort of Irish whole blood donors. Results of this study are presented in chapter 4.
As mentioned, depleted iron stores and iron deficient erythropoiesis, early stages of iron deficiency with still normal Hb levels, are common in blood donors. Chapter 5 describes a study in which the prevalence of this so-called sub-clinical iron deficiency was examined in Dutch whole blood donors that were not deferred for low Hb levels. The prevalence of sub-clinical iron deficiency was based on levels of the iron status parameter zinc protoporphyrin (ZPP). Also, the distribution of other iron parameters was assessed. Following this study, the added value of ZPP levels to the sex-specific prediction models from chapter 3 and 4 was investigated in chapter 6.
The prediction models presented in chapters 2-4 and chapter 6 are logistic regression models with the dichotomous outcome Hb deferral yes/no. For the development of the prediction models, Hb levels were dichotomized at the sex-specific cutoff level for donation. With these prediction models, the risk of Hb deferral can be calculated. In the

final study of this thesis, sex-specific prediction models were developed to predict continuous Hb levels at the next visit. Comparison of the predicted Hb level with the required Hb cutoff level for donation provides information on the risk of Hb deferral. The performance of these linear models was compared with the logistic models. In addition, the longitudinal aspects of the donation history was accounted for in these linear regression models. The results of this study are presented in chapter 7.
This thesis ends with a summary, clinical implications, and perspectives for further research in chapter 8.

Reference List

1. European Commissian Directive 2004/33/EC. 2004.
2. Sanquin Annual Report 2011.
3. Crichton RR. Iron metabolism. [3rd edition]. Chichester, UK: John Wiley & Sons; 2009.
4. Finch CA, Cook JD, Labbe RF, Culala M. Effect of blood donation on iron stores as evaluated by serum ferritin. Blood 1977;50:441-7.
5. Skikne B, Lynch S, Borek D, Cook J. Iron and blood donation. Clin.Haematol. 1984;13:271-87.
6. Hoffbrand AV, Ganeshaguru K, Hooton JW, Tattersall MH. Effect of iron deficiency and desferrioxamine on DNA synthesis in human cells. Br.J.Haematol. 1976;33:517-26.
7. Furukawa T, Naitoh Y, Kohno H, Tokunaga R, Taketani S. Iron deprivation decreases ribonucleotide reductase activity and DNA synthesis. Life Sci. 1992;50:2059-65.
8. Ward RJ, Crichton RR, Taylor DL, Della CL, Srai SK, Dexter DT. Iron and the immune system. J.Neural Transm. 2011;118:315-28.
9. Finch CA, Miller LR, Inamdar AR, Person R, Seiler K, Mackler B. Iron deficiency in the rat. Physiological and biochemical studies of muscle dysfunction. J.Clin.Invest 1976;58:447-53.
10. Ackrell BA, Maguire JJ, Dallman PR, Kearney EB. Effect of iron deficiency on succinate- and NADH-ubiquinone oxidoreductases in skeletal muscle mitochondria. J.Biol.Chem. 1984;259:10053-9.
11. Fowler WM, Barer AP. Rate of hemoglobin regeneration in blood donors. JAMA 1942;118:421-7.
12. Simon TL, Garry PJ, Hooper EM. Iron stores in blood donors. JAMA 1981;245:2038-43.
13. Cable RG, Glynn SA, Kiss JE, Mast AE, Steele WR, Murphy EL, Wright DJ, Sacher RA, Gottschall JL, Vij V, et al. Iron deficiency in blood donors: analysis of enrollment data from the REDS-II Donor Iron Status Evaluation (RISE) study. Transfusion 2011;51:511-22.
14. Cable RG, Glynn SA, Kiss JE, Mast AE, Steele WR, Murphy EL, Wright DJ, Sacher RA, Gottschall JL, Tobler LH, et al. Iron deficiency in blood donors: the REDS-II Donor Iron Status Evaluation (RISE) study. Transfusion 2012;52:702-11.
15. European Directorate for the Quality of Medicines and HealthCare. Guide to the preparation, use and quality assurance of blood components, European Committee (partial agreement) on Blood Transfusion (CD-P-TS), Recommendation No. R(95) 15, 15th edition. Strasbourg: Council of Europe; 2009, p 1899.
16. Halperin D, Baetens J, Newman B. The effect of short-term, temporary deferral on future blood donation. Transfusion 1998;38:181-3.
17. Custer B, Chinn A, Hirschler NV, Busch MP, Murphy EL. The consequences of temporary deferral on future whole blood donation. Transfusion 2007;47:1514-23.
18. Zou S, Musavi F, Notari EP, Rios JA, Trouern-Trend J, Fang CT. Donor deferral and resulting donor loss at the American Red Cross Blood Services, 2001 through 2006. Transfusion 2008;48:2531-9.
19. Hillgrove T, Moore V, Doherty K, Ryan P. The impact of temporary deferral due to low hemoglobin: future return, time to return, and frequency of subsequent donation. Transfusion 2011;51:539-47.
20. Yip R, Johnson C, Dallman PR. Age-related changes in laboratory values used in the diagnosis of anemia and iron deficiency. Am.J.Clin.Nutr. 1984;39:427-36.
21. Micozzi MS, Albanes D, Stevens RG. Relation of body size and composition to clinical biochemical and hematologic indices in US men and women. Am.J.Clin.Nutr. 1989;50:1276-81.
22. Sebok MA, Notari EP, Chambers LA, Benjamin RJ, Eder AF. Seasonal temperature variation and the rate of donor deferral for low hematocrit in the American Red Cross. Transfusion 2007;47:890-4.
23. Hoekstra T, Veldhuizen I, van Noord PA, de Kort WL. Seasonal influences on hemoglobin levels and deferral rates in whole-blood and plasma donors. Transfusion 2007;47:895-900.
24. Garry PJ, Koehler KM, Simon TL. Iron stores and iron absorption: effects of repeated blood donations. Am.J.Clin.Nutr. 1995;62:611-20.
25. Cook JD. Clinical evaluation of iron deficiency. Semin.Hematol. 1982;19:6-18.
26. Kohgo Y, Niitsu Y, Kondo H, Kato J, Tsushima N, Sasaki K, Hirayama M, Numata T, Nishisato T, Urushizaki I. Serum transferrin receptor as a new index of erythropoiesis. Blood 1987;70:1955-8.
27. Punnonen K, Irjala K, Rajamaki A. Serum transferrin receptor and its ratio to serum ferritin in the diagnosis of iron deficiency. Blood 1997;89:1052-7.

28. Beguin Y. Soluble transferrin receptor for the evaluation of erythropoiesis and iron status. Clin.Chim.Acta 2003;329:9-22.
29. Labbe RF, Rettmer RL. Zinc protoporphyrin: a product of iron-deficient erythropoiesis. Semin.Hematol. 1989;26:40-6.
30. Nemeth E, Ganz T. The role of hepcidin in iron metabolism. Acta Haematol. 2009;122:78-86.
31. Kemna EH, Tjalsma H, Willems HL, Swinkels DW. Hepcidin: from discovery to differential diagnosis. Haematologica 2008;93:90-7.
32. Steyerberg EW. Clinical prediction models. New York: Springer; 2009.
33. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat.Med. 1996;15:361-87.
34. Harrell FE, Jr. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer; 2001.
35. Royston P, Moons KG, Altman DG, Vergouwe Y. Prognosis and prognostic research: Developing a prognostic model. BMJ 2009;338:b604.
36. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann.Intern.Med. 1999;130:515-24.
37. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.
38. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Stat.Med. 2004;23:2567-86.
39. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. J.Clin.Epidemiol. 2008;61:1085-94.
40. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann.Intern.Med. 2006;144:201-9.
41. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606.
42. Mast AE, Schlumpf KS, Wright DJ, Custer B, Spencer B, Murphy EL, Simon TL. Demographic correlates of low hemoglobin deferral among prospective whole blood donors. Transfusion 2010;50:1794-802.
43. Oliveira CD, Martins G, Custer B, Proietti FA, Carneiro-Proietti AB, Cesar CC. Hierarchical analysis of anaemia deferral in blood donor candidates: the individual in the population perspective. Transfus.Med. 2011;21:371-7.

Chapter 2

Prediction of low hemoglobin levels in whole blood donors

A. Mireille Baart*†, Wim L.A.M. de Kort*, Karel G.M. Moons†, Yvonne Vergouwe†

* Sanquin Research, dept. Donor Studies, Nijmegen, the Netherlands

† Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands

Published in Vox Sanguinis, 2011;100:204-211

Abstract

Background: Each year, a relevant proportion of whole blood donors is deferred from donation because of low hemoglobin (Hb) levels. Such temporary deferrals are demoralizing, and donors may never return for a donation. Reliable predictions of Hb levels may guide the decision whether donors can be invited for the next donation. In this study, a prediction model was developed for the risk of low Hb levels.

Study design and methods: Individual data from 5191 whole blood donors were analyzed; 143 donors had a low Hb level. Eleven candidate predictors were considered in logistic regression models to predict low Hb levels. The performance of the prediction model was studied with the Receiver Operating Characteristic (ROC) curve. Internal validity was assessed with a bootstrap procedure.

Results: Strong predictors were sex, seasonality, Hb level measured at the previous visit, difference in Hb levels between the previous two visits, time since the previous visit, deferral at the previous visit, and the total number of whole blood donations in the past two years. Internal validation showed an area under the ROC curve of 0.87.

Conclusion: The developed prediction model provides accurate discrimination between donors with low and appropriate Hb levels. The model predictions may be valuable to determine whether donors can be invited for a next donation, or whether some interventions such as postponement of the invitation, are warranted. Potentially, this could decrease the number of donor deferrals for low Hb levels.

Introduction

The policy of many blood establishments include the assessment of hemoglobin (Hb) levels in blood donors prior to donation. Donors with low Hb levels are deferred from donation to protect them from being critically anemized after a donation. Deferral also ensures that blood units for transfusion meet the required standards for Hb content.[1] An Hb level of 8.4 mmol/L (13.5 g/dL) for men and 7.8 mmol/L (12.5 g/dL) for women is widely accepted as the cutoff level for donation.[2]

In 2009, 15,204 out of 598,869 (2.5%) health assessments in Dutch male donors and 25,605 out of 418,186 (6.1%) health assessments in Dutch female donors led to a deferral because of low Hb levels. Deferrals are demoralizing for donors. Furthermore, the risk that donors may never return for a donation is increased,[3,4] although the donors' Hb level could actually be high enough at the time of the next donation invitation.

Several factors are known to be associated with Hb levels and might be used as predictors for low Hb levels, such as sex,[5] age,[5] body mass index (BMI)[6,7] and seasonality.[8] In addition, characteristics of the donation history may be predictive, such as the total number of blood donations given in the past years, Hb levels measured at previous visits, the time interval between the present visit and the previous visit to the blood collection center, and previous deferral.

The aim of this study was to develop and internally validate a prediction model for the presence of low Hb levels in whole blood donors. The model predictions may be used to identify donors with a relatively low risk of low Hb levels. These donors can be invited for the next planned donation with preference.

Materials and methods

Donors

We analyzed data of donors who had visited one of seven blood collection centers in the Sanquin Blood Bank Southeast Region in the Netherlands in 2004. The data were recorded in an anonymized database covering five years (2000 - 2004) of donations. If donors had visited a blood collection center more than once in 2004, one visit was selected as the "intended visit". Selection was random in order to mimic the seasonal distribution. Donors should have donated whole blood at least twice prior to the intended visit within a period of four years. This was required to collect sufficient information on donation history. Donors may change from donating whole blood to plasma, and this may influence Hb levels. As the aim of this study was to develop a prediction model for whole blood donors, the last two donations preceding the intended visit had to be whole blood donations. Finally, donors should have visited the blood collection center in the period between April 28th and June 6th 2005. In this period, body height was collected on top of the routine

health assessment. These data were necessary to determine two of the candidate predictors (body mass index (BMI) and blood volume, see below).
Donors were excluded if occurrence of donation was uncertain at a visit in the preceding four years before the intended visit. Donors were also excluded if the Hb level at the intended visit was unknown.

Outcome variable and candidate predictors

We predicted whether the Hb level measured in the blood at the intended visit was too low, which subsequently resulted in deferral for donation. The cutoff values for donor deferral were 8.4 mmol/L for men and 7.8 mmol/L for women.[2] Hb levels below these sex specific cutoff values were defined as low in this study. Hb levels at or greater than these values were defined as appropriate. Hb levels were determined in finger stick capillary samples using a photometer (HemoCue, Angelholm, Sweden).
The following eleven candidate predictors of low Hb level versus appropriate Hb level were studied: sex, age, BMI, blood volume, seasonality, Hb level measured at the previous visit (previous Hb level), difference in Hb levels between the previous two visits (delta Hb), time since the previous visit, total number of whole blood donations in the past two years, deferral at the previous visit, and a history of plasma donation in the past two years (yes/no).
BMI and blood volume were calculated from data on height and total body weight. BMI was calculated as weight divided by squared height (kg/m^2). Blood volume (BV) was calculated as $BV = 0.604 + 0.367 * height\ (m)^3 + 0.0322 * weight\ (kg)$ for men and $BV = 0.183 + 0.356 * height\ (m)^3 + 0.0331 * weight\ (kg)$ for women.[9] Seasonality was defined as the four meteorological seasons: Winter (visits between December 1st and February 29th), Spring (March 1st to May 31st), Summer (June 1st to August 31st), Fall (September 1st to November 30th). Deferral at the previous visit was defined as either deferral because of a low Hb level, deferral because of reasons other than a low Hb level, or no deferral.
All information except for height was directly obtained from the administrative donor database. Height data were obtained by means of self reporting.

Statistical analysis

Missing data are often not completely at random, but rather selectively missing. Simply deleting the donors with missing values (so-called complete case analysis) would thus yield invalid study results. Imputation of selectively missing values reduces bias and allows for including all donors in the analysis.[10-12] Hence, as recommended, we multiple imputed (MI) these missing values ten times (aregImpute function from the Design library,[13] applicable in R software[14]). The imputation models included all the candidate predictors and the outcome variable.[15] Analyses were performed in each MI dataset. Estimates from the ten MI datasets were then combined into one overall estimate and variance according to Rubin's rules.[16]

Logistic regression analysis was performed with low Hb level (yes/no) as dichotomous outcome. The form of association between continuous candidate predictors and low Hb level was studied with restricted cubic spline functions.[17] A backward stepwise selection procedure was used to select the strongest predictors for low Hb levels. Selection was based on the likelihood ratio test with $p < 0.20$. We hypothesized that the effects of the candidate predictors "previous Hb level" and "the total number of whole blood donations in the past two years" could be different for men and women. Possible differences in the effect of these predictors for men and women were studied with interaction terms. Deferral at the previous visit is related to previous Hb level, and interaction between these two predictors was also examined. We used one overall test for the three interaction terms together, to prevent problems of multiple testing.

Overall performance of the prediction model was assessed with the Nagelkerke R^2, a measure of explained variation. The R^2 indicates the percentage of the total variation in low vs. appropriate Hb levels between donors that can be explained by the predictors of the model. The discriminative ability of the model, being the ability of the model to distinguish donors with low Hb levels from donors with appropriate Hb levels, was determined with the area under the Receiver Operating Characteristic (ROC-) curve (AUC). The AUC indicates the percentage of pairs of donors in which one donor has a low Hb level and the other donor has an appropriate Hb level, for which is correctly assigned a higher risk for a low Hb level to the donor that had indeed a low Hb level.

Developed models may be overfitted, meaning that high predictions are too high and low predictions are too low. Furthermore, apparent performance estimates, such as the R^2 and AUC may be too optimistic. A bootstrap procedure was performed to assess optimism. This procedure provides an optimism-corrected R^2 and AUC. One hundred bootstrap samples were drawn with replacement and models were fitted in each bootstrap sample. The mean difference in performance of the bootstrap models between the bootstrap and original data was used as a measure for optimism. The apparent R^2 and AUC were corrected for optimism by subtracting the optimism estimates. The bootstrap procedure also provides a shrinkage factor for the regression coefficients of the predictors in the model. The shrinkage factor is the average slope of the calibration plots for the bootstrap models that are applied in the original data. The regression coefficients of the predictors were multiplied by the shrinkage factor (shrunken) to prevent the model for giving too extreme predictions in donors that were not used for the development of the model.

Several risk thresholds for the predicted risks were set to divide donors into groups with low versus high risk of low Hb levels. The accuracy at the different thresholds was assessed. To facilitate the application of the model in practice, we transformed the logistic formula into a ready to use score chart. To do so, the shrunken regression coefficients were first converted into scores by rounding to integers. Next, a constant was subtracted or added to rescale the scores in positive integers. The sum scores were then related to predicted risks.

Statistical analyses were performed with SPSS, Version 14, SPSS, Inc., Chicago, IL; and R, Version 2.7.1, http://cran.r-project.org/bin/windows/base/old/2.7.1/.

Results

A total of 5429 donors met the inclusion criteria. Next, 164 donors were excluded because they had an unclear donation history. Another 74 donors were excluded because the Hb level at the intended visit was unknown.
As a result, 5191 donors were included in the study. Of these, 2795 (53.8%) were complete cases. Incompleteness of cases was notably due to missing height data.
We observed 143 deferrals for low Hb levels (2.8%). The distribution of candidate predictors and Hb deferral at the intended visit is presented in Table 1.
All continuous candidate predictors showed a linear association with low Hb level; transformation of candidate predictors was hence not necessary. Results of the univariable logistic regression analysis are presented in Table 2. Table 3 presents the results of the multivariable analysis before shrinkage of the regression coefficients. Seven out of eleven candidate predictors were selected with backward selection. Previous Hb level was the strongest predictor for low Hb levels. The overall test for interaction was significant ($p = 0.0057$); the three interaction terms were hence included in the model.
For a proper interpretation, regression coefficients of the interaction effects need to be added to the regression coefficients of the main effects. For example, to examine the effect of previous Hb level in women, the regression coefficient of the interaction between sex and previous Hb level has to be added to the regression coefficient of previous Hb level. So, for women the effect of previous Hb level is $-3.27 + 0.23 * 1 = -3.04$, and for men the effect is $-3.27 + 0.23 * 0 = -3.27$. Because there is also interaction between previous Hb level and previous deferral, the regression coefficient of interaction between these terms should also be added to interpret the effect of previous Hb level if donors were deferred at the previous visit.
Figure 1 presents the model as a ready to use score chart, intended for easy use in practice. In this score chart, interaction terms are incorporated in the scores. The score for the predictor previous Hb level is different for men and women, and for donors that were deferred and that were approved at the previous visit. The scores per category are the sum of the main effects of previous Hb level, sex, previous deferral, and the accompanying interaction terms.
The R^2 of the apparent model was 0.26, which is reasonable for a logistic regression model. The AUC of the apparent model was 0.88. The model showed good internal validity: after the bootstrap procedure, the R^2 became 0.23 and the AUC 0.87. The shrinkage factor to correct for overfitting was 0.92. The regression coefficients were shrunken with this factor and the intercept was adjusted to 22.65.

Table 1 Distribution of candidate predictors and Hb level at the intended visit for the total study population and for donors with low and appropriate Hb levels separately*

Candidate predictor	Total (n=5191)	Low Hb† (n=143)	Appropriate Hb (n=5048)	% missing
Sex, female	1874 (36.1)	99 (69.2)	1775 (35.2)	0
Age, years	47 (±11)	44 (±12)	47 (±11)	0
BMI, kg/m	25.6 (±3.5)	24.7 (±3.5)	25.6 (±3.5)	44.7
Blood volume, L	5.04 (±0.80)	4.51 (±0.67)	5.05 (±0.79)	44.7
Seasonality				0
Winter	870 (16.8)	13 (9.1)	857 (17.0)	
Spring	1497 (28.8)	43 (30.1)	1454 (28.8)	
Summer	1573 (30.3)	45 (31.5)	1528 (30.3)	
Fall	1251 (24.1)	42 (29.4)	1209 (24.0)	
Previous Hb level, mmol/L				1.2
Men	9.4 (±0.7)	8.6 (±0.5)	9.4 (±0.7)	
Women	8.5 (±0.6)	7.9 (±0.5)	8.5 (±0.6)	
Hb second last visit, mmol/L				1.7
Men	9.4 (±0.6)	8.7 (±0.6)	9.4 (±0.6)	
Women	8.5 (±0.6)	8.0 (±0.5)	8.5 (±0.6)	
Delta Hb, mmol/L	0.0 (-0.4 – 0.4)	-0.1 (-0.6 – 0.3)	0.0 (-0.4 – 0.4)	2.9
Time since previous visit, days	147 (112 – 224)	148 (119 – 218)	147 (112 – 224)	
Deferral at previous visit				1.2
Due to low Hb	178 (3.5)	31 (21.8)	147 (2.9)	
Due to other reason than low Hb	133 (2.6)	2 (1.4)	131 (2.6)	
Number of whole blood donations in past 2 years	4 (2 – 5)	3 (2 – 4)	4 (3 – 5)	0
Plasma donation in past 2 years	50 (1.0)	3 (2.1)	47 (0.9)	0
Hb intended visit, mmol/L				0
Men	9.5 (±0.6)	7.9 (±0.3)	9.5 (±0.6)	
Women	8.5 (±0.6)	7.4 (±0.2)	8.6 (±0.5)	

* Data are reported as n (%), mean (±SD), or median (25^{th} – 75^{th} percentile).

† < 8.4 mmo/L for men and < 7.8 mmo/L for women.

Table 2 Univariable associations between candidate predictors and low Hb levels

Candidate predictor		OR (95% CI)*
Sex, female		4.15 (2.90 – 5.95)
Age ≤ 50 years†		1.49 (1.05 – 2.12)
BMI ≤ 25 kg/m²†		1.35 (0.86 – 2.14)
Blood volume, L (e.g. 4 L vs. 5 L)		1.92 (1.23 – 2.94)
Seasonality	Winter‡	1
	Spring	1.95 (1.04 – 3.65)
	Summer	1.94 (1.04 – 3.62)
	Fall	2.29 (1.22 – 4.29)
Previous Hb level, mmol/L (e.g. 8 mmol/L vs. 9 mmol/L)		6.25 (4.76 – 8.33)
Delta Hb, mmol/L, mmol/L (e.g. 0 mmol/L vs. 1 mmol/L)		1.35 (1.02 – 1.75)
Time since previous visit ≤ 5 months†		0.95 (0.68 – 1.32)
Deferral at previous visit	No deferral‡	1
	Due to low Hb	8.12 (5.28 – 12.47)
	Due to other reason than low Hb	0.60 (0.15 – 2.45)
Number of whole blood donations in past 2 years ≤ 4 times†		2.56 (1.67 – 3.94)
Plasma donation in past 2 years, yes		2.28 (0.70 – 7.41)

* OR = exp (beta). A value < 1 indicates a decreased risk and a value > 1 indicates an increased risk.
† Most continuous predictors were dichotomized for easy interpretation of the odds ratio.
‡ Winter and no deferral are reference categories.

The accuracy related to the different threshold values to classify donors as low versus high at risk of having low Hb levels is shown in Table 4. Figure 2 presents the AUC with the different threshold values.

Table 3 Multivariable association between predictors and low Hb levels

Predictor			β* (95% CI)
Main effects			
Sex, female			-1.75 (-8.53 – 5.04)
Seasonality	Winter†		0
	Spring		0.72 (0.07 – 1.37)
	Summer		0.77 (0.12 – 1.42)
	Fall		0.95 (0.29 – 1.60)
Previous Hb level, mmol/L			-3.27 (-3.99 – -2.54)
Delta Hb, mmol/L			1.03 (0.67 – 1.40)
Time since previous visit, per 100 days			-0.26 (-0.43 – -0.08)
Deferral at previous visit		No deferral†	0
		Due to low Hb	-7.39 (-15.31 – 0.53)
		Due to other reason than low Hb	-19.67 (-39.93 – 0.59)
Number of whole blood donations in past 2 years			-0.01 (-0.18 – 0.16)
Interaction effects			
Sex, female x Previous Hb level, mmol/L			0.23 (-0.58 – 1.05)
Previous Hb level, mmol/L x Deferral at previous visit due to low Hb			0.86 (-0.16 – 1.89)
Previous Hb level, mmol/L x Deferral at previous visit due to other reason than low Hb			2.21 (-0.06 – 4.47)
Sex, female x Number of whole blood donations in past 2 years			-0.38 (-0.64 – -0.13)
Intercept			24.78 (18.43 – 31.13)

All selected predictors are statistically significant ($p < 0.20$).

* β = regression coefficient, shows the strength and the direction of the variable's influence.

† Winter and no deferral are reference categories.

Figure 1 Score chart for the prediction of low Hb levels

predictors									
Hb level measured at previous visit, (mmol/L)		*Value*	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	Men	*Score*							
	No previous deferral		na	na	14	11	8	5*	2*
	Previous deferral low Hb		13	11	9	na	na	na	na
	Previous deferral other		na	na	1*	0*	-1*	-2*	-3*
	Women								
	No previous deferral		na	na	15	12	9	6*	3*
	Previous deferral low Hb		13	11	9	na	na	na	na
	Previous deferral other		na	na	1*	0*	-1*	-2*	-2*
Delta Hb†, (mmol/L)		*Value*	**-2.5**	**-1.5**	**-0.5**	**0.5**	**1.5**	**2.5**	
		Score	0	1	2	3	4	5	
Seasonality		*Value*	**Winter**		**Other**				
		Score	0		1				
Time since previous visit, months		*Value*	**<8**	**8-21**	**22-35**	**>35**			
		Score	3	2	1	0			
Number of whole blood donations in past 2 years		*Value*	**1**	**3**	**5**	**7**			
	Men	*Score*	0	0	0	0			
	Women		3	2	1	0			

Total sum score	**<20**	**20**	**20.5**	**21**	**21.5**	**22**	**22.5**	**23**	**23.5**	**24**
Risk of low Hb level (%)	<1	1	2	3	4	7	10	16	24	34

Use of score chart: For each donor, a total sum score can be calculated by counting the scores that correlate to the characteristics of the donor. The sum score can than be linked to the individual risk in the box below. For example, a woman whose Hb level at the previous visit, 3 months ago (score 3), was 6.1, which is too low and was therefore a reason for deferral (score 13), whose Hb level at the second last donation was 7.8 (delta Hb is -1.7, score 1), who can be invited to a blood collection center in Spring (score 1), and who has given 2 whole blood donations in the past two years (score 2.5), has a total sum score of 20.5. This score refers to a risk of having a low Hb level of 2%.

na: not applicable (Combination of predictor values does not occur).

* The score is lower than 8, the total sum score is therefore always lower than 20 and the accompanying risk of having a low Hb level is lower than 1%.

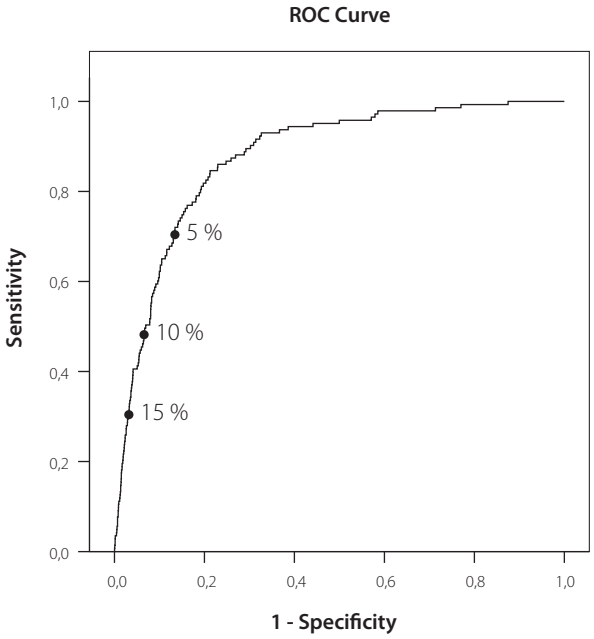
† Hb level measured at the previous visit minus the Hb level at the second last visit.

Table 4 Comparison of the accuracy of different threshold values of the predicted probability for low Hb levels

Probability	Donors %	Low Hb levels* % (95% CI)	NPV % (95% CI)	Sensitivity % (95% CI)	Specificity % (95% CI)
≥ 5%	15.6	12.6 (10.3 – 14.9)	99.1 (98.8 – 99.3)	71.6 (64.2 – 79.0)	86.0 (85.0 – 87.0)
≥ 10%	7.0	18.1 (14.1 – 22.0)	98.4(98.0 – 98.8)	46.2 (38.0 – 54.4)	94.1 (93.4 – 94.8)
≥ 15%	3.5	24.7 (18.4 – 31.0)	98.0 (97.6 – 98.4)	31.2 (23.6 – 39.0)	97.3 (96.9 – 97.7)

* = PPV.

Figure 2 ROC curve with different threshold values of the predicted probability for low Hb levels



Discussion

We assessed the value of various donor and donation characteristics to predict low Hb levels in whole blood donors. The strongest predictors were combined in a prediction model, including sex, seasonality, Hb level measured at the previous visit, difference in Hb levels between the previous two visits, time since the previous visit, deferral at the previous visit, and the total number of whole blood donations in the past two years. We found

different effects between men and women for the predictors previous Hb level and total number of whole blood donations in the past two years. The predictive effect of previous Hb level was also different for donors that were deferred at the previous visit compared to donors that were approved. The model discriminated adequately between donors with low and appropriate Hb levels.

Donors with a low risk of low Hb levels can be invited for a donation with preference. Invitation of donors depends on the available blood stock. When blood stock levels lower, adequate numbers of donors from the available donor base are invited. The model can be applied to identify eligible blood donors from the donor base. For the application of our model, a threshold value is required above which donors will be assigned in a high risk group of having low Hb levels. At lower threshold values more deferrals can be prevented (higher sensitivity), but at cost of more donors that are unnecessarily not invited for a donation (lower specificity) (Table 4). This may affect the productivity of blood units for transfusion. Caution should be given to the situation where uninvited donors might not be replaced by eligible donors. The number of deferrals that can be prevented must be weighted against the number of donors with appropriate Hb levels that are unnecessarily not invited for a donation at that time. At a threshold value of 10%, 46% of all deferrals can be prevented. At the same time, 5.7% (i.e. 100-18.1=81.9 of 7.0%) of the donors is unnecessarily not invited, which we consider acceptable.

Deferral for donation implies prolongation of the donation interval. Donors have more time to recover from the previous donation and the donation will take place in an other season. The risk of low Hb will probably be lower at the new time point. Donors in the high risk group of having low Hb levels may benefit from interventions such as postponement of the invitation for donation, a dietary advice or iron fortification.

Seasonality had an effect on the risk for low Hb levels: donors who visited a blood collection center in Spring, Summer or Fall had higher risks for low Hb levels compared to donors that visited a blood collection center in Winter. This is in agreement with a previous study in which higher deferral rates were observed in summer months compared to winter months.[8] Hb levels decrease with increasing daily temperature. One suggestion for an explanation is the physiological water shift into the vascular system as part of the heat balance system in changing environmental climate conditions. Related explanations might be indirect effects of seasonal differences in factors influencing Hb levels, like nutrition, physical activity, and virus infections during the year.

We found opposite results for the univariable and multivariable effects of delta Hb. The opposing effects are probably the result of high correlation with the previous Hb level (Pearson's correlation coefficient, $r = 0.41$). Although the interpretation of the effect is difficult, the inclusion of delta Hb provides better predictions. The effect of the total number of whole blood donations in the past two years (univariable and multivariable) was in the opposite direction to what was expected. The risk for low Hb levels was lower for donors who donated more often. A possible explanation is the healthy donor effect.

Donors with high Hb levels are more readily approved for donation and therefore donate more frequently.

Not all the candidate predictors were included in the multivariable model. Although younger donors had higher risks for low Hb levels, which is in agreement with other studies,[5,8] age had no added value in the multivariable model. Beforehand, we expected blood volume to be predictive, because the amount of blood given with a whole blood donation is always around 500 ml. Donors with a large blood volume lose relatively little blood with a donation, compared to donors with a small blood volume. Thus, the iron status of donors with a large blood volume might be less affected by a blood donation, resulting in higher Hb levels compared to donors with a small blood volume. Such a predictive effect was indeed found in the univariable analysis, but blood volume was not selected in the multivariable model. Either the added value of blood volume is limited, or our study was underpowered to detect an effect. Many values of blood volume had to be imputed (44.7%), because of the large number of missing values for height. Imputed values do not contribute to the effective sample size and therefore, the power to detect an association is relatively low for this predictor. This may also explain why we did not find a predictive effect of BMI.

Recent plasma donation showed little predictive effect, both univariably and multivariably. We had expected some beneficial effect of recent plasma donation compared to whole blood donation. With a plasma donation, red blood cells are returned to the donor. Only a small amount of red blood cells is lost and therefore plasma donation has an impact on Hb levels only when plasma is donated every week or fortnightly.[18] Most plasma donors do not donate that often, and therefore we had expected an advantageous effect of recent plasma donation.

The discriminative ability of our model was mainly based on the strong predictor previous Hb level. The AUC of a model with only this predictor was 0.84, whereas the AUC of the multivariable model was 0.88. Despite the high discriminative ability, the usefulness in practice might be limited, because previous Hb level and also delta Hb cannot be influenced such as seasonality and time since the previous visit. We therefore investigated if other candidate predictors would have been selected when previous Hb level and delta Hb were not considered. Interestingly, the same predictors and interaction terms as in the presented model were selected (sex, seasonality, time since the previous visit, previous deferral, and the total number of whole blood donations in the past two years). No additional predictors were selected. The R^2 of this model was much lower (0.12), as was the AUC (0.75). The performance of this model is substantially less than the model including the predictors previous Hb level and delta Hb. Therefore, we concluded the latter model to be the best.

A reason for the strong effect of previous Hb level may be that most donors have relatively constant Hb levels over time, far above the cutoff value. Possibly, other candidate predictors have stronger effects in donors that have Hb levels around the cutoff value. For

these donors, approval or deferral may depend on the value of other candidate predictors than previous Hb level. To investigate if other candidate predictors are more valuable in donors with Hb levels that vary around the cutoff value, we developed a model in the donors whose Hb level measured at the previous visit ranged from too low to maximal 1.0 mmol/L (subgroup 1), or 0.5 mmol/L (subgroup 2), above the sex-specific cutoff point for donation. Thus, donors in subgroup 1 had Hb levels measured at the previous visit that varied between too low up to and including 9.4 mmol/L for men and 8.8 mmol/L for women, and for donors in subgroup 2 these previous Hb levels ranged from too low up to and including 8.9 mmol/L for men and 8.3 mmol/L for women. After validation in the total donor population, the AUCs were 0.85 (subgroup 1) and 0.83 (subgroup 2). So, neither of these models performed better than the presented model.

This study has some limitations. One of our inclusion criteria was that donors should have given at least two whole blood donations. This criterion was used in order to study the effect of blood donation. However, newly registered donors and first time donors are not included now. There are more factors that are associated with Hb levels than the ones we have examined, for example smoking,[7,19] nutrition,[20] physical activity,[21] and race.[22] In a subsequent study we intend to examine the predictive effect of other factors, including these factors. These latter factors might especially be meaningful for newly registered and first time donors, because data on donation history are not yet available for these donors. Another limitation is that Hb levels were measure in finger stick capillary samples, which is usual for donor screening. Measurements in these samples are less precise than measurements in venous samples.

In conclusion, we developed a prediction model for low Hb levels in whole blood donors. The model predictions may be valuable to determine whether donors can be invited for a next donation, or whether some interventions such as postponement of the invitation, are warranted. If model performance remains adequate after external validation, which we will do in a subsequent study, the model can be implemented in the invitation process for blood donors.[23,24] Using such a model could help to decrease the number of donor deferrals for low Hb levels.

Reference List

1. European Directorate for the Quality of Medicines and HealthCare. Guide to the preparation, use and quality assurance of blood components, European Committee (partial agreement) on Blood Transfusion (CD-P-TS), Recommendation No. R(95) 15, 15th edition. Strasbourg: Council of Europe; 2009, p 1899.
2. Radtke H, Polat G, Kalus U, Salama A, Kiesewetter H. Hemoglobin screening in prospective blood donors: comparison of different blood samples and different quantitative methods. Transfus.Apher.Sci. 2005;33:31-5.
3. Halperin D, Baetens J, Newman B. The effect of short-term, temporary deferral on future blood donation. Transfusion 1998;38:181-3.
4. Custer B, Chinn A, Hirschler NV, Busch MP, Murphy EL. The consequences of temporary deferral on future whole blood donation. Transfusion 2007;47:1514-23.
5. Yip R, Johnson C, Dallman PR. Age-related changes in laboratory values used in the diagnosis of anemia and iron deficiency. Am.J.Clin.Nutr. 1984;39:427-36.
6. Micozzi MS, Albanes D, Stevens RG. Relation of body size and composition to clinical biochemical and hematologic indices in US men and women. Am.J.Clin.Nutr. 1989;50:1276-81.
7. Skjelbakken T, Dahl IM, Wilsgaard T, Langbakk B, Lochen ML. Changes in haemoglobin levels according to changes in body mass index and smoking habits, a 20-year follow-up of a male cohort: the Tromso Study 1974-1995. Eur.J.Epidemiol. 2006;21:493-9.
8. Hoekstra T, Veldhuizen I, van Noord PA, de Kort WL. Seasonal influences on hemoglobin levels and deferral rates in whole-blood and plasma donors. Transfusion 2007;47:895-900.
9. Nadler SB, Hidalgo JU, Bloch T. Prediction of blood volume in normal human adults. Surgery 1962;51:224-32.
10. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat.Med. 1991;10:585-98.
11. Schafer JL, Graham JW. Missing data: our view of the state of the art. Psychol.Methods 2002;7:147-77.
12. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J.Clin.Epidemiol. 2006;59:1087-91.
13. Harrell, F. E., Jr. Design: Design package. R package version 2.0. 2009. [cited 2009 Feb 10]. Available from: URL: http://biostat.mc.vanderbilt.edu/s/Design.
14. R Development Core Team. R: A language and environment for statistical computing. 2009. [cited 2009 Feb 10]. Available from URL: http://www.R-project.org.
15. Moons KG, Donders RA, Stijnen T, Harrell FE, Jr. Using the outcome for imputation of missing predictor values was preferred. J.Clin.Epidemiol. 2006;59:1092-101.
16. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. BMC.Med.Res.Methodol. 2009;9:57.
17. Harrell FE, Jr. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer; 2001.
18. Bier-Ulrich AM, Haubelt H, Anders C, Nagel D, Schneider S, Siegler KE, Seiler D, Hellstern P. The impact of intensive serial plasmapheresis and iron supplementation on iron metabolism and Hb concentration in menstruating women: a prospective randomized placebo-controlled double-blind study. Transfusion 2003;43:405-10.
19. Nordenberg D, Yip R, Binkin NJ. The effect of cigarette smoking on hemoglobin levels and anemia screening. JAMA 1990;264:1556-9.
20. Brussaard JH, Brants HA, Bouman M, Lowik MR. Iron intake and iron status among adults in the Netherlands. Eur.J.Clin.Nutr. 1997;51 Suppl 3:S51-S58.
21. Beard J, Tobin B. Iron status and exercise. Am.J.Clin.Nutr. 2000;72:594S-7S.
22. Johnson-Spear MA, Yip R. Hemoglobin difference between black and white women with comparable iron status: justification for race-specific anemia criteria. Am.J.Clin.Nutr. 1994;60:117-21.
23. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.
24. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606.

Chapter 3

Development and validation of a prediction model for low hemoglobin deferral in a large cohort of whole blood donors

A. Mireille Baart*†, Wim L.A.M. de Kort*, Femke Atsma*, Karel G.M. Moons†, Yvonne Vergouwe†

* Sanquin Research, dept. Donor Studies, Nijmegen, the Netherlands

† Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands

Published in Transfusion, 2012;52:2559-2569

Abstract

Background: Each year, around 5% of the invited blood donors is eventually deferred from donation because of low hemoglobin (Hb) levels. Estimating the risk of Hb deferral in blood donors can be helpful in the management of the donation program. We developed and validated a prediction model for Hb deferral in whole blood donors, separately for men and women.

Study design and methods: Data from a Dutch prospective cohort of 220,946 whole blood donors were used to identify predictors for Hb deferral using multivariable logistic regression analyses. Validity of the prediction models was assessed with a cross-validation.

Results: 12,865 donors (5.8%) were deferred because of a low Hb level. The strongest predictors of Hb deferral were Hb level measured at the previous visit, age, seasonality, difference in Hb levels between the previous two visits, time since the previous visit, deferral at the previous visit, and the total number of whole blood donations in the past two years for both men and women. The prediction models had an area under the receiver operating characteristic curve (AUC) of 0.89 for men and 0.84 for women. Cross-validation showed similar results and good calibration.

Conclusion: Using a limited number of easy to measure characteristics enables a good prediction of Hb deferral risk in whole blood donors. The prediction models may guide the decision which donors to invite for a next donation, and for which donors the invitation should be postponed. Potentially, this could decrease the number of Hb deferrals in blood donors.

Introduction

Whole blood donation may lead to iron deficiency. To protect donors from developing iron deficiency after a donation, the iron status is assessed in blood donors prior to donation. Most commonly, the iron status is assessed by measuring hemoglobin (Hb) levels. Donors with relatively low Hb levels are not allowed to donate in order to prevent them from developing iron deficiency anemia. In addition, deferral of these donors guarantees that blood units for transfusion meet the required standards for Hb content.[1]
A substantial number of health assessments in donors lead to a deferral because of low Hb levels. Deferrals are demoralizing for donors. Furthermore, deferrals increase the risk of donor lapse,[2-5] although the donor may actually meet the Hb criterion at the time of the next donation invitation.
For these reasons, prediction of Hb deferral risk would be valuable. Such predictions may be helpful to determine whether donors can be invited for a next donation, or whether some interventions such as postponement of the invitation or a dietary advice, are warranted. Eventually, this could decrease the number of donor deferrals for low Hb levels. In a previous study we already developed a prediction model for the risk of having an Hb level below the sex specific cutoff level for donation as described in the European Commission Directive.[6] However, we developed a single model to be applied in both men and women, and the model was developed in a relatively small sample of whole blood donors from only one Dutch region. Based on the results of this previous research and having now collected the individual data from a forty times larger sample of whole blood donors from all four geographical regions in the Netherlands, we had the unique opportunity for the current study to develop and validate much more robust prediction models for Hb deferral, separately for men and women. We chose to develop and validate models separately for men and women, in order to allow for different predictor effects between men and women.

Study design and methods

Study population

Whole blood donors, who visited a blood collection center in the Netherlands in the years 2007 until 2009 were eligible for the study. We defined whole blood donors as blood donors whose last two donations are whole blood donations. All these donors fulfilled the Dutch criteria for donation with relation to the time interval between donations and donation frequency. The minimum time interval between two donations is 56 days, and the maximum number of donations per year is five for men and three for women. From these donors, data of all visits from 2005 until 2009 were extracted from the donor database. To be included, donors should have donated whole blood at least twice in this

period, prior to the visit of which the measured Hb level was used as outcome (see below). Also, donors should not have uncertainties in their history, e.g. whether or not a donation took place. A total of 226,513 donors fulfilled the inclusion criteria.
Donors were excluded if the outcome was unknown (n=5567). Finally, 220,946 whole blood donors (112,491 men and 108,455 women) were included in the study.

Outcome: Hb deferral

We predicted the risk of Hb deferral at the so called "intended visit". For donors who quit donating in 2007 or 2008, the last visit was indicated as the intended visit. For active donors, a randomly selected visit in 2009 was used as the intended visit. This random selection was done to allow for equal distributions across the seasons, that is, to avoid a clustering of donations at the end of 2009. The proportions of intended visits in seasons in the years 2007 and 2008 were similar.
Hb levels were routinely measured in finger stick capillary samples using a photometer (HemoCue, Angelholm, Sweden). Hb deferral was defined as not meeting the sex specific cutoff level for donation as described in the European Commission Directive,[6] which is 8.4 mmol/L (13.5 g/dL) for men and 7.8 mmol/L (12.5 g/dL) for women.

Candidate predictors

We studied the following candidate predictors for the risk of Hb deferral: age, seasonality, Hb level measured at the previous visit (previous Hb level), difference in Hb level between the previous two visits (delta Hb), time since the previous visit, deferral at the previous visit, total number of whole blood donations in the past two years, a history of plasma donation in the past two years (yes/no), BMI and blood volume.
Seasonality was defined as the four meteorological seasons: winter (visits between December 1 and February 29), spring (March 1 to May 31), summer (June 1 to August 31), fall (September 1 to November 30). Deferral at the previous visit was defined as either deferral because of a low Hb level, deferral because of reasons other than a low Hb level, or no deferral. BMI and blood volume were calculated from data on height and total body weight. BMI was calculated as weight divided by squared height (kg/m^2). Blood volume (BV) was calculated using Nadler's formula:
$BV = 0.604 + 0.367 * height\ (m)^3 + 0.0322 * weight\ (kg)$ for men and
$BV = 0.183 + 0.356 * height\ (m)^3 + 0.0331 * weight\ (kg)$ for women.[7]
All information except for height was directly obtained from the administrative donor database. Height data were obtained by means of self-reporting.
Since height and weight are not routinely measured during the health assessments, we studied the added value of BMI and blood volume rather than considering them as candidate predictors in a full model.

Statistical analysis

A total of 6159 donors (2.8%) had missing values for one or more of the candidate predictors (BMI and blood volume left out of consideration). These missing values were observed for the variables previous Hb level, delta Hb and deferral at the previous visit. In order to be able to use the observed information of the other known variables, we imputed missing values once.[8,9]

For the development of the prediction models, logistic regression analysis was used with Hb deferral (yes/no) as dichotomous outcome variable. Univariate regression coefficients and odds ratios with 95% CIs were estimated for each predictor, separately for men and women. The nature of association between continuous candidate predictors and risk of Hb deferral was studied with restricted cubic spline functions with three, four or five knots.[10] The restricted cubic splines were plotted and in most cases approximated with simple transformations. The simplest transformation is a linear term. Another transformation is a piecewise linear function, which consists of several linear pieces with one or more knots.

A multivariable model that included all candidate predictors with chosen transformations was fitted separately for men and women. A backward stepwise selection procedure was used to select the strongest predictors for Hb deferral. Selection was based on the Akaike's Information Criterion (AIC), which corresponds to a p value of 0.157 for a predictor with one regression coefficient.[11] Additionally, interaction terms with previous Hb level were assessed, and also a possible interaction effect between deferral at the previous visit and time since the previous visit was examined.

The predictive performance of the models was assessed with calibration and discrimination. Calibration refers to the agreement between predicted probabilities and observed frequencies of Hb deferral.[12] Calibration was studied with a logistic regression model with the linear predictor as the only covariable. Ideally, the intercept of this calibration model is 0 and the slope is 1. Discrimination refers to the ability of the model to discriminate between deferred and approved donors based on their Hb level.[12] Discrimination was assessed with the area under the receiver operating characteristic (ROC-) curve (AUC). An AUC-value of 1 indicates perfect discrimination, a value of 0.5 indicates poor discrimination, equivalent to flipping a coin.

In order to validate the models, we developed prediction models in donors from three of the four Dutch geographical regions (Northeast, Northwest, Southeast or Southwest), and we cross-validated these models in the remaining region in terms of calibration and discrimination. The number of whole blood donors in each of the four geographical regions was as follows: 16,467 in the Northeast region, 67,609 in the Northwest region, 62,411 in the Southeast region and 74,459 in the Southwest region.

We studied the added value of BMI and blood volume with the net reclassification improvement (NRI). NRI focuses on reclassification tables constructed separately for deferred and approved donors, and quantifies the correct movement in categories (to a

higher risk group for deferred donors and to a lower risk group for approved donors).[13] The threshold level to classify donors as deferred was chosen at a predicted risk greater than 10%. For sensitivity analysis, we also estimated the NRI at a threshold level of 5% and 15% risk.

BMI and blood volume could not be calculated for 94,516 donors (42.8%) because of missing height or weight data. Most donors with missing values were from the Northern regions: 74,218 donors (88.3%) in the two Northern regions had missing values for height or weight. We therefore examined the added value of BMI and blood volume in donors from the Southern regions only (61,630 men and 54,942 women).

Finally, we derived ready to use score charts based on the prediction models. To do so, the regression coefficients of all predictors except previous Hb level were first multiplied by the predictor values and then converted into scores by rounding to integers. Next, a constant was subtracted or added to rescale the scores into positive integers. The sum scores were then combined with the effect of previous Hb level and related to predicted risks.

Statistical analyses were performed with SPSS, Version 18, SPSS, Inc., Chicago, IL; and R, Version 2.9.2, http://cran.r-project.org/bin/windows/base/old/2.9.2/.

Results

A total of 4568 male donors (4.1%) and 8,297 female donors (7.7%) were deferred because of a low Hb level. The distribution of candidate predictors and Hb deferral is presented in Table 1.

Linear associations of continuous candidate predictors with risk of Hb deferral were observed for the total number of whole blood donations in the past two years and for blood volume both in men and women. A good transformation for previous Hb level, delta Hb, time since the previous visit and BMI was a piecewise linear function based on two pieces both in men and women. In men, the association between age and the risk of Hb deferral was linear, whereas in women the association showed several curves (Figure 1). All candidate predictors except a history of plasma donation in the past two years were retained in the multivariable models. We observed no interaction effects. Table 2 presents the results of the univariate and multivariable logistic regression analyses in terms of odds ratios. See Appendix I for the exact formulas to calculate the risk of Hb deferral.

Calibration plots assessed with the development data showed good model fit over the complete range of predictions. The models could discriminate well between donors with low and appropriate Hb levels. The AUC in the development data set was 0.89 (95% CI 0.88 - 0.89) for the model for men and 0.84 (95% CI 0.83 - 0.84) for the model for women.

In order to study the validity of our models, we also developed prediction models in three of the four regions and cross-validated these models in the remaining region. For models

Table 1 Distribution of candidate predictors and Hb deferral at the intended visit for men and women separately*

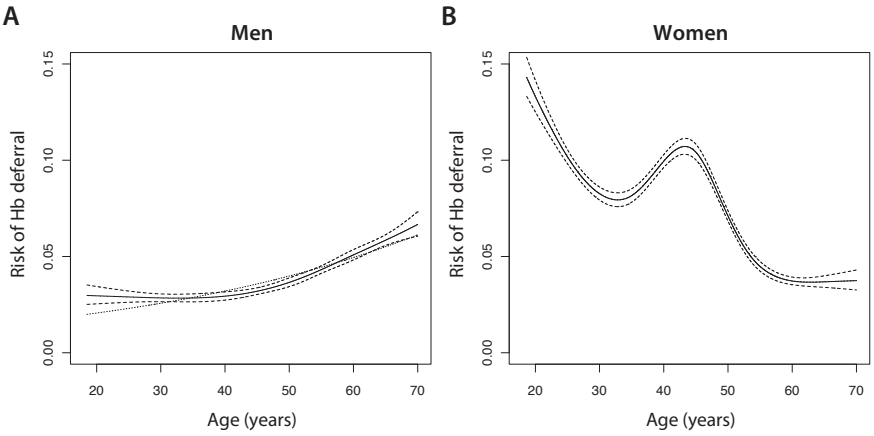
Candidate predictor	Total	Hb deferral†	Hb approval
Men	**n=112,491**	**n=4,568**	**n=107,923**
Age, years	49 (±12)	52 (±12)	49 (±12)
Seasonality			
Winter	22,789 (20.3)	775 (17.0)	22,014 (20.4)
Spring	29,040 (25.8)	1,399 (30.6)	27,641 (25.6)
Summer	31,315 (27.8)	1,477 (32.3)	29,838 (27.6)
Fall	29,347 (26.1)	917 (20.1)	28,430 (26.3)
Previous Hb level, mmol/L	9.4 (±0.7)	8.6 (±0.5)	9.4 (±0.6)
Delta Hb, mmol/L	0 (-0.4 – 0.4)	-0.1 (-0.4 – 0.2)	0 (-0.4 – 0.4)
Time since previous visit, days	112 (77 – 210)	86 (70 – 124)	113 (78 – 217)
Deferral at previous visit			
Due to low Hb	3,607 (3.2)	683 (15.0))	2,924 (2.7)
Due to other reason than low Hb	3,675 (3.3)	52 (1.1)	3,623 (3.4)
Number of whole blood donations in past 2 years	4 (2 – 6)	5 (3 – 7)	4 (2 – 6)
Plasma donation in past 2 years, yes	404 (0.4)	13 (0.3)	391 (0.4)
BMI, kg/m^2‡	26.1 (±3.2)	25.5 (±3.2)	26.1 (±3.2)
Blood volume, L‡	5.56 (±0.54)	5.41 (±0.52)	5.57 (±0.54)
Hb intended visit, mmol/L	9.3 (±0.7)	7.9 (±0.3)	9.4 (±0.6)
Women	**n=108,455**	**n=8,297**	**n=100,158**
Age, years	44 (±13)	40 (±12)	45 (±13)
Seasonality			
Winter	20,487 (18.9)	1,452 (17.5)	19,035 (19.0)
Spring	28,089 (25.9)	2,364 (28.5)	25,725 (25.7)
Summer	31,645 (29.2)	2,682 (32.3)	28,963 (28.9)
Fall	28,234 (26.0)	1,799 (21.7)	26,435 (26.4)
Previous Hb level, mmol/L	8.5 (±0.6)	7.9 (±0.5)	8.5 (±0.6)
Delta Hb, mmol/L	0 (-0.4 – 0.4)	-0.1 (-0.5 – 0.3)	0 (-0.4 – 0.4)
Time since previous visit, days	154 (120 – 259)	137 (117 – 185)	158 (121 – 266)
Deferral at previous visit			
Due to low Hb	6,658 (6.1)	1,683 (20.3)	4,975 (5.0)
Due to other reason than low Hb	4,311 (4.0)	131 (1.6)	4,180 (4.2)

Table 1 Continued

Candidate predictor	Total	Hb deferral†	Hb approval
Women	**n=108,455**	**n=8,297**	**n=100,158**
Number of whole blood donations in past 2 years	3 (2 – 4)	3 (2 – 4)	3 (2 – 4)
Plasma donation in past 2 years, yes	584 (0.5)	32 (0.4)	552 (0.6)
BMI, kg/m²‡	25.1 (±3.9)	24.3 (±3.7)	25.1 (±3.9)
Blood volume, L‡	4.28 (±0.48)	4.24 (±0.46)	4.29 (±0.48)
Hb intended visit, mmol/L	8.4 (±0.6)	7.3 (±0.3)	8.5 (±0.5)

* Data are reported as mean (±SD), n (%) or median (25th – 75th percentile).
† Hb < 8.4 mmol/L for men and < 7.8 mmol/L for women.
‡ Only measured in donors from the Southwest and Southeast regions.

Figure 1 Restricted cubic spline functions for age in men (**A**) and women (**B**)



Nature of univariate association between age and risk of Hb deferral in men (A) and women (B). Continuous lines represent restricted cubic spline functions with 5 knots with 95% CIs. Dotted lines represent transformations. Transformations: A linear, B as restricted cubic spline function.

developed in men, calibration slopes were close to 1 for each of the four models (Figures 2A1-2A4). The AUCs ranged from 0.88 to 0.89. For models developed in women, the AUCs of the four models ranged from 0.80 to 0.84. The calibration slopes of three models were close to 1, and of one model (developed in the Northwest, Southwest and Southeast region and validated in the Northeast region) it was 0.64 (Figures 2B1-2B4).

Figure 3 presents the models as ready to use score charts, intended for easy use in practice. Figure 3A can be used to calculate sum scores based on six of the seven predictors in the model (all except previous Hb level) for men and women separately. Figure 3B relates the total sum scores (vertical axis) and the previous Hb level (horizontal axis) with the predicted risk for Hb deferral. The cross point of a vertical line drawn from the x-axis and a horizontal line drawn from the y-axis shows the corresponding predicted risk for Hb deferral. The lines represent risks of 2%, 4%, 6%, 8%, 10%, 15%, 20% and 25%.

For example, a man at the age of 55 (score 0.5), is invited to the blood collection center in summer (score 0.5). He has a difference in Hb level between the previous two visits of -0.5 (score 2.5). His last visit was four months ago (score 1.5). At that time he donated blood (no deferral, score 1). Further, he has given six whole blood donations in the past two years (score 0). Hence, his sum score is 6. With a previous Hb level of 8.6, his risk of Hb deferral is 15%, i.e. the cross point of a horizontal line from 6 points and a vertical line from 8.6 mmol/L is at 15%.

To derive the score chart we used the exact regression coefficients. However, because we rounded the scores in the last step, small differences in risk between different predictor values disappear in the score chart. For example, a 20 year old man has the same age score as a 40 year old man, although a 40 year old man has a greater risk of Hb deferral according to the prediction model.

Application of the prediction models using a threshold value of 10% risk of Hb deferral results in a decrease of the percentage of Hb deferrals in men from 4.1% to 2.6%, whereas the percentage of men for which the donation invitation is unnecessarily postponed is 10.4%. In women, the percentage of Hb deferrals decreases from 7.7% to 5.6%, whereas the percentage of women for which the donation invitation is unnecessarily postponed is 20.3%.

To study the added value of BMI and blood volume, models with and without these variables were fitted in donors from the Southwest and Southeast regions. In this sub population, 2522 men (4.1%) and 4,270 women (7.8%) were deferred from donation because of a low Hb level. Based on a likelihood ratio test, BMI and blood volume had additive predictive value ($P < 0.05$). The AUCs for the models with and without these variables were 0.89 and 0.88 respectively for men, and the AUCs for both models in women were 0.84.

The addition of BMI and blood volume to the model for men resulted in the following net reclassification: for 32 deferred donors classification improved (they moved to the high risk group) and for 76 approved donors classification improved (they moved to the low risk group). The NRI was estimated at 1.4%. The addition of BMI and blood volume to the

Table 2 Association between predictors and Hb deferral

Predictor		OR (95% CI)*			
		Univariate		Multivariable	
		Men	Women	Men	Women
Age (years)† 25 vs. 20		1.12 (1.11 – 1.13)	0.74 (0.70 – 0.77)	1.05 (1.03 – 1.06)	0.85 (0.80 – 0.89)
Age (years)† 40 vs. 35		1.12 (1.11 – 1.13)	1.24 (1.19 – 1.29)	1.05 (1.03 – 1.06)	1.17 (1.12 – 1.22)
Age (years)† 55 vs. 50		1.12 (1.11 – 1.13)	0.61 (0.59 – 0.64)	1.05 (1.03 – 1.06)	0.75 (0.72 – 0.78)
Age (years)† 65 vs. 60		1.12 (1.11 – 1.13)	0.99 (0.93 – 1.05)	1.05 (1.03 – 1.06)	1.03 (0.99 – 1.07)
Seasonality	Winter‡	1	1	1	1
	Spring	1.44 (1.31 – 1.57)	1.20 (1.13 – 1.29)	1.63 (1.48 – 1.80)	1.28 (1.19 – 1.38)
	Summer	1.41 (1.29 – 1.54)	1.21 (1.14 – 1.30)	1.51 (1.37 – 1.66)	1.33 (1.24 – 1.43)
	Fall	0.92 (0.83 – 1.01)	0.89 (0.83 – 0.96)	0.82 (0.74 – 0.91)	0.87 (0.81 – 0.94)
Previous Hb level, per mmol/L below sex specific cutoff value†		0.71 (0.61 – 0.83)	0.44 (0.40 – 0.49)	0.10 (0.08 – 0.14)	0.12 (0.10 – 0.15)
Previous Hb level, per mmol/L at or above sex specific cutoff value†		0.05 (0.04 – 0.05)	0.08 (0.07 – 0.09)	0.02 (0.02 – 0.03)	0.04 (0.03 – 0.04)
Delta Hb, per 0.1mmol/L equal to or below 0 mmol/L†		1.02 (1.01 – 1.03)	0.99 (0.98 – 0.99)	1.14 (1.13 – 1.16)	1.11 (1.10 – 1.12)
Delta Hb, per 0.1 mmol/L above 0 mmol/L†		0.93 (0.92 – 0.94)	0.99 (0.98 – 1.00)	1.13 (1.12 – 1.15)	1.15 (1.14 – 1.16)
Time since previous visit, per month below 1 year†		0.84 (0.84 – 0.84)	0.89 (0.89 – 0.89)	0.84 (0.82 – 0.85)	0.85 (0.84 – 0.87)
Deferral at previous visit	No deferral‡	1	1	1	1
	Due to low Hb	6.16 (5.64 – 6.73)	4.82 (4.54 – 5.12)	0.54 (0.46 – 0.64)	0.59 (0.53 – 0.66)
	Due to other reasons than low Hb	0.35 (0.27 – 0.44)	0.47 (0.41 – 0.55)	0.38 (0.30 – 0.49)	0.40 (0.34 – 0.46)
Number of whole blood donations in past 2 years		1.14 (1.12 – 1.15)	0.92 (0.90 – 0.93)	0.95 (0.93 – 0.96)	0.85 (0.83 – 0.87)
Plasma donation in past 2 years, yes vs. no		0.78 (0.45 – 1.37)	0.70 (0.49 – 1.00)		

* OR = exp (beta). A value < 1 indicates a decreased risk and a value > 1 indicates an increased risk.

† Nonlinear association with risk of Hb deferral.

‡ Winter and no deferral are reference categories.

Figure 2 Calibration plots for men (**A**) and women (**B**) in the four regions

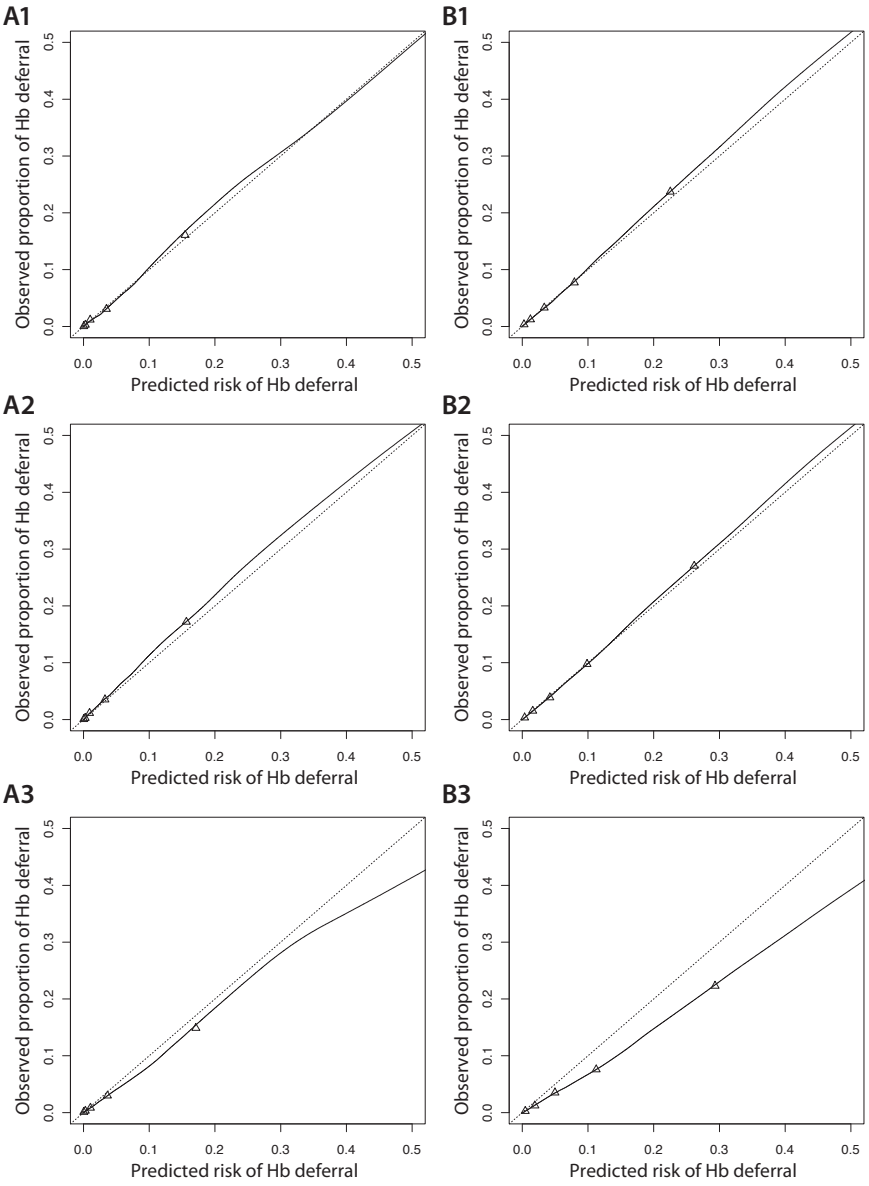
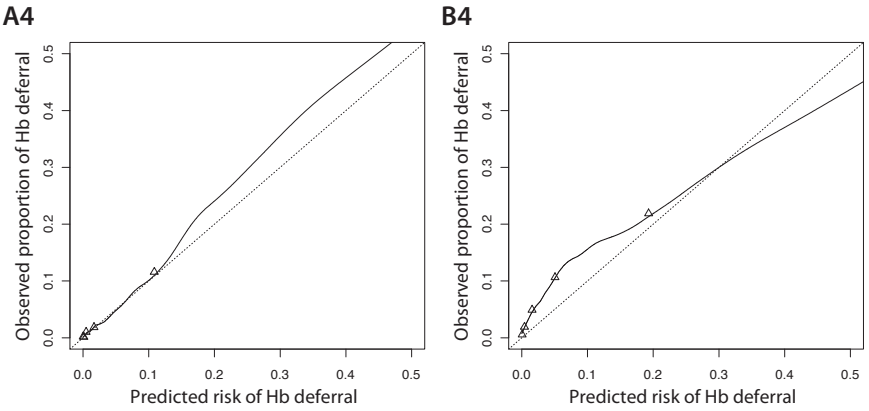


Figure 2 Continued



Calibration plots for men (A) and women (B) in the four regions. Figures 1-4 show calibration plots of models validated in the regions Northwest (1), Southwest (2), Southeast (3) and Northeast (4).
Triangles indicate the proportion of donors who were deferred because of low Hb levels per pentile of the predicted probabilities. The solid line shows the relation between observed proportions and predicted probabilities. Ideally, this line equals the dotted line.

Figure 3 Score chart for the prediction of Hb deferral, separately for men and women

A
Men

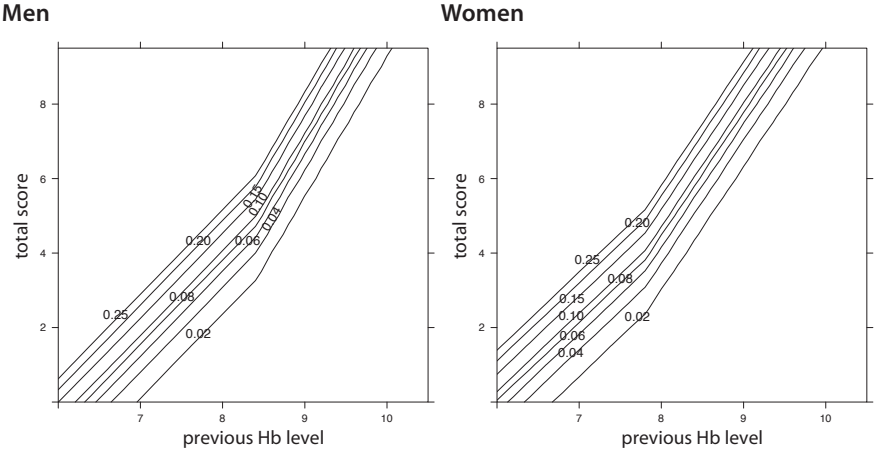
Predictors									
Age, years	*Value*	**18-45**	**46-70**						
	Score	0	0.5						
Seasonality	*Value*	**Fall**	**Winter**	**Spring**	**Summer**				
	Score	0	0	0.5	0.5				
Delta Hb, mmol/L	*Value*	**-2**	**-1.5**	**-1**	**-0.5**	**0**	**0.5**	**1**	**1.5**	**2**
	Score	0	1	1.5	2.5	3	3.5	4	5	5.5
Time since previous visit, months	*Value*	**1-2**	**3-5**	**6-7**	**8-10**	**>10**				
	Score	2	1.5	1	0.5	0				
Deferral at previous visit	*Value*	**No deferral**		**Due to low Hb**		**Due to other reasons**				
	Score	1		0.5		0				
Number of whole blood donations in past 2 years	*Value*	**0-4**	**≥5**						
	Score	0.5	0						

Figure 3 Continued

A
Women

Predictors										
Age, years	*Value*	**18-50**	**51-70**							
	Score	0.5	0							
Seasonality	*Value*	**Fall**	**Winter**	**Spring**	**Summer**					
	Score	0	0	0.5	0.5					
Delta Hb, mmol/L	*Value*	**-2**	**-1.5**	**-1**	**-0.5**	**0**	**0.5**	**1**	**1.5**	**2**
	Score	0	0.5	1	1.5	2	2.5	3.5	4	4.5
Time since previous visit, months	*Value*	**1-4**	**5-7**	**8-10**	**>10**					
	Score	1.5	1	0.5	0					
Deferral at previous visit	*Value*	**No deferral**		**Due to low Hb**		**Due to other reasons**				
	Score	1		0.5		0				
Number of whole blood donations in past 2 years	*Value*	**0-1**	**2-4**	**≥5**						
	Score	1	0.5	0						

B



Score chart for the prediction of Hb deferral. (A) Can be used to calculate sum scores based on six of the seven predictors in the model (all except previous Hb level) for men and women separately. (B) Relates the total sum scores (vertical axis) and the previous Hb level (horizontal axis) with the predicted risk for Hb deferral. The cross point of a vertical line drawn from the x-axis and a horizontal line drawn from the y-axis shows the corresponding predicted risk for Hb deferral. The lines represent risks of 2%, 4%, 6%, 8%, 10%, 15%, 20% and 25%.

model for women resulted in a better classification for two deferred donors and for 99 approved donors. The NRI was estimated at 0.2%. In both men and women BMI and blood volume improved classification in a minor way. NRI estimates at threshold levels of 5% and 15% risk were similar. See Appendix II for the exact formulas to calculate the risk of Hb deferral with the models that also include BMI and blood volume.

Discussion

In this paper we described the development and validation of a model to predict deferral for donation because of a low Hb level in whole blood donors. Models were developed separately for men and women. Predictive factors for Hb deferral in both men and women were Hb level measured at the previous visit, age, seasonality, difference in Hb level between the previous two visits, time since the previous visit, deferral at the previous visit, and total number of whole blood donations in the past two years. Adding BMI and blood volume did not improve the performance of the models. The performance of the models was in general good at cross-validation. Calibration was excellent and the models could discriminate well between donors with low and appropriate Hb levels.

Hb level measured at the previous visit was the strongest predictor for Hb deferral. The AUC of a model with only previous Hb level as predictor was 0.85 for men and 0.79 for women, compared to 0.89 and 0.84 for the multivariable models. This indicates that the other predictors in the model add less, but relevant information for the prediction of Hb deferral.

The predictive effect of age was opposite for men and women. In men, the risk of Hb deferral increased with increasing age. In women we found a decreasing effect with increasing age, with a temporary increasing effect between the age of 35 and 45. The opposite effect of age in men and women is consistent with the literature.[14-16] It is generally known that young women have a high risk of Hb deferral. The decrease in risk after the age of 45 can be explained by the effect of the menopause: women do not lose iron anymore by menstruation.[17]

The risk of Hb deferral was higher in spring and summer compared to fall and winter. This is in agreement with a previous study in which higher deferral rates were observed in summer months compared to winter months.[18] In that study, an inverse relationship was found between Hb levels and daily temperature. The underlying mechanism might be a water shift into the blood vessels as part of the heat regulation process. However, other explanations might be indirect effects of seasonal differences in factors influencing Hb levels, like nutrition and physical activity.

The longer the time interval since the previous visit, the lower the risk of Hb deferral. We chose to study the effect of time interval since the previous visit and not time since previous donation because this interval corresponds with the time of measurement of the

previous Hb level. Further, measurement of Hb level at the previous visit, whether or not a donation took place, results in the most recent assessment of Hb level.

The risk of Hb deferral was lower for donors with more donations. This effect was observed in the multivariable models for both men and women. In a previous study, an inverse association between donation frequency and risk of Hb deferral was also found.[15] A possible explanation for this finding is the healthy donor survivor effect.[19] This effect occurs when donors who are relatively healthy and are rarely deferred from donation remain donor and those who are unhealthy and have a repeatedly low Hb level do not. As a result the relatively healthy group of donors will be capable to achieve a large number of donations during their donor career in comparison to the more unhealthy group of donors.

According to the multivariable models, donors with no deferral at the previous visit had a higher risk of Hb deferral than donors that were deferred at the previous visit. This contra intuitive result is a consequence of the multivariable modelling. The univariate effects clearly show that donors with a previous deferral due to low Hb have a high risk of new deferral (ORs are 6 and 5 for men and women, respectively). We like to stress that the multivariable effects can only be used for predictive purposes. It does not imply that donors with a deferral at the previous visit have a lower risk than donors that were not deferred at the previous visit.

BMI and blood volume were associated with the risk of Hb deferral, although the predictive effects were small. Hence, the addition of these variables to the multivariable models did not substantially improve the model performance. Therefore, we did not include BMI and blood volume in our models, particularly since weight measurement should become part of the regular health assessment of donors at each visit.

A prediction model for low Hb levels in whole blood donors has been developed earlier by our research group.[20] A single model for men and women was developed. Although this model was developed in a much smaller dataset, the predictors included in the model were almost the same as in the newly developed sex specific models. Compared to the current models, age was not selected in the previous model. In both studies we used sophisticated prognostic techniques to develop the prediction models. For example, we studied the nature of association between continuous predictive factors and the risk of Hb deferral, we investigated calibration and discrimination properties, and we studied the validity of the prediction models. In the current study, we used a much larger dataset, resulting in robuster models with stable regression coefficients.

The prediction models can be applied in the invitation process of blood donors. Donors with a low risk of Hb deferral can be invited for a donation with preference. Invitation of donors depends on the available blood stock. When blood stock levels decrease, a sufficient number of donors from the available donor base will be invited. The model can be applied to identify eligible blood donors from the donor base. For the application of our model, a threshold value is required above which donors will be assigned in a high risk

group of Hb deferral. At lower threshold values more deferrals can be prevented, but at cost of more donors for which the donation invitation is unnecessarily postponed. This may affect the productivity of blood units for transfusion. Caution should be given to the situation where uninvited donors might not be replaced by eligible donors. The number of deferrals that can be prevented must be weighted against the number of donors with appropriate Hb levels for which the donation invitation is unnecessarily postponed at that time. At a threshold value of 10% risk of Hb deferral, application of the prediction model will help to decrease the percentage of Hb deferrals in men from 4.1% to 2.6%, whereas the percentage of men for which the donation invitation is unnecessarily postponed is 10.4%. In women, the percentage of Hb deferrals decreases from 7.7% to 5.6%, whereas the percentage of women for which the donation invitation is unnecessarily postponed is 20.3%. We consider the percentages of unnecessary postponements of invitation acceptable in relation to the decrease in percentages of Hb deferrals.

A limitation of this study is the large number of missing values for height and weight, which forced us to study the predictive value of BMI and blood volume in donors from the Southern regions only. We compared the observed values of predictive factors between donors from the Northern and the Southern regions. We observed no differences in the distribution of these variables between the two groups. Therefore, exclusion of the donors has probably not influenced our results.

To study the validity of the models, we performed a cross-validation. This type of validation gives a first indication of generalizability. External validation, particularly in other countries, is a necessary step before implementation in practice.[21,22] If the models show good performance in a group of independent donors, the models can be applied to identify eligible blood donors from the donor base. Donors with a low predicted risk of Hb deferral can be invited for a donation with preference, whereas those with a high risk of Hb deferral may benefit from interventions such as postponement of the invitation for donation, preferential invitation in fall and winter, iron fortification or a dietary advice. Ultimately, these models may be useful to decrease the number of donor deferrals for low Hb levels.

In conclusion, we developed and cross-validated sex specific prediction models for Hb deferral with good predictive accuracy. After external validation of the prediction models, these models may be useful to decrease the number of donor deferrals for low Hb levels.

Acknowledgement

The authors thank Dr. I. Veldhuizen for her contribution in fund raising.

Reference List

1. European Directorate for the Quality of Medicines and Healthcare. Guide to the preparation, use and quality assurance of blood components, European Committee (partial agreement) on Blood Transfusion (CD-P-TS), Recommendation No. R(95) 15, 15th ed. Strasbourg: Council of Europe; 2009, p. 1899.
2. Halperin D, Baetens J, Newman B. The effect of short-term, temporary deferral on future blood donation. Transfusion 1998;38:181-3.
3. Custer B, Chinn A, Hirschler NV, Busch MP, Murphy EL. The consequences of temporary deferral on future whole blood donation. Transfusion 2007;47:1514-23.
4. Zou S, Musavi F, Notari EP, Rios JA, Trouern-Trend J, Fang CT. Donor deferral and resulting donor loss at the American Red Cross Blood Services, 2001 through 2006. Transfusion 2008;48:2531-9.
5. Hillgrove T, Moore V, Doherty K, Ryan P. The impact of temporary deferral due to low hemoglobin: future return, time to return, and frequency of subsequent donation. Transfusion 2011;51:539-47.
6. The Commission of the European Communities. European Commission Directive 2004/33/EC. 2004.
7. Nadler SB, Hidalgo JU, Bloch T. Prediction of blood volume in normal human adults. Surgery 1962;51:224-32.
8. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J.Clin.Epidemiol. 2006;59:1087-91.
9. Marshall A, Altman DG, Royston P, Holder RL. Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. BMC.Med.Res.Methodol. 2010;10:7.
10. Harrell FE, Jr. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer; 2001.
11. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. Stat.Med. 2007; 26:5512-28.
12. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat.Med. 1996 28;15:361-87.
13. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat.Med. 2008;27:157-72.
14. Yip R, Johnson C, Dallman PR. Age-related changes in laboratory values used in the diagnosis of anemia and iron deficiency. Am.J.Clin.Nutr. 1984;39:427-36.
15. Mast AE, Schlumpf KS, Wright DJ, Custer B, Spencer B, Murphy EL, Simon TL. Demographic correlates of low hemoglobin deferral among prospective whole blood donors. Transfusion 2010;50:1794-802.
16. Cable RG, Glynn SA, Kiss JE, Mast AE, Steele WR, Murphy EL, Wright DJ, Sacher RA, Gottschall JL, Vij V, et al. Iron deficiency in blood donors: analysis of enrollment data from the REDS-II Donor Iron Status Evaluation (RISE) study. Transfusion 2011;51:511-22.
17. Milman N, Rosdahl N, Lyhne N, Jorgensen T, Graudal N. Iron status in Danish women aged 35-65 years. Relation to menstruation and method of contraception. Acta Obstet.Gynecol.Scand. 1993;72:601-5.
18. Hoekstra T, Veldhuizen I, van Noord PA, de Kort WL. Seasonal influences on hemoglobin levels and deferral rates in whole-blood and plasma donors. Transfusion 2007;47:895-900.
19. Atsma F, Veldhuizen I, Verbeek A, de Kort W, de Vegt F. Healthy donor effect: its magnitude in health research among blood donors. Transfusion 2011;51:1820-8.
20. Baart AM, de Kort WL, Moons KG, Vergouwe Y. Prediction of low haemoglobin levels in whole blood donors. Vox Sang. 2011;100:204-11.
21. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.
22. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606.

Appendix

Appendix I Logistic regression coefficients to calculate the risk of Hb deferral*

Predictor	Value or coding	Beta†	
		Men	Women
Age, per year‡	age	0.01	-0.03
	$(age-22.5)_+$^3		$1.22x10^{-4}$
	$(age-34.6)_+$^3		$-5.22x10^{-4}$
	$(age-45.0)_+$^3		$9.05x10^{-4}$
	$(age-54.0)_+$^3		$-6.79x10^{-4}$
	$(age-65.0)_+$^3		$1.73x10^{-4}$
Seasonality	winter‖	0	0
	spring	0.49	0.25
	summer	0.41	0.28
	fall	-0.20	-0.14
Previous Hb level, per mmol/L below sex specific cutoff value§	men: Hb-8.4, women: Hb-7.8	-2.27	-2.10
Previous Hb level, per mmol/L at or above sex specific cutoff value§	men: Hb-8.4, women: Hb-7.8	-3.77	-3.31
Delta Hb, per mmol/L equal to or below 0 mmol/L§	delta Hb	1.35	1.04
Delta Hb, per mmol/L above 0 mmol/L§	delta Hb	1.24	1.37
Time since previous visit, per month below 1 year§	months-12	-0.18	-0.16
Deferral at previous visit	no deferral‖	0	0
	due to low Hb	-0.62	-0.53
	due to other reasons than low Hb	-0.96	-0.93
Number of whole blood donations in past 2 years	nr of donations	-0.05	-0.17
Intercept		-2.77	-0.70

* With those regression coefficients the linear predictor can be calculated. The risk of Hb deferral equals: risk (Hb deferral) = $1/(1 + e^{-\text{linear predictor}})$.

† Beta = regression coefficient, shows the strength and the direction of the variable's influence.

‡ In men the association between age and risk of Hb deferral was linear, in women the association showed several curves that were modeled with a restricted cubic spline function (see Figure 1). Negative values of the cubic terms become 0; e.g. $(age-22.5)_+$ indicates age-22.5 for positive values, 0 for negative values.

§ Nonlinear association with risk of Hb deferral.

‖ Winter and no deferral are reference categories, 1 if true, 0 if false.

Appendix II Logistic regression coefficients to calculate the risk of Hb deferral for models that also include BMI and blood volume*

Predictor	Value or coding	Beta†	
		Men	Women
Age, per year‡	age	0.01	-0.03
	$(age-22.5)_+$^3		0.88x10-4
	$(age-34.6)_+$^3		-4.12x10-4
	$(age-45.0)_+$^3		7.63x10-4
	$(age-54.0)_+$^3		-5.92x10-4
	$(age-65.0)_+$^3		1.53x10-4
Seasonality	winter‖	0	0
	spring	0.45	0.24
	summer	0.39	0.22
	fall	-0.25	-0.26
Previous Hb level, per mmol/L below sex specific cutoff value§	men: Hb-8.4, women: Hb-7.8	-2.36	-2.09
Previous Hb level, per mmol/L at or above sex specific cutoff value§	men: Hb-8.4, women: Hb-7.8	-3.70	-3.27
Delta Hb, per mmol/L equal to or below 0 mmol/L§	delta Hb	1.41	1.01
Delta Hb, per mmol/L above 0 mmol/L§	delta Hb	1.17	1.39
Time since previous visit, per month below 1 year§	months-12	-0.18	-0.15
Deferral at previous visit	no deferral‖	0	0
	due to low Hb	-0.76	-0.60
	due to other reasons than low Hb	-1.13	-1.00
Number of whole blood donations in past 2 years	nr of donations	-0.06	-0.16
BMI, per kg/m^2 equal to or below 25 kg/m^2§	BMI-25	-0.08	-0.03
BMI, per kg/m^2 above 25 kg/m^2§	BMI-25	0.02	0.01
Blood volume, per L		-0.23	-0.17
Intercept		-1.74	-0.31

* With those regression coefficients the linear predictor can be calculated. The risk of Hb deferral equals: risk (Hb deferral) = $1/(1 + e^{-\text{linear predictor}})$.

† Beta = regression coefficient, shows the strength and the direction of the variable's influence.

‡ In men the association between age and risk of Hb deferral was linear, in women the association showed several curves that were modeled with a restricted cubic spline function (see Figure 1). Negative values of the cubic terms become 0; e.g. $(age-22.5)_+$ indicates age-22.5 for positive values, 0 for negative values.

§ Nonlinear association with risk of Hb deferral.

‖ Winter and no deferral are reference categories, 1 if true, 0 if false.

Chapter 4

External validation and updating of a Dutch prediction model for low hemoglobin deferral in Irish whole blood donors

A. Mireille Baart*†, Femke Atsma*, Ellen N. McSweeney‡, Karel G.M. Moons†, Yvonne Vergouwe†, Wim L.A.M. de Kort*

* Sanquin Research, dept. Donor Studies, Nijmegen, the Netherlands
† Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands
‡ Irish Blood Transfusion Service, Dublin, Ireland

Accepted for publication in Transfusion

Abstract

Background: Recently, sex specific prediction models for low hemoglobin (Hb) deferral have been developed in Dutch whole blood donors. In the present study, we validated and updated the models in a cohort of Irish whole blood donors.

Study design and methods: Prospectively collected data from 45,031 Irish whole blood donors were used. Hb cutoff levels for donation were approximately 0.35 mmol/L lower in Ireland than the Dutch cutoff levels (8.07 vs. 8.40 mmol/L in men; 7.45 vs. 7.80 mmol/L in women). The predictive performance of the models was assessed with calibration plots, calibration-in-the-large and the concordance (c)-statistic. The models were updated by revising the strength of the individual predictors in the models.

Results: A total of 613 men (2.4%) and 1624 women (8.4%) were deferred from donation because of a low Hb level. Validation demonstrated underestimation of predicted risks and lower c-statistics for men and women compared to the Dutch cohort. The strength of most predictive factors, particularly previous Hb level, was lower in Irish donors. The updated models showed a c-statistic of 0.83 (95% CI 0.81 – 0.84) for men and 0.76 (95% CI 0.74 – 0.77) for women.

Conclusion: The performance of Dutch prediction models for Hb deferral was limited when validated in Irish whole blood donors. Updating the models resulted in different predictor effects. This improved mainly the model calibration; the improvement in discrimination was small.

Introduction

Before a blood donation, donors are screened for hemoglobin (Hb) levels. Donors with relatively low Hb levels are deferred from donation to protect them from developing iron deficiency anemia after a donation. In addition, deferral of donors with low Hb levels guarantees that blood units for transfusion meet the required standards for Hb content[1]. Although deferrals are meant to protect donors, they are demoralizing for donors. Therefore, the risk of donor lapse is increased,[2-4] although the donor may actually meet the Hb criterion at the time of the next donation invitation. For these reasons, prediction of Hb deferral risk would be valuable. Such predictions may be helpful to determine whether donors can be invited for a next donation, or whether some interventions such as postponement of the invitation for donation or a dietary advice, are warranted. Eventually, the application of such a model could help to decrease the number of donor deferrals due to low Hb levels.

Prediction models for Hb deferral risk have been previously developed. A first prediction model was developed in a sample of Dutch whole blood donors[5]. Subsequently, more robust models for men and women separately were developed with data from all Dutch whole blood donors[6]. These prediction models include donor characteristics, visit characteristics and characteristics from the donation history. The models performed well at cross-validation within Dutch regions.

Since the prediction models were based on Dutch data only, the aim of this study was to externally validate the prediction models for Hb deferral in donors from another country. We externally validated the prediction models in a cohort of Irish whole blood donors. Hb cutoff levels for donation in Ireland are lower than in the Netherlands. This probably influenced the validity of the models and we therefore updated the prediction models for the Irish situation.

Donors and methods

Donors

All Irish whole blood donors, who visited any blood collection center in Ireland in the years 2008 until 2010 were eligible for the study. We defined whole blood donors as blood donors whose last two donations were whole blood donations. From these donors, data of all visits from January 2008 until December 2010 were extracted from the donor database.

For each donor an "intended visit" was defined. The Hb level measured at this visit was used as outcome in the analyses. For donors who quit donating in 2008 or 2009, their last visit was indicated as the intended visit. For active donors, a randomly selected visit in 2010 was used as the intended visit. This random selection was done to allow for equal

distributions across the seasons, that is, to avoid a clustering of donations at the end of 2010. The proportions of intended visits in seasons in the years 2008 and 2009 were similar. Inclusion and exclusion criteria were the same as those used for the development of the prediction model in Dutch donors. Donors should have donated whole blood at least twice prior to the intended visit since January 2008. Additionally, donors should not have uncertainties in their history, e.g. whether or not a donation took place. Furthermore, donors were excluded if the outcome was unknown. Finally, 45,031 Irish whole blood donors (25,766 men and 19,265 women) were included in the study.

Prediction model

The existing prediction models were developed in 112,491 Dutch male whole blood donors and 108,455 Dutch female whole blood donors[6]. The sex specific models predict the risk of Hb deferral at the intended visit. Hb deferral was defined as having an Hb level below the sex specific cutoff level for donation as described in the European Commission Directive[7], which is 8.4 mmol/L (13.5 g/dL) for men and 7.8 mmol/L (12.5 g/dL) for women. Predictive factors of the models are age, seasonality, Hb level measured at the previous visit (previous Hb level), difference in Hb level between the previous visit and the second-last visit prior to the intended visit (delta Hb), time since the previous visit, total number of whole blood donations in the past two years and deferral at the previous visit. Seasonality was defined as the four meteorological seasons: Winter (donations between December 1st and February 29th), Spring (March 1st to May 31st), Summer (June 1st to August 31st), Fall (September 1st to November 30th). Previous Hb level and delta Hb were included in the model with a piecewise linear function based on two pieces. The breakpoint for previous Hb level was chosen at the sex specific cutoff level for donation and for delta Hb at 0 mmol/L. Deferral at the previous visit was categorized as deferral because of a low Hb level, deferral because of reasons other than a low Hb level, and no deferral. See Appendix I for the exact formulas to calculate the risk of Hb deferral.

Data collection

Data on Hb levels and the relevant predictors were obtained from the Irish administrative donor database. Hb levels were routinely measured during donor screening in finger stick capillary samples using a photometer (HemoCue, Angelholm, Sweden). Irish Hb levels were assessed in g/dL rather than mmol/L. Cutoff levels for donation used in Ireland are 13.0 g/dL for men and 12.0 g/dL for women. For the analyses, Hb levels were converged into mmol/L. Converged Irish cutoff levels for donation are 8.07 mmol/L for men and 7.45 mmol/L for women. These values are approximately 0.35 mmol/L lower than the Dutch cutoff levels. Since the lower cutoff levels are used in the Irish donation practice, these cutoff levels were applied in the current study to define Hb deferral. Also, the breakpoint in the piecewise linear function for the predictor previous Hb level was set at these Irish cutoff levels.

Statistical analysis

External validation

Missing values occurred in the following variables: previous Hb level (3%), delta Hb (6%), and deferral at the previous visit (3%). In order to be able to use the observed information of other known variables, we single imputed missing values[8].

For the external validation, the predictive performance of the models was assessed in terms of calibration and discrimination[9]. Calibration is the agreement between predicted probabilities and observed frequencies. Calibration was studied with a logistic regression model with Hb deferral as dichotomous outcome and the linear predictor as the only covariate. The regression coefficient of the linear predictor (the calibration slope, visualized in a calibration plot) reflects whether the effects of the predictors in the Irish data are on average similar as the effects in the Dutch models, and is ideally 1. We also assessed calibration-in-the-large by fitting a logistic regression model with the linear predictor as an offset variable (setting the regression coefficient to 1). The intercept indicates whether predictions are in general correct, and is ideally 0. Discrimination is the ability of the model to differentiate between deferred and approved donors. Discrimination was determined with the concordance(c)-statistic. A value of 1 indicates perfect discrimination; a value of 0.5 indicates poor discrimination, equivalent to flipping a coin. We also calculated a benchmark value for the c-statistic to obtain more insight in the cause of a possible lower discriminative ability of the model in Irish donors. The benchmark c-statistic aims to disentangle two possible reasons for disappointing discrimination: a case mix effect and incorrectness of regression coefficients.[10] The benchmark c-statistic indicates the discriminative ability of the model under the condition that the model predictions are statistically correct in the validation data. To calculate the benchmark c-statistic, the outcome Hb deferral was simulated for the Irish donors with their own predictor values (case mix). Predicted risks were calculated for each donor and then the outcome value was generated based on the prediction. This was repeated 10 times for each donor and in this manner a stable benchmark c-statistic could be estimated.

Model updating

Usually, model performance is poorer in external validation compared to the performance in the development data. If this is the case, the models should be updated and adjusted to the conditions in the validation cohort to improve performance.[11-13]

Results of the external validation prompted us to update the models. We adjusted the intercept and regression coefficients of the prediction models to the Irish setting. The most important difference with the Dutch setting is the lower Hb cutoff level for donation, which affects the outcome and the breakpoint in the piecewise linear function for the predictor previous Hb level. Two methods were applied for updating: recalibration of the model and model revision[11]. Recalibration included adjustment of the intercept and adjustment of the individual regression coefficients with the same factor, i.e. the calibration

slope. For the revised models, individual regression coefficients were separately adjusted. This was done by adding the predictors to the recalibrated model in a step forward manner, and to test with a likelihood ratio test ($p<0.05$) if they had added value. If so, the regression coefficient for that predictor was adjusted further.
Statistical analyses were performed with SPSS, Version 19, SPSS, Inc., Chicago, IL; and R, Version 2.12.2, http://cran.r-project.org/bin/windows/base/old/2.12.2/.

Results

A total of 613 male donors (2.4%) and 1624 female donors (8.4%) were deferred because of a low Hb level (Table 1). In Irish donors mean Hb levels were lower than in Dutch donors at the intended visit (women) and at the previous visit (men and women). The difference for men was 0.1 mmol/L at the previous visit, and for women 0.2 mmol/L at the intended visit and 0.3 mmol/L at the previous visit. Irish donors were on average younger than Dutch donors (men 6 years and women 4 years) and Irish donors had donated less often in the past two years compared to Dutch donors (median value was 1 lower). For the other predictive factors, no substantial differences between Irish and Dutch donors were observed. The distribution of predictors at the intended visit in deferred and approved donors is presented in Appendix II.
Table 2 presents the results of the external validation. Predicted risks from the Dutch models were systematically too low for the Irish donors (Table 2) as indicated with the average predicted risks compared to the observed proportions of donors with Hb deferral and with the calibration-in-the-large. The calibration slopes deviated from the ideal value of 1: 0.65 (95% CI 0.60 – 0.70) for men and 0.63 (95% CI 0.59 – 0.67) for women. Calibration plots for men and women are presented in Figures 1A and 1B respectively.
Discrimination of the models was lower than in the Dutch development data: the c-statistic for men was 0.82 (95% CI 0.80 – 0.83) versus 0.89 (95% CI 0.88 – 0.89) in the development data), and for women 0.75 (95% CI 0.73 – 0.76) versus 0.84 (95% CI 0.83 – 0.84). The benchmark c-statistics were almost similar to the values in the development data: 0.90 for men and 0.83 for women.
The limited performance of the models in Irish donors prompted us to update the models. We updated the models with recalibration and with model revision. For the recalibrated models, all regression coefficients were multiplied by the slope of the calibration model (0.65 for men and 0.63 for women). The intercept was adjusted by multiplying the original value by the calibration slope and adding the accompanying intercept of the calibration model (-0.66 for men and -0.36 for women). To derive the revised models, regression coefficients of predictors that had added value in the recalibrated model were further adjusted. For men, regression coefficients were further adjusted for the predictors deferral at the previous visit, time since the previous visit, delta Hb and seasonality. For women,

Table 1 Distribution of predictors at the intended visit*

Predictor	Irish validation data		Dutch development data	
	Men	Women	Men	Women
	n=25,766	n=19,265	n=112,491	n=108,455
Age, years	43 (±12)	40 (±12)	49 (±12)	44 (±13)
Seasonality				
Winter	4,859 (18.9)	3,570 (18.5)	22,789 (20.3)	20,487 (18.9)
Spring	6,647 (25.8)	5,108 (26.5)	29,040 (25.8)	28,089 (25.9)
Summer	7,156 (27.8)	5,703 (29.6)	31,315 (27.8)	31,645 (29.2)
Fall	7,104 (27.6)	4,884 (25.4)	29,347 (26.1)	28,234 (26.0)
Previous Hb level, mmol/L	9.3 (±0.7)	8.2 (±0.6)	9.4 (±0.7)	8.5 (±0.6)
Delta Hb, mmol/L	0 (-0.4 – 0.4)	0 (-0.4 – 0.4)	0 (-0.4 – 0.4)	0 (-0.4 – 0.4)
Time since previous visit, days	155 (114 – 272)	167 (115 – 280)	112 (77 – 210)	154 (120 – 259)
Deferral at previous visit				
Due to low Hb (based on mmol/L)	419 (1.7)	1,163 (6.3)	3,607 (3.2)	6,658 (6.1)
Due to other reason than low Hb	212 (0.8)	330 (1.8)	3,675 (3.3)	4,311 (4.0)
Number of whole blood donations in past 2 years	3 (2 – 4)	2 (2 – 3)	4 (2 – 6)	3 (2 – 4)
Hb intended visit, mmol/L	9.3 (±0.7)	8.2 (±0.6)	9.3 (±0.7)	8.4 (±0.6)
Hb deferral†	613 (2.4)	1,624 (8.4)	4,568 (4.1)	8,297 (7.7)

* Data are reported as mean (±SD), n (%), or median (25^{th} – 75^{th} percentile).

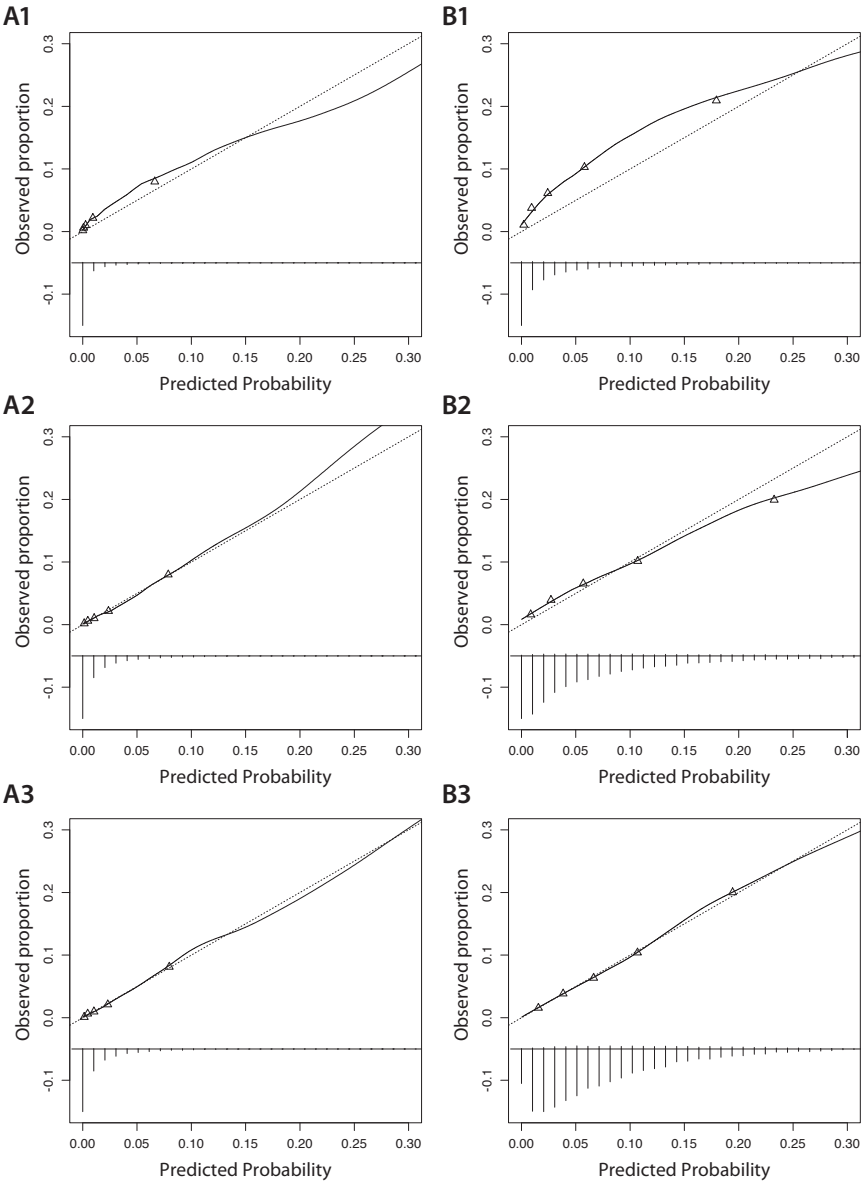
† Validation data: Hb < 8.07 mmol/L (= 13.0 g/dL) for men and < 7.45 mmol/L (=12.0 g/dL) for women.
Development data: Hb < 8.4 mmol/L (= 13.5 g/dL) for men and < 7.8 mmol/L (=12.5 g/dL) for women.

regression coefficients were further adjusted for deferral at the previous visit and delta Hb. See Appendix III for the exact formulas of the recalibrated and revised models to calculate the risk of Hb deferral. The adjusted regression coefficients in the revised models were generally lower than in the original models. This was especially true for previous Hb level. After updating, the models were (by definition) well calibrated (Figures 1A and 1B) and had slightly better discriminative ability than the original models: the c-statistic for the revised model in men was 0.83 (95% CI 0.81 – 0.84), and for women 0.76 (95% CI 0.74 – 0.77).

Table 2 Performance of the prediction models in the Irish validation data and the Dutch development data

Model	Predicted Hb deferral (%)	Observed Hb deferral (%)	Calibration-in-the-large (95% CI)	Calibration slope (95% CI)	C-statistic (95% CI)
Men					
Irish validation data	1.6	2.4	0.47 (0.38 – 0.55)	0.65 (0.60 – 0.70)	0.82 (0.80 – 0.83)
Dutch development data	4.1	4.1	0.00 (-0.06 – 0.06)	1.00 (0.97 – 1.03)	0.89 (0.88 – 0.89)
Women					
Irish validation data	5.5	8.4	0.54 (0.48 – 0.60)	0.63 (0.59 – 0.67)	0.75 (0.74 – 0.76)
Dutch development data	7.7	7.7	0.00 (-0.02 – 0.02)	1.00 (0.98 – 1.02)	0.84 (0.83 – 0.84)

Figure 1 Calibration plots for men (**A**) and women (**B**)



Calibration plots for men (A) and women (B). 1. original model, 2. recalibrated model, 3. revised model. Triangles indicate the proportion of donors with low Hb level per pentile of the predicted probabilities. The solid line shows the relation between observed proportions and predicted probabilities. Ideally, this line equals the dotted line. Vertical lines at the bottom indicate the distribution of the predicted probabilities; lines upward represent donors with low Hb levels, lines downward represent donors with appropriate Hb levels.

Discussion

We assessed the validity of sex specific Dutch prediction models for low Hb deferral in Irish whole blood donors. The performance of the Dutch prediction models was limited when validated in Irish donors: calibration was poor and the discriminative ability was lower than in the Dutch donors, particularly in women. Updating the models in Irish donors improved calibration, however the improvement in discrimination was very small. Important differences between the Dutch development data and the Irish validation data were lower mean Hb levels and lower Hb cutoff levels for donation in the Irish validation data. These differences probably had a negative influence on the model validity.

The Dutch prediction models had a lower ability to discriminate between deferred and approved donors in the Irish donor cohort. The two most likely reasons for this observation are differences in case mix and differences in predictor effects between the Dutch and Irish cohorts. Differences in case mix refer to different distributions of predictors in the development and the validation cohorts. If the predictor distributions indicate less heterogeneity in the validation data, the discriminative ability of the model is automatically lower.[10] In order to obtain more insight in the cause of the lower discriminative ability of the model in Irish donors we calculated the benchmark value for the c-statistic. This value was similar to the c-statistics estimated in the development cohort, both for men and women. The similarity indicates that the distributions of predictors were comparable in the development and the validation cohorts and that we could not identify a difference in case mix. As a consequence, the lower discriminative ability in Irish donors is probably the result of different predictor effects. Updating the prediction models by means of model revision resulted in new regression coefficients and thus different predictor effects. The predictive effects of most predictors, and particularly of previous Hb level, were weaker in Irish donors.

The poor calibration was also related to the weaker effect of previous Hb level. Furthermore, predicted risks were in generally too low. Possibly distributions of variables that were not included in the model but are related to Hb deferral are different between the two cohorts. This may for example be the case for dietary factors.

Important differences between the Dutch development data and the Irish validation data include lower mean Hb levels and lower Hb cutoff levels for donation in Irish donors compared to Dutch donors. These differences could also be the reason for the lower validity of the models in Irish donors.

The lower mean Hb levels are not a consequence of a higher donation frequency in Ireland as Irish donors had donated less often in the past two years compared to Dutch donors. Moreover, in a separate study we examined mean Hb levels in new donors (who presented at a blood collection for the first time and had not given a blood donation before) in the years 2008-2010, and we observed that mean Hb levels in Irish new donors were 0.1 mmol/L lower than mean Hb levels in Dutch new donors, in both men and

women. This finding indicates that mean Hb levels in the general Irish population are systematically lower than in the general Dutch population.

With respect to the lower Hb cutoff levels for donation it is difficult to disentangle the consequence on the validity of the Dutch models in Irish donors. The cutoff level is not only used to define the outcome Hb deferral; it is also used to define the variable "deferral at the previous visit" and the breakpoint used in the piecewise linear transformation for previous Hb level. For all these variables, Dutch cutoff levels were used during development of the models and Irish cutoff levels were used in the current validation study. To investigate whether the validity was affected by the lower Irish cutoff levels, we have also applied the Dutch cutoff levels in the Irish data, both in the outcome and in the predictors, and studied the validity of the Dutch models. Using Dutch cutoff levels, the c-statistic for men was lower (0.79 versus 0.82) and for women equal (0.75); however calibration-in-the-large was worse for both men (1.00 versus 0.47) and women (0.91 versus 0.54). Thus, the validity of the models was best when we used the Irish cutoff levels. Since these levels are the ones that are used in practice, we believe that these are most appropriate. Furthermore, to assess the generalisability of a prediction model, it is recommended to validate the prediction model in the broad sense in which even different definitions of outcome and predictive factors may be used.[9,13-18] However, despite this recommendation, to our knowledge no validation studies have been published before in which cutoff levels used to define the outcome were different from cutoff levels used during development of the model. The current study can therefore be considered unique.

The limited performance of the models in Irish donors prompted us to update the Dutch models for the Irish setting. We updated the models rather than fitting new prediction models to be able to use the information captured in the large development study.[13] Updating involved adjustment of the model intercept and the individual regression coefficients. Two methods were applied for updating: model recalibration and model revision. With recalibration all the regression coefficients are multiplied by the same calibration slope. This slope was mainly influenced by the predictor "previous Hb level" because this predictor was much stronger than the other predictors in the model. Revision therefore, resulted in weaker effects for previous Hb deferral, but similar effects as in the original Dutch models for most other predictors (Appendix I and III).

Results of this study led us to the recommendation to repeat the external validation in donors from a country in which the same Hb cutoff levels for donation are used. Besides, the lower discriminative ability of the model in Irish donors, particularly in women, indicates that it would be valuable to identify extra predictors. Beside the predictive factors in the Dutch prediction models, other factors are also associated with Hb levels. These factors could also be predictive for Hb deferral, especially in other countries than the Netherlands. Other factors that might have added value in the prediction models include for example ethnicity,[19,20] smoking,[21,22] nutrition,[23] and physical activity.[24-26] Furthermore, values of iron parameters in blood that can indicate early stages of iron

deficiency preceding a decrease of Hb levels, such as ferritin[27-29] or zinc protoporphyrin,[30-32] might also be predictive for Hb deferral. Addition of these new predictors might improve the model performance, both in the Dutch population and in other populations.
When prediction models show adequate performance at external validation, the models might eventually be applied in the invitation process of blood donors. The risk predictions can be used to decide for each individual donor at the moment of possible invitation for a donation whether they can indeed be invited. Donors with a low predicted risk of Hb deferral can be invited with preference, whereas those with a high risk of Hb deferral may benefit from interventions such as postponement of the invitation for donation, preferential invitation in fall and winter, iron fortification or a dietary advice. Ultimately, these models may be useful to decrease the number of donor deferrals for low Hb levels.
In conclusion, the predictive performance of Dutch prediction models for low Hb deferral was limited when validated in Irish whole blood donors. The different Hb cutoff levels for donation in the Irish setting had probably a negative influence on the model validity. Updating the prediction models for the Irish setting resulted in good calibration, but only slightly better discrimination. We believe that further studies will be required to assess the performance of prediction models for Hb deferral risk in countries with the same cutoff levels of either the Netherlands or Ireland.

Acknowledgements

The authors thank Ingrid Veldhuizen for her contribution in fund raising. Karel G.M. Moons gratefully acknowledges financial contribution for all his research by the Netherlands Organisation for Scientific Research (project 9120.8004 and 918.10.615).

Reference List

1. European Directorate for the Quality of Medicines and HealthCare. Guide to the preparation, use and quality assurance of blood components, European Committee (partial agreement) on Blood Transfusion (CD-P-TS), Recommendation No. R(95) 15, 15th ed. Strasbourg: Council of Europe; 2009, p1899.
2. Halperin D, Baetens J, Newman B. The effect of short-term, temporary deferral on future blood donation. Transfusion 1998;38:181-3.
3. Custer B, Chinn A, Hirschler NV, Busch MP, Murphy EL. The consequences of temporary deferral on future whole blood donation. Transfusion 2007;47:1514-23.
4. Zou S, Musavi F, Notari EP, Rios JA, Trouern-Trend J, Fang CT. Donor deferral and resulting donor loss at the American Red Cross Blood Services, 2001 through 2006. Transfusion 2008;48:2531-9.
5. Baart AM, de Kort WL, Moons KG, Vergouwe Y. Prediction of low haemoglobin levels in whole blood donors. Vox Sang. 2011;100:204-11.
6. Baart AM, de Kort WL, Atsma F, Moons KG, Vergouwe Y. Development and validation of a prediction model for low hemoglobin deferral in a large cohort of whole blood donors. Transfusion 2012;52:2559-2569.
7. The Commission of the European Communities. European Commissian Directive 2004/33/EC. 2004.
8. Marshall A, Altman DG, Royston P, Holder RL. Comparison of techniques for handling missing covariate data within prognostic modelling studies: a simulation study. BMC.Med.Res.Methodol. 2010;10:7.
9. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann.Intern.Med. 1999;130:515-24.
10. Vergouwe Y, Moons KG, Steyerberg EW. External validity of risk models: Use of benchmark values to disentangle a case-mix effect from incorrect coefficients. Am.J.Epidemiol. 2010;172:971-80.
11. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Stat.Med. 2004;23:2567-86.
12. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. J.Clin.Epidemiol. 2008;61:76-86.
13. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, Woodward M. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98:691-8.
14. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat.Med 2000;19:453-73.
15. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann.Intern.Med. 2006;144:201-9.
16. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. J.Clin.Epidemiol. 2008;61:1085-94.
17. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.
18. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606.
19. Perry GS, Byers T, Yip R, Margen S. Iron nutrition does not account for the hemoglobin differences between blacks and whites. J.Nutr. 1992;122:1417-24.
20. Johnson-Spear MA, Yip R. Hemoglobin difference between black and white women with comparable iron status: justification for race-specific anemia criteria. Am.J.Clin.Nutr. 1994;60:117-21.
21. Nordenberg D, Yip R, Binkin NJ. The effect of cigarette smoking on hemoglobin levels and anemia screening. JAMA 1990;264:1556-9.
22. Skjelbakken T, Dahl IM, Wilsgaard T, Langbakk B, Lochen ML. Changes in haemoglobin levels according to changes in body mass index and smoking habits, a 20-year follow-up of a male cohort: the Tromso Study 1974-1995. Eur.J.Epidemiol. 2006;21:493-9.
23. Zimmermann MB, Hurrell RF. Nutritional iron deficiency. Lancet 2007;370:511-20.
24. Cook JD. The effect of endurance training on iron metabolism. Semin.Hematol. 1994;31:146-54.
25. Beard J, Tobin B. Iron status and exercise. Am.J.Clin.Nutr. 2000;72:594S-7S.
26. Ottomano C, Franchini M. Sports anaemia: facts or fiction? Blood Transfus. 2012;10:252-4.
27. Cook JD. Clinical evaluation of iron deficiency. Semin.Hematol. 1982;19:6-18.
28. Skikne B, Lynch S, Borek D, Cook J. Iron and blood donation. Clin.Haematol. 1984;13:271-87.

29. Pasricha SR, McQuilten ZK, Keller AJ, Wood EM. Hemoglobin and iron indices in nonanemic premenopausal blood donors predict future deferral from whole blood donation. Transfusion 2011;51:2709-13.
30. Lamola AA, Yamane T. Zinc protoporphyrin in the erythrocytes of patients with lead intoxication and iron deficiency anemia. Science 1974;186:936-8.
31. Labbe RF, Rettmer RL. Zinc protoporphyrin: a product of iron-deficient erythropoiesis. Semin.Hematol. 1989;26:40-6.
32. Harthoorn-Lasthuizen EJ, Lindemans J, Langenhuijsen MM. Zinc protoporphyrin as screening test in female blood donors. Clin.Chem. 1998;44:800-4.

Appendix

Appendix I Logistic regression coefficients to calculate the risk of Hb deferral*

Predictor	Value or coding	Beta†	
		Men	Women
Age, per year‡	age	0.01	-0.03
	$(age-22.5)_+$^3		1.22x10^{-4}
	$(age-34.6)_+$^3		-5.22x10^{-4}
	$(age-45.0)_+$^3		9.05x10^{-4}
	$(age-54.0)_+$^3		-6.79x10^{-4}
	$(age-65.0)_+$^3		1.73x10^{-4}
Seasonality	winter‖	0	0
	spring	0.49	0.25
	summer	0.41	0.28
	fall	-0.20	-0.14
Previous Hb level, per mmol/L below sex specific cutoff value§	men: Hb-8.4, women: Hb-7.8	-2.27	-2.10
Previous Hb level, per mmol/L at or above sex specific cutoff value§	men: Hb-8.4, women: Hb-7.8	-3.77	-3.31
Delta Hb, per mmol/L equal to or below 0 mmol/L§	delta Hb	1.35	1.04
Delta Hb, per mmol/L above 0 mmol/L§	delta Hb	1.24	1.37
Time since previous visit, per month below 1 year§	months-12	-0.18	-0.16
Deferral at previous visit	no deferral‖	0	0
	due to low Hb	-0.62	-0.53
	due to other reasons than low Hb	-0.96	-0.93
Number of whole blood donations in past 2 years	nr of donations	-0.05	-0.17
Intercept		-2.77	-0.70

* With those regression coefficients the linear predictor can be calculated. The risk of Hb deferral equals: risk (Hb deferral) = $1/(1 + e^{-\text{linear predictor}})$.

† Beta = regression coefficient, shows the strength and the direction of the variable's influence.

‡ In men the association between age and risk of Hb deferral was linear, in women the association showed several curves that were modeled with a restricted cubic spline function. Negative values of the cubic terms become 0; e.g. $(age-22.5)_+$ indicates age-22.5 for positive values, 0 for negative values.

§ Nonlinear association with risk of Hb deferral.

‖ Winter and no deferral are reference categories, 1 if true, 0 if false.

Appendix II Distribution of predictors at the intended visit in deferred and approved donors*

Predictor	Irish validation data		Dutch development data	
	Hb deferral†	Hb approval	Hb deferral†	Hb approval
Men	**n=613**	**n=25,153**	**n=4,568**	**n=107,923**
Age, years	47 (±12)	43 (±12)	52 (±12)	49 (±12)
Seasonality				
Winter	102 (16.6)	4,757 (18.9)	775 (17.0)	22,014 (20.4)
Spring	226 (36.9)	6,421 (25.5)	1,399 (30.6)	27,641 (25.6)
Summer	169 (27.6)	6,987 (27.8	1,477 (32.3)	29,838 (27.6)
Fall	116 (18.9)	6,988 (27.8)	917 (20.1)	28,430 (26.3)
Previous Hb level, mmol/L	8.6 (±0.6)	9.3 (±0.7)	8.6 (±0.5)	9.4 (±0.6)
Delta Hb, mmol/L	-0.1 (-0.4 – 0.4)	0 (-0.4 – 0.4)	-0.1 (-0.4 – 0.2)	0 (-0.4 – 0.4)
Time since previous visit, days	140 (112 – 236)	155 (114 – 273)	86 (70 – 124)	113 (78 – 217)
Deferral at previous visit				
Due to low Hb	63 (10.3)	356 (1.5)	683 (15.0))	2,924 (2.7)
Due to other reason than low Hb	9 (1.5)	203 (0.8)	52 (1.1)	3,623 (3.4)
Number of whole blood donations in past 2 years	3 (2 – 4)	3 (2 – 4)	5 (3 – 7)	4 (2 – 6)
Hb intended visit, mmol/L	7.7 (±0.4)	9.3 (±0.6)	7.9 (±0.3)	9.4 (±0.6)

Women	**n=1,624**	**n=17,641**	**n=8,297**	**n=100,158**
Age, years	38 (±12)	41 (±12)	40 (±12)	45 (±13)
Seasonality				
Winter	272 (16.7)	3,298 (18.7)	1,452 (17.5)	19,035 (19.0)
Spring	502 (30.9)	4,606 (26.1)	2,364 (28.5)	25,725 (25.7)
Summer	497 (30.6)	5,206 (29.5)	2,682 (32.3)	28,963 (28.9)
Fall	353 (21.7)	4,531 (25.7)	1,799 (21.7)	26,435 (26.4)
Previous Hb level, mmol/L	7.8 (±0.5)	8.3 (±0.6)	7.9 (±0.5)	8.5 (±0.6)
Delta Hb, mmol/L	-0.1 (-0.5 – 0.3)	0 (-0.4 – 0.4)	-0.1 (-0.5 – 0.3)	0 (-0.4 – 0.4)
Time since previous visit, days	139 (112 – 238)	173 (117 – 288)	137 (117 – 185)	158 (121 – 266)
Deferral at previous visit				
Due to low Hb	203 (12.9)	960 (5.6)	1,683 (20.3)	4,975 (5.0)
Due to other reason than low Hb	14 (0.9)	316 (1.9)	131 (1.6)	4,180 (4.2)
Number of whole blood donations in past 2 years	2 (2 – 3)	2 (2 – 3)	3 (2 – 4)	3 (2 – 4)
Hb intended visit, mmol/L	7.1 (±0.3)	8.3 (±0.6)	7.3 (±0.3)	8.5 (±0.5)

* Data are reported as mean (±SD), n (%) or median (25^{th} – 75^{th} percentile).

† Validation data: Hb < 8.07 mmol/L (= 13.0 g/dL) for men and < 7.45 mmol/L (=12.0 g/dL) for women.
Development data: Hb < 8.4 mmol/L (= 13.5 g/dL) for men and < 7.8 mmol/L (=12.5 g/dL) for women.

Appendix III Logistic regression coefficients of the recalibrated and revised prediction models to calculate the risk of Hb deferral*

Predictor	Value or coding	Beta†			
		Recalibrated model		Revised model	
		Men	Women	Men	Women
Age, per year‡	age	0.01	-0.02	0.01	-0.02
	$(age-22.5)_+$^3		0.78x10-4		0.83x10-4
	$(age-34.6)_+$^3		-3.33x10-4		-3.55x10-4
	$(age-45.0)_+$^3		5.72x10-4		6.05x10-4
	$(age-54.0)_+$^3		-4.22x10-4		-4.43x10-4
	$(age-65.0)_+$^3		1.05x10-4		1.09x10-4
Seasonality	winter‖	0	0	0	0
	spring	0.32	0.16	0.54	0.17
	summer	0.27	0.18	0.18	0.19
	fall	-0.13	-0.09	-0.24	-0.09
Previous Hb level, per mmol/L below sex specific cutoff level§	men: Hb-8.4, women: Hb-7.8	-1.48	-1.33	-1.51	-1.41
Previous Hb level, per mmol/L at or above sex specific cutoff level§	men: Hb-8.4, women: Hb-7.8	-2.46	-2.10	-2.50	-2.22
Delta Hb, per mmol/L equal to or smaller than 0 mmol/L§	delta Hb	0.88	0.66	0.96	0.81
Delta Hb, per mmol/L greater than 0 mmol/L§	delta Hb	0.81	0.87	1.18	0.67
Time since previous visit, per month smaller than 1 year§	months-12	-0.12	-0.10	-0.12	-0.11
Deferral at previous visit	no deferral‖	0	0	0	0
	due to low Hb	-0.40	-0.34	-0.02	-0.66
	due to other reasons than low Hb	-0.63	-0.59	0.22	-0.22
Number of whole blood donations in past 2 years	nr of donations	-0.03	-0.11	-0.03	-0.11
Intercept		-2.46	-0.80	-2.31	-0.62

* With those regression coefficients the linear predictor can be calculated. The risk of Hb deferral equals: risk (Hb deferral) = $1/(1+e^{-\text{linear predictor}})$.

† Beta = regression coefficient, shows the strength and the direction of the variable's influence.

‡ In men the association between age and risk of Hb deferral was linear, in women the association showed several curves that were modeled with a restricted cubic spline function. Negative values of the cubic terms become 0; e.g. $(age - 22.5)_+$ indicates age – 22.5 for positive values, 0 for negative values.

§ Nonlinear association with risk of Hb deferral.

‖ Winter and no deferral are reference categories, other categories: 1 if true, 0 if false.

Chapter 5

High prevalence of sub-clinical iron deficiency in whole blood donors not deferred for low hemoglobin

A. Mireille Baart*†, Paulus A.H. van Noord*, Yvonne Vergouwe†, Karel G.M. Moons†, Dorine W. Swinkels‡, Erwin T. Wiegerinck‡, Wim L.A.M. de Kort*, Femke Atsma*

* Sanquin Research, dept. Donor Studies, Nijmegen, the Netherlands
† Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands
‡ Department of Laboratory Medicine, Laboratory of Genetic, Endocrine and Metabolic Disorders, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands

Accepted for publication in Transfusion

Abstract

Background: Blood donors that meet the hemoglobin (Hb) criteria for donation may have undetected sub-clinical iron deficiency. The aim of this study was to assess the prevalence of sub-clinical iron deficiency in whole blood donors with Hb levels above cutoff levels for donation by measuring zinc protoporphyrin (ZPP) levels. In addition, prevalence rates based on other iron parameters were assessed for comparison.

Study design and methods: The study population comprised 5280 Dutch whole blood donors, who passed the Hb criteria for donation. During donor screening, Hb levels were measured in capillary samples (finger prick), and venous blood samples were taken for measurements of ZPP and other iron parameters. These parameters included ferritin, transferrin saturation, soluble transferrin receptor (sTfR), hepcidin, erythrocyte mean corpuscular volume (MCV) and mean cell hemoglobin (MCH).

Results: With a ZPP cutoff level of ≥100 µmol/mol heme, sub-clinical iron deficiency was present in 6.9% of male donors and in 9.8% of female donors. Based on other iron parameters, iron deficiency was also observed. Prevalence rates ranged from 4.8% (based on transferrin saturation) to 27.4% (based on hepcidin concentration) in men, and from 5.6% (based on sTfR concentration) to 24.7% (based on hepcidin concentration) in women.

Conclusion: Results from this study showed that sub-clinical iron deficiency is prevalent among blood donors that meet the Hb criteria for blood donation, based on ZPP levels and on other iron parameters. This finding needs attention because these donors are at increased risk of developing iron deficiency affecting Hb formation and other cellular processes.

Introduction

Prior to blood donation, donors are screened for hemoglobin (Hb) levels. Donors with relatively low Hb levels are not allowed to donate in order to prevent them from developing overt iron deficiency anemia. In addition, deferral of donors with low Hb levels guarantees that blood units for transfusion meet the required standards for Hb content.[1] An Hb level at or above 8.4 mmol/L (13.5 g/dL) for men and at or above 7.8 mmol/L (12.5 g/dL) for women has been defined as the criterion for blood donation in the European Commission Directive.[2]

Most iron in the body is distributed between three compartments: the storage compartment, the transport compartment, and the functional compartment. During development of iron deficiency, these compartments become subsequently affected, which marks three different consecutive stages of iron deficiency: iron depletion, iron-deficient erythropoiesis and iron deficiency anemia.[3] A negative iron balance first decreases body iron reserves. When iron becomes absent in the storage compartment, but the functional iron compartment is not affected, iron depletion exists. Once iron stores are depleted, the amount of iron in the transport compartment starts to decrease. The stage of iron-deficient erythropoiesis has been reached; the iron supply to the erythropoietic bone marrow is insufficient for erythropoiesis. However, Hb levels are still normal. This condition is what we call sub-clinical iron deficiency. When finally the functional compartment is affected and the iron supply is no longer sufficient to produce a normal amount of Hb, iron deficiency anemia becomes apparent. So Hb levels are only low in this advanced stage of iron deficiency. As a consequence, it may occur that donors with appropriate Hb levels have depleted iron stores[4] or even iron deficient erythropoiesis.[3,5] Consequently, these conditions remain undetected by Hb screening. Especially these donors are at high risk of developing iron deficiency anemia after a blood donation and are more likely to be deferred at their next visit to the blood collection center.

There are several iron parameters available to assess iron depletion or iron-deficient erythropoiesis. Iron depletion can be assessed by measuring serum ferritin levels.[6] Tests for the diagnosis of iron deficient erythropoiesis include measurement of plasma iron,[3] total iron binding capacity or transferrin,[3] transferrin saturation, soluble transferrin receptor (sTfR) concentration,[7] the sTfR index (sTfR divided by log-transformed ferritin values)[8,9] and zinc protoporphyrin (ZPP).[10] Another iron parameter with which sub-clinical iron deficiency can be assessed is the recently discovered iron regulatory protein hepcidin.[11,12] Each of the above mentioned tests has its own advantages and disadvantages. ZPP is measured by an automated technology and its attractive features are its ability to perform immediate point-of-care assays and its relative low price. It is therefore a useful test for donor screening.

ZPP is formed during heme synthesis under conditions of iron deficiency.[13] In the last step of heme synthesis a ferrous ion is incorporated into protoporphyrin IX, which is catalyzed

by the enzyme ferrochelatase. This enzyme also catalyzes the reaction for the incorporation of zinc into protoporphyrin IX. When iron levels are low, more zinc instead of iron is incorporated, the ratio of zinc to iron incorporation increases, and ZPP accumulates in the red blood cells. ZPP levels start to increase in the early stage of iron deficient erythropoiesis[10] and thus ZPP measurement can detect iron deficiency in the sub-clinical state.
The main objective of this study was to estimate the prevalence of sub-clinical iron deficiency by measuring ZPP levels, in whole blood donors who passed the Hb screening test for blood donation. In addition, the distribution of other iron parameters (i.e. ferritin, transferrin saturation, sTfR, hepcidin, erythrocyte mean corpuscular volume (MCV) and mean cell hemoglobin (MCH)) was assessed in a sub group of donors. Prevalence rates of iron deficiency based on these other parameters were also assessed and compared with the prevalence based on ZPP levels.

Methods

Donors

Dutch whole blood donors who passed the Hb screening test for blood donation (an Hb level at or above 8.4 mmol/L (13.5 g/dL) for men and at or above 7.8 mmol/L (12.5 g/dL) for women) were eligible for inclusion in the present study. We defined whole blood donors as blood donors whose last two donations are whole blood donations. All these whole blood donors also fulfilled the other eligibility criteria for donation according to the European Commission Directive.[2] These include an age between 18 and 70 years, a weight of at least 50 kg, a systolic blood pressure between 90 and 180 mmHg, and a diastolic blood pressure between 50 and 100 mmHg. Furthermore, donors have to fill out an eligibility questionnaire to identify known medical conditions and perilous behaviour. All included donors passed the specific health and lifestyle criteria. In addition, criteria with relation to the time interval between donations and donation frequency were also met. In the Netherlands, the minimum time interval between two donations is 56 days, and the maximum number of donations allowed per year is five for men and three for women.
Data collection took place in two blood collection centers, between 1 December 2008 and 30 November 2009 in center 1 and between 1 August 2009 and 31 July 2010 in center 2. A total of 5280 donors (2897 men, 2383 women) were included in the study.

Design

At each visit in the inclusion period, venous blood samples for measurement of ZPP levels were collected from all donors. In a sub group of donors additional iron parameters were assessed, including ferritin, transferrin saturation, sTfR, hepcidin, MCV and MCH. For the selection of donors in the sub group, the total 5280 donors were stratified by sex and a ZPP level category. The five ZPP categories were as follows (in µmol/mol heme): <40,

40-59, 60-79, 80-99, ≥100. Next, from each of the 10 strata 40 donors were randomly selected into the sub group. In one stratum only 30 donors were available, and so the total sub group consisted of 390 donors.

Laboratory methods

Hb levels were routinely measured during donor screening in finger prick capillary samples using a photometer (HemoCue, Angelholm, Sweden). In addition, two venous blood samples were collected prior to blood donation, one blood sample was collected in lithium heparin and another in EDTA. ZPP levels were measured with a hematofluorometer (Model 206D, Aviv Biomedical, Lakewood, NJ) on the next day in the samples collected in lithium heparin. Before ZPP measurement red blood cells were washed with phosphate-buffered saline to minimize the possible influence of other fluorescent elements in plasma.[14]

In the sub group of 390 donors, additional iron parameters were measured. Erythrocyte MCV and MCH levels were measured on a hematology analyzer (Model XT-1800i, Sysmex, Kobe, Japan) within eight hours after blood collection in the whole blood samples collected in EDTA. The MCV and MCH measurements, as well as the ZPP measurements, were performed in the Sanquin laboratory, which is ISO 9001 certified and fulfills the GMP guidelines.

Plasma from samples collected in lithium heparin was frozen at -80 ^{0}C on the day after collection. Between March and August 2011, we measured plasma iron, ferritin, transferrin, sTfR and hepcidin. The sTfR index (sTfR divided by log-transformed ferritin values)[8,9] was calculated. In addition to the described iron parameters, C-reactive protein (CRP) was measured in order to assess inflammation that could influence ferritin levels. Plasma concentrations of iron, transferrin and CRP were determined on an immunology analyser (Architect C16000, Abbott Diagnostics, Abbott Park, IL). The total iron binding capacity (TIBC, µmol/L) was calculated from transferrin (g/L) multiplied by 25. The percent transferrin saturation was calculated as 100 times the plasma iron concentration divided by the TIBC. Ferritin levels were quantified with a solid-phase 2-site chemiluminescent immunometric assay (Immulite 2500, Siemens Medical Solutions, Erlangen, Germany). Plasma concentrations of sTfR were measured immunonephelometrically on a nephelometer (BN II System, Dade-Behring, Deerfield, IL). Plasma hepcidin-25 measurements (further in the text denoted as "hepcidin") were performed by a combination of weak cation exchange chromatography and time-of-flight mass spectrometry as described previously.[15,16] Hepcidin concentrations were expressed as nmol/L (1 nmol/L = 2.789 µg/mL) and the lower limit of detection of this method was 0.5 nmol/L; ranges for the coefficients of variation were 2.2-3.7% (intra-run) and 3.9-9.1% (inter-run). For this study, hepcidin concentrations below the lower limit of detection were included as 0.25 nmol/L for the statistical calculations. Measurements of plasma iron, ferritin, transferrin, sTfR, hepcidin and CRP were performed at the Department of

Laboratory Medicine of the Radboud University Nijmegen Medical Centre, the Netherlands. This laboratory has received accreditation by the CCKL (the Dutch Accreditation Board for Medical Laboratories) according to the accreditation criteria for medical laboratories as laid down in EN/ISO 15189 and specified in the CCKL code of practice for the implementation of a quality system. We used reference ranges for iron parameters based on manuals and standard operating procedures from equipment used in this laboratory.
A total of 24 donors had CRP levels above 10 mg/L and were excluded in the analyses with additional iron parameters, as to prevent false-high levels of ferritin. So, finally 366 donors were used to study the distribution of additional iron parameters in blood donors.

Statistical analysis

Descriptive analyses were performed. Donors were stratified by sex and the predefined ZPP categories (as for the selection of donors into the sub group). For each stratum mean values and standard deviation (SD) of the measured iron parameters were calculated.
In the first instance, the prevalence of sub-clinical iron deficiency was calculated as the proportion of donors with ZPP levels ≥100 µmol/mol heme. This cutoff level is recommended by the hematofluorometer firm AVIV. Besides, the prevalence of sub-clinical iron deficiency based on reference values for the other iron parameters was assessed. With the use of weight factors the measured prevalence in the sub group of 366 donors could be extrapolated to the total group of 5280 donors. Hereto, the sampling fraction was calculated for each stratum by dividing per stratum the number of donors in the sub group by the number of donors in the total population. Sampling fractions ranged from 0.03 for men with ZPP levels of 40-59 µmol/mol heme to 0.28 for women with ZPP levels <40 µmol/mol heme. Weight factors for each stratum were then calculated as 1 divided by the sampling fractions. Next, prevalence rates in the total population were calculated using the weight factors per stratum.
A total of 3760 donors (71%), visited the blood collection center again within half a year after the inclusion period. For these donors, the Hb level at the subsequent visit was also examined per ZPP stratum. This was done to examine if increased ZPP levels, indicating sub-clinical iron deficiency, are associated with future overt clinical iron deficiency in whole blood donors.
Statistical analyses were performed with SPSS, Version 19, SPSS, Inc., Chicago, IL.

Results

Characteristics of the total donor population (n=5280) are presented in Table 1A and characteristics of the sub group in which additional iron parameters were measured (n=366) are presented in Table 1B. In the total population, the mean age in men was 51 (SD=11) years and in women 46 (SD=13) years. Mean Hb levels were 9.3 (SD=0.6, range 8.4

Table 1A Donor characteristics*

Characteristic	Number	Total (n =5280)	Men (n=2897)	Women (n=2383)
Age, years	5280	49 (±12)	51 (±11)	46 (±13)
Height, cm	5260	175 (±9)	180 (±11)	167 (±10)
Weight, kg	5279	80 (±14)	86 (±12)	73 (±13)
BMI, kg/m^2	5259	26 (±4)	27 (±3)	26 (±4)
Hb, mmol/L	5280	9.0 (±0.7)	9.3 (±0.6)	8.6 (±0.6)
ZPP, µmol/mol heme	5280	66 (±24)	63 (±24)	69 (±25)

* Data are reported as mean (±SD).

Table 1B Donor characteristics of sub group*

Characteristic	Number	Total (n =366)	Men (n=190)	Women (n=176)
Age, years	366	50 (±12)	52 (±11)	47 (±12)
Height, cm	366	174 (±9)	180 (±7)	168 (±6)
Weight, kg	366	80 (±14)	87 (±11)	72 (±13)
BMI, kg/m^2	366	26 (±4)	27 (±3)	26 (±4)
Hb, mmol/L	366	8.9 (±0.7)	9.3 (±0.6)	8.6 (±0.6)
ZPP, µmol/mol heme	366	74 (±35)	74 (±37)	74 (±32)
Ferritin, µg/L	356	29 (14 – 50)	31 (15 – 60)	27 (14 – 46)
Transferrin saturation, %	361	24 (16 – 32)	24 (17 – 33)	23 (16 – 32)
sTfR, mg/L	359	1.26 (1.06 – 1.63)	1.32 (1.09 – 1.89)	1.22 (1.04 – 1.47)
sTfR index, mg/µg	351	0.8 (0.6 – 1.4)	0.8 (0.6 – 1.6)	0.7 (0.8 – 1.2)
Hepcidin, nmol/L	366	1.2 (0.25 – 3.0)	1.1 (0.25 – 2.9)	1.4 (0.25 – 3.4)
MCV, fL	366	88 (±5)	88 (±5)	89 (±5)
MCH, fmol	366	1.82 (±0.16)	1.81 (±0.18)	1.83 (±0.13)

* Data are reported as mean (±SD) or median (25th – 75th percentile).

– 11.4) mmol/L for men and 8.6 (SD=0.6, range 7.8 – 11.5) mmol/L for women. Mean ZPP levels for men were 63 (SD=24 range 19 - 260) µmol/mol heme and for women 69 (SD=25, range 17 - 254) µmol/mol heme. Demographic characteristics of the sub group of 366 donors were comparable. Due to stratification, mean ZPP levels in the sub group were higher and mean Hb levels lower.

The number of donors in each ZPP stratum and mean Hb levels per stratum are presented in Table 2. The cumulative percentage of donors shows for each ZPP cutoff level the

Table 2 Distribution of donors over different ZPP strata and mean Hb levels per ZPP stratum

ZPP (µmol/mol heme)	Number (%)	Cumulative %	Hb (mmol/L*) Mean (±SD)
Men			
≥100	201 (6.9)	6.9	8.8 (±0.5)
80-99	292 (10.1)	17.0	9.2 (±0.6)
60-79	876 (30.2)	47.3	9.3 (±0.6)
40-59	1286 (44.4)	91.6	9.5 (±0.6)
<40	242 (8.4)	100.0	9.5 (±0.6)
Women			
≥100	234 (9.8)	9.8	8.3 (±0.5)
80-99	379 (15.9)	25.7	8.4 (±0.5)
60-79	834 (35.0)	60.7	8.5 (±0.5)
40-59	838 (35.2)	95.9	8.7 (±0.6)
<40	98 (4.1)	100.0	8.8 (±0.6)

* Conversion factor for g/dL: 1.61.

percentage of donors with sub-clinical iron deficiency. With a ZPP cutoff level of 100 µmol/mol heme for example, sub-clinical iron deficiency was present in 6.9% of male donors and in 9.8% of female donors.

Table 3 shows the distribution of additional iron parameters per ZPP stratum in the sub group of 366 donors. The table clearly shows that higher ZPP levels are associated with lower Hb levels, ferritin levels, transferrin saturations, hepcidin levels, erythrocyte MCV levels and MCH levels, and higher sTfR concentrations and sTfR indices in both men and women.

Table 4 shows the percentage of donors with sub-clinical iron deficiency based on the other iron parameters. Prevalence rates ranged from 4.8% (based on transferrin saturation) to 27.4% (based on hepcidin concentration) in men, and from 5.6% (based on sTfR concentration) to 24.7% (based on hepcidin concentration) in women. For the analysis with the additional iron parameters we excluded 24 donors (6%) because of high CRP levels. For comparison, we performed the same analysis in donors with high CRP levels, and we found comparable results.

When looking at Hb values at the subsequent visit, mean Hb levels were 0.1 mmol/L lower than mean Hb levels at the visit of ZPP measurement in male donors in the stratum with ZPP levels ≥100 µmol/mol heme and in female donors in the strata with ZPP levels 80-99 and ≥100 µmol/mol heme (Table 5).

Table 3 Distribution of iron parameters in donors, per ZPP stratum*

ZPP (µmol/mol heme)	Number	Hb (mmol/L)	Ferritin (µg/L)	Transferrin saturation (%)	sTfR (mg/L)	sTfR index (mg/µg)	Hepcidin (nmol/L)	MCV (fL)	MCH (fmol)
Men									
≥100	40	8.8 (±0.4)	13 (9 – 18)	12 (9 – 22)	2.11 (1.77 – 2.71)	1.9 (1.4 – 2.6)	0.25 (0.25 – 0.25)	83 (±6)	1.67 (±0.13)
80-99	37	9.2 (±0.6)	16 (12 – 32)	22 (16 – 27)	1.57 (1.25 – 1.97)	1.3 (0.9 – 1.7)	0.25 (0.25 – 1.2)	87 (±4)	1.78 (±0.10)
60-79	37	9.5 (±0.6)	37 (24 – 61)	23 (21 – 34)	1.43 (1.08 – 1.72)	0.8 (0.7 – 1.2)	1.7 (0.6 – 4.5)	88 (±4)	1.85 (±0.10)
40-59	38	9.5 (±0.5)	41 (29 – 79)	27 (22 – 35)	1.17 (0.98 – 1.27)	0.7 (0.5 – 0.8)	2.6 (1.2 – 3.7)	91 (±4)	1.84 (±0.31)
<40	38	9.4 (±0.5)	60 (41 – 96)	33 (26 – 40)	1.08 (0.97 – 1.21)	0.6 (0.5 – 0.7)	2.7 (1.1 – 4.0)	90 (±4)	1.90 (±0.08)
Women									
≥100	35	8.4 (±0.5)	11 (8 – 16)	16 (9 – 23)	1.65 (1.28 – 2.09)	1.6 (1.2 – 2.3)	0.25 (0.25 – 1.1)	84 (±5)	1.70 (±0.11)
80-99	38	8.4 (±0.5)	22 (14 – 33)	20 (15 – 27)	1.27 (1.14 – 1.44)	0.9 (0.8 – 1.2)	1.2 (0.25 – 2.5)	88 (±5)	1.81 (±0.12)
60-79	37	8.6 (±0.5)	22 (14 – 35)	22 (18 – 32)	1.24 (1.09 – 1.46)	0.9 (0.7 – 1.1)	1.0 (0.25 – 2.2)	89 (±3)	1.85 (±0.08)
40-59	38	8.8 (±0.7)	37 (28 – 55)	27 (21 – 35)	1.03 (0.94 – 1.25)	0.7 (0.6 – 0.8)	2.4 (1.5 – 4.3)	92 (±3)	1.90 (±0.07)
<40	28	8.7 (±0.7)	42 (32 – 72)	31 (23 – 40)	1.09 (0.87 – 1.27)	0.6 (0.5 – 0.8)	2.9 (1.3 – 5.7)	94 (±5)	1.92 (±0.11)

* Data are reported as mean (±SD) or median (25^{th} – 75^{th} percentile).

Table 4 Percentages of donors with iron deficiency based on different iron parameters

Iron parameter	Reference range*	Iron deficiency in men (%)	Iron deficiency in women (%)
ZPP, µmol/mol heme	< 100	6.9	9.8
Ferritin, µg/L	Men: 15 – 280 Women premenopausal: 6 – 80 Women postmenopausal: 15 – 190	14.3	8.9
Transferrin saturation, %	10 – 30	4.8	8.4
sTfR, mg/L	0.76 – 1.76	19.3	5.6
sTfR index, mg/µg	Not available		
Hepcidin, nmol/L	Men: 0.5 – 14.7 Women premenopausal: 0.5 – 12.3 Women postmenopausal: 0.5 – 15.6†	27.4†	24.7†
MCV, fL	85 – 100	15.7	11.8
MCH, fmol	1.70 – 2.20	9.5	6.3

* Reference ranges are based on manuals and standard operating procedures from the laboratory equipment and reagents used in this study (see methods).

† The percentage of donors with iron deficiency is based on setting the lower reference value for hepcidin at 0.5 nmol/L. Note that the lower reference value as assessed from the general population (based on < 2.5th percentile) lies below 0.5 nmol/L (below the lower limit of detection).[29] In other words, when using a value <0.5 nmol/L as the lower reference value, the percentage of donors with iron deficiency will be lower.

Table 5 Mean Hb levels at the subsequent visit per ZPP stratum*

ZPP (µmol/mol heme)	Number (%)	Hb at visit of ZPP measurement (mmol/L†)	Hb at subsequent visit (mmol/L†)	Hb-difference‡ (mmol/L†)	Time interval (days)
Men					
≥100	170 (7.8)	8.8 (±0.4)	8.7 (±0.7)	-0.1 (-0.5 – 0.3)§	78 (70 – 112)
80-99	239 (11.0)	9.1 (±0.6)	9.1 (±0.7)	0 (-0.4 – 0.4)	78 (70 – 118)
60-79	662 (30.5)	9.3 (±0.6)	9.3 (±0.7)	0 (-0.3 – 0.4)	91 (70 – 145)
40-59	932 (42.9)	9.4 (±0.6)	9.5 (±0.7)	0 (-0.3 – 0.4)	92 (71 – 175)
<40	171 (7.9)	9.4 (±0.6)	9.5 (±0.6)	0 (-0.3 – 0.4)	98 (70 – 175)
Women					
≥100	162 (10.2)	8.2 (±0.4)	8.1 (±0.6)	-0.1 (-0.5 – 0.1)§	140 (119 – 182)
80-99	253 (16.0)	8.4 (±0.5)	8.4 (±0.6)	-0.1 (-0.4 – 0.2)§	154 (119 – 189)
60-79	563 (35.5)	8.5 (±0.5)	8.5 (±0.6)	0 (-0.4 – 0.4)	147 (119 – 187)
40-59	549 (34.6)	8.7 (±0.6)	8.7 (±0.6)	0 (-0.3 – 0.4)	154 (119 – 183)
<40	59 (3.7)	8.8 (±0.7)	8.9 (±0.6)	0.1 (-0.3 – 0.5)	168 (133 – 195)

* Data are reported as n (%), mean (SD), or median (25^{th} – 75 th percentile).

† Conversion factor for g/dL: 1.61.

‡ Hb at subsequent visit minus Hb at visit of ZPP measurement.

§ Significant, $p < 0.05$.

Discussion

Sub-clinical iron deficiency is prevalent among blood donors that meet the Hb criteria for blood donation. In our Dutch donor population, sub-clinical iron deficiency was present in 6.9% of non-deferred male donors and in 9.8% of non-deferred female donors, based on a ZPP cutoff level of 100 µmol/mol heme. Also based on other iron parameters, we found sub-clinical iron deficiency in donors meeting the Hb criteria for donation.

The existence of iron depletion or iron deficient erythropoiesis in blood donors has been reported before.[4,5,17] However, estimated prevalence rates varied widely. The REDS-II Donor Iron Status Evaluation (RISE) study[17] investigated the prevalence of iron depletion (defined as ferritin levels <12 ng/ml) and of iron deficient erythropoiesis (defined as log(sTfR/ferritin) ≥2.07) in whole blood donors that met the Hb criteria for blood donation. Hb cutoff levels for donation used in the RISE study were 12.5 g/dL (7.8 mmol/L) for both men and women. Based on the used definitions of iron deficiency, the RISE study showed that in male frequent donors (at least 3 donations in the past year) the prevalence of iron depletion and iron deficient erythropoiesis was 47% and 18% respectively. Among female frequent donors (at least 2 donations in the past year) these prevalence rates were 62% and 27%, respectively. The observed prevalence rates in the RISE study are different from our prevalence rates. This may be due to the use of different iron parameters with different reference ranges, and because the Hb criterion for donation for males was lower and the donation frequency of both men and women was higher than those in the current study.

Detection of sub-clinical iron deficiency in blood donors is important not only to prevent donors from becoming anemic after subsequent blood donations, but also because iron deficiency affects various other cellular processes as iron is an important element of many enzymes. For example, early studies suggested that iron deficiency also affects DNA synthesis,[18,19] the immune system,[20] and energy metabolism through impaired mitochondrial electron transport.[21,22] After detection of sub-clinical iron deficiency interventions such as a dietary advice, iron fortification and prolonged donation intervals could help to protect donors from developing iron deficiency anemia or other iron deficiency related disorders after blood donation.

In order to investigate if increased ZPP levels are associated with future overt clinical iron deficiency in whole blood donors, we examined the Hb level at the subsequent visit after the visit of ZPP measurement. Mean Hb levels at the subsequent visit were lower than mean Hb levels at the visit of ZPP measurement in donors with high ZPP levels. The time interval between the two subsequent visits was a little shorter in these donors, although they fulfilled the donation criterion of a minimal donation interval of 56 days. For other iron parameters we observed comparable patterns: values indicative of a low iron status were associated with lower Hb levels at the subsequent visit (data not shown). These results indicate that Hb levels in donors with sub-clinical iron deficiency do not fully return to the normal value within 56 days after a blood donation. Consequently, these donors are

at high risk of developing iron deficiency anemia after a blood donation and subsequently being deferred at the next visit to the blood collection centre. Prolongation of the donation interval for donors with sub-clinical iron deficiency might be useful to protect them from developing iron deficiency anemia and to prevent them from being deferred for a subsequent donation.

In our study we showed that by using Hb measurement as a screening tool in blood donors, iron deficiency cannot be detected in the sub-clinical state. Therefore, it is imperative to search for alternative tests. Several iron parameters have been proposed as suitable alternative methods to assess sub-clinical iron deficiency in blood donors. In a previous study it was concluded that routine ferritin measurement in blood donors allowed an optimized management of donors with (sub-clinical) iron deficiency and prevention of development of overt anemia.[23] However, an important drawback of ferritin measurement is its susceptibility to increase in case of inflammation. As a result, false positive results may occur. STfR has also been proposed as a screening tool for iron deficiency in blood donors, and it was concluded to be a better screening test for iron deficiency than ferritin.[24] However, this immunochemical test is rather costly and therefore less suitable as a screening tool for the iron status in large numbers of blood donors.

In the past, ZPP measurement has already been proposed as a useful method in donor screening, because of it's good diagnostic value.[25,26] As stated before, ZPP measurements are automated in a point-of-care apparatus and therefore relatively cheap and easily applicable in clinical blood bank practice. Therefore, we argue that measurement of ZPP is most appropriate as an additional test for donor screening.

In order to confirm the presence of a substantial proportion of sub-clinical iron deficiency as measured by ZPP among blood donors who passed the Hb screening test for blood donation, we performed additional analyses in a sub group of blood donors of whom we also had other iron parameters at our disposal. We investigated the distribution of these other iron parameters, and observed if we also found donors to be sub-clinically iron deficient based on the other parameters. Results showed clear associations between the different iron parameters. Moreover, for each iron parameter, a certain proportion of blood donors was sub-clinical iron deficient. This strengthens our conclusion of the presence of sub-clinical iron deficiency in blood donors who passed the Hb criteria for blood donation.

Finally, this study has some strengths and limitations. Strengths are the availability of a large number of ZPP measurements, and the availability of six other iron parameters in a sub group of donors with which we could confirm our conclusion about the prevalence of sub-clinical iron deficiency.

A limitation of the current study is the use of reference ranges for iron parameters that have not been related to immediate or future functional outcomes, simply because they are not available. Moreover, reference ranges used in this study and other studies are assessed in different reference populations that might not be comparable for the various aspects that are iron related, such as age, sex and nutritional status. Furthermore, ZPP, sTfR

and hepcidin measurements lack as yet standardization,[27,28] implying that results of these tests and the reference ranges differ between methodologies and laboratories. Without universal and functional reference ranges the exact scale of the existence of sub-clinical iron deficiency in blood donors is unclear. Prevalence rates may be under- or overestimated. The uncertainties concerning reference ranges make it difficult to compare the results from different studies in which the same iron parameters was used. Moreover, within the same study, prevalence rates may vary when based on different iron parameters, which we also observed in this study. Nevertheless, the results of our study do provide indications for the presence of sub-clinical iron deficiency among whole blood donors with Hb levels above cutoff levels for donation. The observed differences in prevalence of iron deficiency based on different iron parameters in this study confirm the need for universal and functional standardized decision limits for the diagnosis of (sub-clinical) iron deficiency in blood donors, as mentioned by others.[27,28]

In conclusion, our data suggest that sub-clinical iron deficiency is prevalent among blood donors that meet the Hb criteria for blood donation, based on ZPP levels and on other iron parameters. This finding needs attention because these donors are at increased risk of developing iron deficiency affecting Hb formation and many other cellular processes.

Acknowledgements

The authors thank Giny Hetterscheit and Mirianne van Gerwen for their help with the practical work and Ingrid Veldhuizen for her contribution in fund raising.

Karel G.M. Moons gratefully acknowledges financial contribution for all his research by the Netherlands Organisation for Scientific Research (project 9120.8004 and 918.10.615).

Reference List

1. European Directorate for the Quality of Medicines and HealthCare. Guide to the preparation, use and quality assurance of blood components, European Committee (partial agreement) on Blood Transfusion (CD-P-TS), Recommendation No. R(95) 15, 15th edition. Strasbourg: Council of Europe; 2009, p 1899.
2. The Commission of the European Communities. European Commissian Directive 2004/33/EC. 2004.
3. Skikne B, Lynch S, Borek D, Cook J. Iron and blood donation. Clin.Haematol. 1984;13:271-87.
4. Finch CA, Cook JD, Labbe RF, Culala M. Effect of blood donation on iron stores as evaluated by serum ferritin. Blood 1977;50:441-7.
5. Simon TL, Garry PJ, Hooper EM. Iron stores in blood donors. JAMA 1981;245:2038-43.
6. Cook JD. Clinical evaluation of iron deficiency. Semin.Hematol. 1982;19:6-18.
7. Kohgo Y, Niitsu Y, Kondo H, Kato J, Tsushima N, Sasaki K, Hirayama M, Numata T, Nishisato T, Urushizaki I. Serum transferrin receptor as a new index of erythropoiesis. Blood 1987;70:1955-8.
8. Punnonen K, Irjala K, Rajamaki A. Serum transferrin receptor and its ratio to serum ferritin in the diagnosis of iron deficiency. Blood 1997;89:1052-7.
9. Beguin Y. Soluble transferrin receptor for the evaluation of erythropoiesis and iron status. Clin.Chim.Acta 2003;329:9-22.
10. Labbe RF, Rettmer RL. Zinc protoporphyrin: a product of iron-deficient erythropoiesis. Semin.Hematol. 1989;26:40-6.
11. Nemeth E, Ganz T. The role of hepcidin in iron metabolism. Acta Haematol. 2009;122:78-86.
12. Kemna EH, Tjalsma H, Willems HL, Swinkels DW. Hepcidin: from discovery to differential diagnosis. Haematologica 2008;93:90-7.
13. Lamola AA, Yamane T. Zinc protoporphyrin in the erythrocytes of patients with lead intoxication and iron deficiency anemia. Science 1974;186:936-8.
14. Hastka J, Lasserre JJ, Schwarzbeck A, Strauch M, Hehlmann R. Washing erythrocytes to remove interferents in measurements of zinc protoporphyrin by front-face hematofluorometry. Clin.Chem. 1992;38:2184-9.
15. Kroot JJ, Laarakkers CM, Geurts-Moespot AJ, Grebenchtchikov N, Pickkers P, van Ede AE, Peters HP, van Dongen-Lases E, Wetzels JF, Sweep FC, et al. Immunochemical and mass-spectrometry-based serum hepcidin assays for iron metabolism disorders. Clin.Chem. 2010;56:1570-9.
16. Hepcidinanalysis.com. Homepage. [cited 2011 Sep 12]. Available from: URL: http://www.hepcidinanalysis.com.
17. Cable RG, Glynn SA, Kiss JE, Mast AE, Steele WR, Murphy EL, Wright DJ, Sacher RA, Gottschall JL, Tobler LH, Simon TL; NHLBI Retrovirus Epidemilogy Donor Study-II. Iron deficiency in blood donors: the REDS-II Donor Iron Status Evaluation (RISE) study. Transfusion 2012;52:702-11.
18. Hoffbrand AV, Ganeshaguru K, Hooton JW, Tattersall MH. Effect of iron deficiency and desferrioxamine on DNA synthesis in human cells. Br.J.Haematol. 1976;33:517-26.
19. Furukawa T, Naitoh Y, Kohno H, Tokunaga R, Taketani S. Iron deprivation decreases ribonucleotide reductase activity and DNA synthesis. Life Sci. 1992;50:2059-65.
20. Ward RJ, Crichton RR, Taylor DL, Della CL, Srai SK, Dexter DT. Iron and the immune system. J.Neural Transm. 2011;118:315-28.
21. Finch CA, Miller LR, Inamdar AR, Person R, Seiler K, Mackler B. Iron deficiency in the rat. Physiological and biochemical studies of muscle dysfunction. J.Clin.Invest 1976;58:447-53.
22. Ackrell BA, Maguire JJ, Dallman PR, Kearney EB. Effect of iron deficiency on succinate- and NADH-ubiquinone oxidoreductases in skeletal muscle mitochondria. J.Biol.Chem. 1984;259:10053-9.
23. O'Meara A, Infanti L, Stebler C, Ruesch M, Sigle JP, Stern M, Buser A. The value of routine ferritin measurement in blood donors. Transfusion 2011;51:2183-8.
24. Flesland O, Eskelund AK, Flesland AB, Falch D, Solheim BG, Seghatchian J. Transferrin receptor in serum. A new tool in the diagnosis and prevention of iron deficiency in blood donors. Transfus.Apher.Sci. 2004;31:11-6.
25. Schifman RB, Rivers SL, Finley PR, Thies C. RBC zinc protoporphyrin to screen blood donors for iron deficiency anemia. JAMA 1982;248:2012-5.
26. Harthoorn-Lasthuizen EJ, Lindemans J, Langenhuijsen MM. Zinc protoporphyrin as screening test in female blood donors. Clin.Chem. 1998;44:800-4.

27. Lynch S. Case studies: iron. Am.J.Clin.Nutr. 2011;94:673S-8S.
28. Kroot JJ, Kemna EH, Bansal SS, Busbridge M, Campostrini N, Girelli D, Hider RC, Koliaraki V, Mamalaki A, Olbina G, et al. Results of the first international round robin for the quantification of urinary and plasma hepcidin assays: need for standardization. Haematologica 2009;94:1748-52.
29. Galesloot TE, Vermeulen SH, Geurts-Moespot AJ, Klaver SM, Kroot JJ, van TD, Wetzels JF, Kiemeney LA, Sweep FC, den HM, et al. Serum hepcidin: reference ranges and biochemical correlates in the general population. Blood 2011;117:e218-e225.

Chapter 6

Zinc protoporphyrin levels have added value in the prediction of low hemoglobin deferral in whole blood donors

A. Mireille Baart*†, Wim L.A.M. de Kort*, Karel G.M. Moons†, Femke Atsma*, Yvonne Vergouwe†

* Sanquin Research, dept. Donor Studies, Nijmegen, the Netherlands
† Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands

Accepted for publication in Transfusion

Abstract

Background: Increased zinc protoporphyrin (ZPP) levels can indicate iron deficiency, and may be predictive for low hemoglobin (Hb) deferral in blood donors. Prediction models for Hb deferral in whole blood donors have already been developed. In this study, we examined if addition of ZPP to these prediction models improves risk estimation of Hb deferral.

Study design and methods: This study included 4598 Dutch whole blood donors. Information on ZPP levels measured at the previous visit was added to the existing prediction models to estimate the risk of Hb deferral. Models were compared using the following measures: concordance(c)-statistic, continuous net reclassification improvement (NRI), and clinical net benefit (NB).

Results: Seventy-six male donors (2.9%) and 69 female donors (3.5%) were deferred because of a low Hb level. Previous ZPP level was associated with risk of Hb deferral (OR for interquartile range of previous ZPP level, men: 2.0 (95% CI 1.7 – 2.3); women: 2.2 (95% CI 1.9 – 2.4)) in a multivariable risk model. Addition of ZPP into the models resulted in an increase of the c-statistic from 0.93 to 0.94 for men, and from 0.80 to 0.85 for women. The added value of ZPP was confirmed by measures of clinical usefulness. NRI for men was 0.42, for women 0.56. At relevant threshold probabilities between 10% and 15%, NB was higher for models considering ZPP.

Conclusion: This study shows that ZPP measurements obtained at the previous visit may have added value in the risk prediction of Hb deferral in whole blood donors.

Introduction

Prior to blood donation, the iron status is assessed in blood donors. Most commonly, the iron status is assessed by measuring hemoglobin (Hb) levels. Donors with low Hb levels are deferred from donation to protect them from developing iron deficiency anemia after a donation. Besides, deferral of donors with low Hb levels guaranties that blood units for transfusion meet the required standards for Hb content.[1]

A substantial number of donors is deferred from donation because of low Hb levels. Deferrals are demoralizing for donors. Furthermore, the risk of donor lapse is increased,[2-4] although the donor's Hb level could actually be high enough at the time of the next invitation to donate.

In contrast, donors that meet the Hb criterion for blood donation may have depleted iron stores[5] or even iron deficient erythropoiesis.[6,7] These conditions are early stages in the development of iron deficiency and may proceed to iron deficiency anemia. As these conditions remain undetected by Hb screening, donors in these conditions might be at increased risk of developing anemia after a donation and subsequently being deferred at the next visit to the blood bank.

With respect to the iron status, some donors need more time to recover after a donation than other donors until they can donate blood again, and the guidelines for the inter-donation intervals may not be safe for each individual donor. Recently, prediction models for Hb deferral have been developed using data from all Dutch whole blood donors. These models predict the risk of Hb deferral for individual donors at the moment of possible invitation for a next donation. Based on the individual risk predictions, decisions can be made about whether to invite a donor for a next donation, or whether it is better to postpone the invitation, or to apply another intervention such as a dietary advice. Eventually, prediction of Hb deferral may be valuable to decrease the number of donor deferrals for low Hb levels.

Zinc protoporphyrin (ZPP) is another indicator of the iron status.[8] ZPP levels start to increase in the early stage of iron deficient erythropoiesis and thus ZPP measurement can detect iron deficiency in an early stage before Hb levels decrease. ZPP is formed during heme synthesis in case of iron deficiency. When iron levels are low, more zinc rather than iron is incorporated into protoporphyrin IX during heme synthesis. This results in the formation of more ZPP and less heme. As a result, ZPP accumulates in the red blood cells.[9] Increased ZPP levels may be predictive for low Hb levels and ZPP levels may therefore have added value in the prediction of Hb deferral. Although other indicators for assessment of the iron status are available, such as ferritin levels[10] and the soluble transferrin receptor concentration,[11] we focus on ZPP, because ZPP measurement seems appropriate as a screening test for blood donors. The measurement can be performed in a finger stick capillary sample as a point-of-care test and has a relative low price.

The aim of this study was to assess the added value of ZPP levels for the prediction of Hb deferral in the previously developed prediction models. We investigated if addition of ZPP levels to these prediction models could further improve risk estimation and stratification of Hb deferral.

Donors and methods

Donors

The present study was performed among a cohort of Dutch whole blood donors who visited one of two blood collection centers in an inclusion period (1 December 2008 – 30 November 2009 in center 1 and 1 August 2009 – 31 July 2010 in center 2). We defined whole blood donors as blood donors whose last two donations are whole blood donations. All these donors fulfilled the Dutch criteria for donation with relation to the time interval between donations and donation frequency. The minimum time interval between two donations is 56 days, and the maximum number of donations allowed per year is five for men and three for women. At each visit in the inclusion period, blood samples for measurement of ZPP levels were collected.

To be included in the study, donors should have a subsequent visit after the visit of ZPP measurement. This subsequent visit could be in the inclusion period or within six months after this period. We call this subsequent visit the "intended visit", and the Hb level measured at this visit was used as outcome in the analyses.

Donors were excluded if one of the two whole blood donations prior to the intended visit took place longer than four years ago. This was required to collect sufficient information on donation history. Donors with uncertainties in their history, e.g. whether or not a donation took place, and donors with an unknown Hb level at the intended visit were also excluded. Finally, data from 4598 whole blood donors (2605 men and 1993 women) were analyzed in this study. Most of these donors were also used for the development of the existing models, however, the intended visit used for the current study was at a later point in time (between 1 December 2008 and 31 July 2010 in the current study, and between 1 January 2007 and 31 December 2009 in the previous study).

Previously developed prediction models

The previously developed models predict the risk of Hb deferral at the intended visit, for men and women separately.[12] Hb deferral was defined as having an Hb level below the sex specific cutoff level for donation as described in the European Commission Directive,[13] which is 8.4 mmol/L (13.5 g/dL) for men and 7.8 mmol/L (12.5 g/dL) for women.

Predictive factors included in these models were age, seasonality, Hb level measured at the previous visit (previous Hb level), difference in Hb level between the previous two visits (delta Hb), time since the previous visit, total number of whole blood donations in

the past two years, and deferral at the previous visit. Seasonality was defined as the four meteorological seasons: winter (visits between December 1 and February 29), spring (March 1 to May 31), summer (June 1 to August 31), fall (September 1 to November 30). Deferral at the previous visit was categorized as deferral because of a low Hb level, deferral because of reasons other than a low Hb level, and no deferral.

These models were developed in data from all Dutch whole blood donors who fulfilled the inclusion criteria, which were comparable to the criteria in the current study (n=220,946). Visits that were used as intended visits occurred between 1 January 2007 and 31 December 2009.

See Appendix I for the exact formulas to calculate the risk of Hb deferral.

Data collection

Information about the predictors of the previously developed models was obtained from the administrative donor database. Hb levels were routinely measured during donor screening in finger stick capillary samples using a photometer (HemoCue, Angelholm, Sweden).

Venous blood samples for ZPP measurement were collected in tubes containing lithium heparin at each donor visit. On the next day after blood collection, ZPP levels in red blood cells were measured with a hematofluorometer (Model 206D, Aviv Biomedical, Lakewood, NJ). Before ZPP measurement, red blood cells were washed with phosphate-buffered saline to minimize the influence of other fluorescent elements in plasma.[14] ZPP levels that were measured at the visit prior to the intended visit were used for the analysis. We will further use the term "previous ZPP level". If ZPP was measured more than once per donor, we also calculated "delta ZPP" for these donors. Delta ZPP was defined as the difference between the ZPP level measured at the previous visit and at the second-last visit prior to the intended visit.

Statistical analysis

Missing values occurred in the following variables: previous ZPP level (22%), previous Hb level (1%), delta Hb (2%), and deferral at the previous visit (1%). In order to be able to use the observed information of other known variables, we multiple imputed (MI)[15] the missing values five times with the aregImpute function from the Design library,[16] applicable in R software.[17] The imputation models used all the predictors in our models and also the Hb level at the intended visit. The five completed datasets were identical in known information, but could differ on imputed values for missing information. Analyses were performed in each MI data set. Estimates from the five MI datasets were then pooled into one overall estimate and variance according to Rubin's Rules.[18]

Logistic regression analysis was used with Hb deferral (yes/no) as dichotomous outcome variable. The form of association between previous ZPP level and the risk of Hb deferral was studied on the logodds scale using restricted cubic spline functions with five knots.

The odds ratio (OR) was scaled so that it corresponded to a change from the 25th percentile to the 75th percentile of the ZPP distribution. This scaling allowed for a direct interpretation of the prognostic value of ZPP that had been recorded on a continuous scale. Interaction terms between previous ZPP level and delta ZPP, and between previous ZPP level and previous Hb level were examined with likelihood ratio tests, but none were statistically significant.

We used the previously developed sex specific prediction models[12] to calculate the risk of Hb deferral for each donor. Then, two logistic regression models were fitted:

Model 1: intercept + risk of Hb deferral previous model (basic model)

Model 2: intercept + risk of Hb deferral previous model + previous ZPP level (extended model)

Risk of Hb deferral was included as the log-odds (linear predictor) of the previously developed model. Rather than including the separate predictors of the previous prediction model individually, we explicitly chose to use the fixed regression coefficients of the original model, summarized into the average predicted risk (on the log-odds scale). This was because the regression coefficients of the previously developed model were estimated in a much larger cohort than the current cohort, making them highly robust estimates. For comparison purposes however, we also fitted a model in which the regression coefficients of all the predictors were newly estimated in the current cohort, and then this fitted model was again extended with previous ZPP level.

The statistical added value of previous ZPP level was assessed by comparing the basic model (model 1) with the extended model (model 2) in model fit and discriminative ability. The model fit was assessed with the model X^2. Discrimination refers to the ability of the models to distinguish donors who were deferred from donation based on their Hb level[19] from donors who were approved. The discriminative ability was assessed with the concordance(c)-statistic. A value of 1 indicates perfect discrimination; a value of 0.5 indicates poor discrimination, equivalent to flipping a coin.

The clinical usefulness of previous ZPP level was assessed with the net reclassification improvement (NRI) and with differences in net benefit (NB) estimated with decision curve analysis. NRI focuses on reclassification of donors, when the model including previous ZPP level was used for risk calculation compared to a model without previous ZPP level. NRI quantifies the correct movement into categories. Donors should move to a higher risk category if deferred, and to a lower risk category if approved.[20] To avoid the use of threshold levels that divide donors into high and low risk groups, we used the continuous NRI (NRI>0). In this measure, all possible threshold levels are used together.[21] Decision curve analysis is a method that incorporates consequences of different risk thresholds to divide donors into high and low risk groups. NB can be estimated by summing the benefits (true positives) and subtracting the harms (false positives). The latter are weighted by a factor related to the relative harm of an Hb deferral versus an unnecessary postponement of invitation for a donation. The weighting is derived from the threshold risk that is used to

select donors with a high risk of Hb deferral and for which the invitation of donation will be postponed. When a 10% threshold risk is used, inviting a donor with a low Hb level is 9 times (90/10) worse than postponing the invitation of a donor with an appropriate Hb level. Decision curve analysis is very useful to compare the impact of the models with and without previous ZPP level on donor management, and it can also help to decide whether either model is worth using at all.[22,23]

All analysis were performed separately for men and women, since the prediction models are also sex specific. We used statistical packages (SPSS, Version 18, SPSS, Inc., Chicago, IL; and R, Version 2.12.2, http://cran.r-project.org/bin/windows/base/old/2.12.2/).

Results

Table 1 shows the characteristics of the whole blood donors. A total of 76 men (2.9%) and 69 women (3.5%) were deferred from blood donation because of a low Hb level.

The association of previous ZPP level with risk of Hb deferral could be well described with a log transformation of ZPP. High ZPP levels were associated with an increased risk of Hb deferral (Figure 1). In men, the univariate OR for previous ZPP level at the interquartile (IQ) range was 4.0 (95% confidence interval (CI): 3.8 – 4.3). In women, the OR at the IQ range was 2.9 (95% CI: 2.7 – 3.2). The univariate effect of delta ZPP on the risk of Hb deferral was not statistically significant in men and women. This variable was not further used in subsequent analyses.

The multivariable OR for previous ZPP level in model 2 was 2.0 (95% CI: 1.7 – 2.3) for men and 2.2 (95% CI: 1.9 – 2.4) for women (Table 2). When previous ZPP level was added to the model, the model X^2 increased with 11% in men (from 238 to 264), and with 37% in women (from 75 to 103). Estimations of the c-statistics also showed a stronger added value for women than for men. Also in terms of the total continuous NRI, the added value of previous ZPP level was stronger for women than for men. In men, particularly approved donors were better classified when previous ZPP level was considered in risk prediction. In women, deferred and approved donors showed similar improvement in reclassification (Table 2). Decision curve analysis showed also added value of previous ZPP level (Figure 2). Previous ZPP level is particularly clinically useful in men for threshold levels above 12% risk of Hb deferral. At threshold levels between 5% and 15% risk, previous ZPP level has added value for female donors.

Previous ZPP level had also added predictive value in the model in which the regression coefficients of all the predictors were newly estimated in the current cohort, both in men and women. Refitted models without previous ZPP level had a c-statistic of 0.93 in men, and 0.87 in women. When previous ZPP level was added to the refitted models, the c-statistic in men increased to 0.94 and in women to 0.89.

Table 1 Distribution of predictors at the intended visit*

Predictor	Total n=4598	Men n=2605	Women n=1993
Age, years	50 (±12)	52 (±11)	47 (±12)
Seasonality			
Winter	1573 (34.2)	918 (35.2)	655 (32.9)
Spring	581 (12.6)	246 (9.4)	335 (16.8)
Summer	865 (18.8)	577 (22.1)	288 (14.5)
Fall	1579 (34.3)	864 (33.2)	715 (35.9)
Previous Hb level, mmol/L	8.9 (±0.7)	9.3 (±0.7)	8.5 (±0.6)
Delta Hb, mmol/L	0 (-0.4 – 0.3)	0 (-0.4 – 0.3)	0 (-0.4 – 0.4)
Time since previous visit, days	125 (85 – 180)	96 (70 – 167)	154 (119 – 189)
Deferral at previous visit			
Due to low Hb	171 (3.8)	86 (3.3)	85 (4.3)
Due to other reason than low Hb	99 (2.2)	46 (1.8)	53 (2.7)
Number of whole blood donations in past 2 years	4 (3 – 6)	6 (3 – 8)	3 (2 – 4)
Previous ZPP level, µmol/mol heme	66 (±23)	63 (±23)	69 (±24)
Delta ZPP, µmol/mol heme	1 (-7 – 9)	1 (-7 – 8)	0 (-10 – 11)
Hb intended visit, mmol/L	9.0 (±0.8)	9.3 (±0.7)	8.6 (±0.6)
Hb deferral†	145 (3.2)	76 (2.9)	69 (3.5)

* Data are reported as mean (±SD), n (%), or median (25^{th} – 75^{th} percentile).
† Hb < 8.4 mmol/L for men and < 7.8 mmol/L for women.

Figure 1 Relationship between previous ZPP level and risk of Hb deferral with 95% CI for men (**A**) and women (**B**)

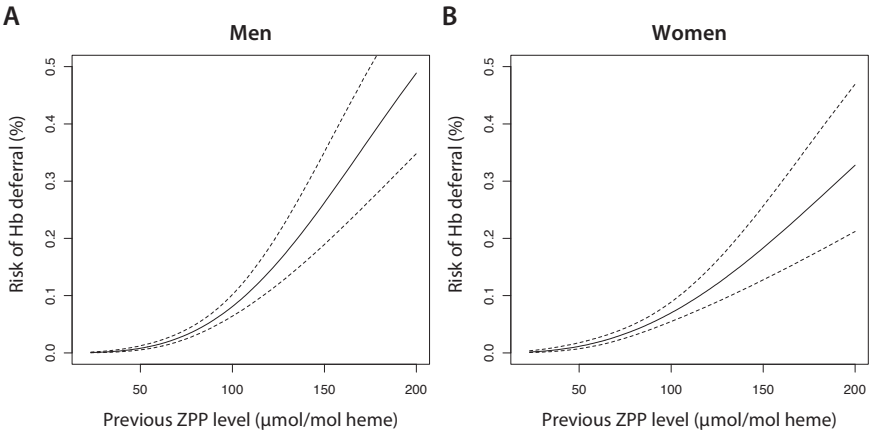
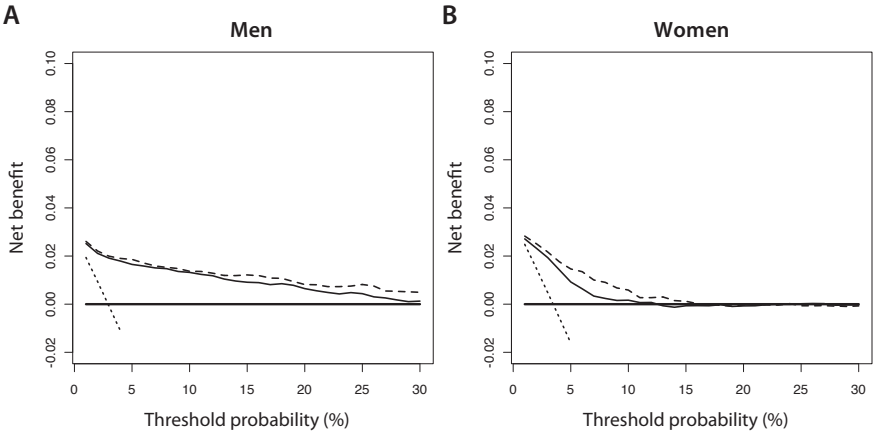


Figure 2 Decision curve for prediction of Hb deferral in men (**A**) and women (**B**)



Dotted line: assume all donors have a low Hb level, postpone invitation for all donors. Thick black line: assume no donors have a low Hb level, invite all donors. Black line: prediction model with 7 predictors. Dashed line: prediction model with previous ZPP level added. The graph gives the expected net benefit per donor relative to no postponement of invitation for any donor ("treat none"). The unit is the benefit associated with one deferred donor whose invitation is duly postponed.

Table 2 Performance of original prediction models for Hb deferral and models with previous ZPP level added

Model		OR previous ZPP level (95% CI) (75th vs. 25 th percentile)*	Model X^2	C-statistic (95% CI)	NRI (>0)†
Men	Original model		238	0.93 (0.90 – 0.96)	
	Previous ZPP level added	2.02 (1.74 – 2.30)	264	0.94 (0.92 – 0.96)	0.38+0.04=0.42
Women	Original model		75	0.80 (0.76 – 0.84)	
	Previous ZPP level added	2.15 (1.86 – 2.44)	103	0.85 (0.82 – 0.88)	0.28+0.28=0.56

* Interquartile range: men: 48 to 72 µmol/mol heme, women: 53 to 80 µmol/mol heme.

† NRI(>0) was calculated using all decision thresholds,
numbers represent the NRI (deferred donors) + NRI (approved donors) = total NRI.

Discussion

In this study, we assessed the added value of ZPP levels measured at the previous visit for the prediction of Hb deferral in whole blood donors. We found that previous ZPP level clearly improved the prediction of Hb deferral. Although previous ZPP level had added predictive value in both men and women, the model fit and discriminative ability improved particularly in women. The greater improvement in women could be explained by a lower discriminative ability of the original model in women: there is more space for improvement when a c-statistic is 0.80 (women) than when it is 0.93 (men).

Implications of using ZPP levels in practice were evaluated with decision curve analysis. We found that previous ZPP levels can help in making decisions about immediate versus postponed invitation for a donation. In men, previous ZPP levels are valuable if the threshold for postponing the invitation is above 12% risk of Hb deferral. In women, previous ZPP levels may be helpful if the threshold levels are between 5% and 15% risk. Hence, it depends on the chosen threshold level if previous ZPP level has added value in the prediction of Hb deferral. We found that around 10% of donors fell in these risk categories. The model without ZPP may be used to identify donors for which additional ZPP measurement may result in a better risk prediction of Hb deferral.

Donors with a low predicted risk of Hb deferral can be invited for a donation with preference, whereas those with a high risk of Hb deferral may benefit from interventions such as postponement of the invitation for donation, preferential invitation in fall and winter, iron fortification or a dietary advice. Ultimately, these models may be useful to decrease the number of donor deferrals for low Hb levels and to retain blood donors that would have quit donating after one or more deferrals.

Increased ZPP levels are associated with iron deficiency.[8] We indeed found a higher risk of Hb deferral in donors with high ZPP levels at the previous visit. In previous studies it has been concluded that ZPP testing may be useful in screening donors for iron depletion and potential risk of iron deficiency anemia.[24-28] It should be realised that ZPP levels are high in all conditions that result in an impaired iron supply for erythropoiesis, not only in iron deficiency. Such conditions include lead poisoning,[8] anemia of chronic disease,[29] haemolytic anemias, hemoglobinopathies,[30] and malignant diseases.[31] However, lead poisoning is rare and the other diseases mentioned are not relevant among blood donors as donors with these conditions are deferred from donation. Therefore, high ZPP levels in blood donors will most likely be the result of a low iron status.

In this study, we used different measures to assess the added value of previous ZPP level. The model X^2 and c-statistic are traditional statistical measures for evaluating prediction models. These measures give information about the model fit and the predictive accuracy of the model and are relevant to obtain insight into the incremental value of a marker. However, they don't give information about consequences of using the model and cannot be used for making decisions.[32] Decision curve analysis is a method that incorporates such

consequences of classifying patients as deferred, which may be right (true positives) or wrong (false positives). It weighs the relative value of the true positives to the false positives in terms of net benefit.

The following example illustrates how previous ZPP level can change the calculated risk of Hb deferral and accordingly decision making. Consider a man at the age of 50 who is invited to the blood collection center in spring. With other values of predictive factors that belong to this man, a risk of Hb deferral of 37% is calculated according to the original prediction model. Taking his previous ZPP level into account, this man would have a risk of Hb deferral of 35% when his ZPP level measured at the previous visit was 80 µmol/mol heme, and a risk of Hb deferral of 14% when his previous ZPP level was 40 µmol/mol heme. Depending on the chosen risk threshold level below which donors can be invited for a donation, addition of previous ZPP level to the prediction model can change the decision about invitation. The demonstrated added value of previous ZPP levels implies that the decision about whether or not inviting a donor for a donation is more often the correct decision when previous ZPP level is added to the model. This should be assessed in future research with an impact analysis.[33]

A limitation of this study is the small effective sample size. Despite the inclusion of 4,598 donors, the number of Hb deferrals is relatively low (76 men and 69 women). We therefore fixed the regression coefficients of the previously developed models and included the risk of Hb deferral as one variable in a new model. As a result only two regression coefficients needed to be estimated. A disadvantage of this approach may be that the estimated effect of previous ZPP level is confounded by predictors of the models. Therefore, predictors that were correlated with ZPP were also added to the model that included previous ZPP level. We found that the estimated effect of previous ZPP level was stable across models.

A second limitation of this study is the different Hb deferral rate in the current study compared to the study in which the original models were developed. Deferral rates in the development study were 4.1% for men and 7.7% for women, compared to 2.9% (men) and 3.5% (women) in the current study. An explanation for these differences in deferral rates is difficult to give, but it may be related to the included blood collection centers. In the current study, donors were recruited from two blood collection centers in the southeast region of the Netherlands, whereas in the previous study donors were recruited from all blood collection centers in the Netherlands.

Another limitation of the study is that we measured ZPP levels in venous blood samples rather than in finger stick capillary samples, which will be the practical implication. We measured ZPP levels in venous blood samples to obtain a first insight in the possible added value of ZPP. Next, the added value should also be examined by using ZPP levels measured in finger stick capillary samples.

In this study we have shown that ZPP measurements obtained at the previous visit have added value in prediction of Hb deferral. It is, however, too early to recommend adding

previous ZPP levels to the prediction model for implementation in practice. For this reason we do not present the prediction models with previous ZPP level added. Instead, we recommend to investigate the added value of ZPP levels measured in finger stick capillary samples. It may be that additional measurement of ZPP is only valuable in donors with a high risk of Hb deferral. The model without ZPP may then be used to identify donors for which additional ZPP measurement may result in a better risk prediction of Hb deferral. Furthermore, in future research, the added value of ZPP should be compared with the added value of other iron parameters in the prediction model. Additionally, it would be informative to consider also the costs of the different measurements to decide which iron parameter to use.[34]

In conclusion, ZPP levels measured at the previous visit may have added value in the prediction of Hb deferral in whole blood donors. Regular measurement of ZPP levels may therefore be helpful in the management of the donation program and may eventually lead to a decrease in donor deferrals for low Hb levels. However, more research is needed on the value of finger stick ZPP in Hb deferral prediction, as well as on the added value of other iron parameters, before extended models can be implemented in practice.

Acknowledgements

The authors thank Giny Hetterscheit and Mirianne van Gerwen for their help with the practical work and Ingrid Veldhuizen for her contribution in fund raising. Karel G.M. Moons gratefully acknowledges financial contribution for all his research by the Netherlands Organisation for Scientific Research (project 9120.8004 and 918.10.615).

Reference List

1. European Directorate for the Quality of Medicines and HealthCare. Guide to the preparation, use and quality assurance of blood components, European Committee (partial agreement) on Blood Transfusion (CD-P-TS), Recommendation No. R(95) 15, 15th ed. Strasbourg: Council of Europe; 2009, p1899.
2. Halperin D, Baetens J, Newman B. The effect of short-term, temporary deferral on future blood donation. Transfusion. 1998;38:181-183.
3. Custer B, Chinn A, Hirschler NV, Busch MP, Murphy EL. The consequences of temporary deferral on future whole blood donation. Transfusion. 2007;47:1514-1523.
4. Zou S, Musavi F, Notari EP, Rios JA, Trouern-Trend J, Fang CT. Donor deferral and resulting donor loss at the American Red Cross Blood Services, 2001 through 2006. Transfusion. 2008;48:2531-2539.
5. Finch CA, Cook JD, Labbe RF, Culala M. Effect of blood donation on iron stores as evaluated by serum ferritin. Blood. 1977;50:441-447.
6. Skikne B, Lynch S, Borek D, Cook J. Iron and blood donation. Clin Haematol. 1984;13:271-287.
7. Simon TL, Garry PJ, Hooper EM. Iron stores in blood donors. JAMA. 1981;245:2038-2043.
8. Lamola AA, Yamane T. Zinc protoporphyrin in the erythrocytes of patients with lead intoxication and iron deficiency anemia. Science. 1974;186:936-938.
9. Labbe RF, Rettmer RL. Zinc protoporphyrin: a product of iron-deficient erythropoiesis. Semin Hematol. 1989;26:40-46.
10. Cook JD. Clinical evaluation of iron deficiency. Semin Hematol. 1982;19:6-18.
11. Kohgo Y, Niitsu Y, Kondo H et al. Serum transferrin receptor as a new index of erythropoiesis. Blood. 1987;70:1955-1958.
12. Baart AM, de Kort WL, Atsma F, Moons KG, Vergouwe Y. Development and validation of a prediction model for low hemoglobin deferral in a large cohort of whole blood donors. Transfusion. 2012;52:2559-2569.
13. The Commission of the European Communities. European Commissian Directive 2004/33/EC. 2004.
14. Hastka J, Lasserre JJ, Schwarzbeck A, Strauch M, Hehlmann R. Washing erythrocytes to remove interferents in measurements of zinc protoporphyrin by front-face hematofluorometry. Clin Chem. 1992;38:2184-2189.
15. Donders AR, van der Heijden GJ, Stijnen T, Moons KG. Review: a gentle introduction to imputation of missing values. J Clin Epidemiol. 2006;59:1087-1091.
16. Harrell, F. E., Jr. Design: Design package. R package version 2.0. 2011. [cited 2011 Feb 10]. Available from: URL: http://biostat.mc.vanderbilt.edu/s/Design.
17. R Development Core Team. R: A language and environment for statistical computing. 2011. [cited 2011 Feb 10]. Available from URL: http://www.R-project.org.
18. Rubin DB, Schenker N. Multiple imputation in health-care databases: an overview and some applications. Stat Med. 1991;10:585-598.
19. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996;15:361-387.
20. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med. 2008;27:157-172.
21. Pencina MJ, D'Agostino RB, Sr., Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. Stat Med. 2011;30:11-21.
22. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006;26:565-574.
23. Vickers AJ, Cronin AM. Traditional statistical methods for evaluating prediction models are uninformative as to clinical value: towards a decision analytic framework. Semin Oncol. 2010;37:31-38.
24. Schifman RB, Rivers SL, Finley PR, Thies C. RBC zinc protoporphyrin to screen blood donors for iron deficiency anemia. JAMA. 1982;248:2012-2015.
25. Raftos J, Schuller M, Lovric VA. Iron stores assessed in blood donors by hematofluorometry. Transfusion. 1983;23:226-228.
26. Morse EE, Cable R, Pisciotto P, Kakaiya R, Kiraly T. Evaluation of iron status in women identified by copper sulfate screening as ineligible to donate blood. Transfusion. 1987;27:238-241.

27. Jensen BM, Sando SH, Grandjean P, Wiggers P, Dalhoj J. Screening with zinc protoporphyrin for iron deficiency in non-anemic female blood donors. Clin Chem. 1990;36:846-848.
28. Harthoorn-Lasthuizen EJ, Lindemans J, Langenhuijsen MM. Zinc protoporphyrin as screening test in female blood donors. Clin Chem. 1998;44:800-804.
29. Hastka J, Lasserre JJ, Schwarzbeck A, Strauch M, Hehlmann R. Zinc protoporphyrin in anemia of chronic disorders. Blood. 1993;81:1200-1204.
30. Graham EA, Felgenhauer J, Detter JC, Labbe RF. Elevated zinc protoporphyrin associated with thalassemia trait and hemoglobin E. J Pediatr. 1996;129:105-110.
31. Gorodetsky R, Fuks Z, Peretz T, Ginsburg H. Elevation of erythrocyte zinc- and free protoporphyrins with metastatic spread in cancer patients. Eur J Cancer Clin Oncol. 1986;22:1515-1521.
32. Steyerberg EW, Pencina MJ, Lingsma HF, Kattan MW, Vickers AJ, Van CB. Assessing the incremental value of diagnostic and prognostic markers: a review and illustration. Eur J Clin Invest. 2012;42:216-228.
33. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann.Intern.Med. 2006;144:201-9.
34. Henriksson M, Palmer S, Chen R, Damant J, Fitzpatrick NK, Abrams K, Hingorani AD, Stenestrand U, Janzon M, Feder G, et al. Assessing the cost effectiveness of using prognostic biomarkers with decision models: case study in prioritising patients waiting for coronary artery surgery. BMJ 2010;340:b5606.

Appendix

Appendix I Logistic regression coefficients to calculate the risk of Hb deferral*

Predictor	Value or coding	Beta†	
		Men	Women
Age, per year‡	age	0.01	-0.03
	$(age-22.5)_+$^3		$1.22x10^{-4}$
	$(age-34.6)_+$^3		$-5.22x10^{-4}$
	$(age-45.0)_+$^3		$9.05x10^{-4}$
	$(age-54.0)_+$^3		$-6.79x10^{-4}$
	$(age-65.0)_+$^3		$1.73x10^{-4}$
Seasonality	winter‖	0	0
	spring	0.49	0.25
	summer	0.41	0.28
	fall	-0.20	-0.14
Previous Hb level, per mmol/L below sex specific cutoff value§	men: Hb-8.4, women: Hb-7.8	-2.27	-2.10
Previous Hb level, per mmol/L at or above sex specific cutoff value§	men: Hb-8.4, women: Hb-7.8	-3.77	-3.31
Delta Hb, per mmol/L equal to or below 0 mmol/L§	delta Hb	1.35	1.04
Delta Hb, per mmol/L above 0 mmol/L§	delta Hb	1.24	1.37
Time since previous visit, per month below 1 year§	months-12	-0.18	-0.16
Deferral at previous visit	no deferral‖	0	0
	due to low Hb	-0.62	-0.53
	due to other reasons than low Hb	-0.96	-0.93
Number of whole blood donations in past 2 years	nr of donations	-0.05	-0.17
Intercept		-2.77	-0.70

* With those regression coefficients the linear predictor can be calculated. The risk of Hb deferral equals: risk (Hb deferral) = $1/(1 + e^{-\text{linear predictor}})$.

† Beta = regression coefficient, shows the strength and the direction of the variable's influence.

‡ In men the association between age and risk of Hb deferral was linear, in women the association showed several curves that were modeled with a restricted cubic spline function. Negative values of the cubic terms become 0; e.g. $(age-22.5)_+$ indicates age-22.5 for positive values, 0 for negative values.

§ Nonlinear association with risk of Hb deferral.

‖ Winter and no deferral are reference categories, 1 if true, 0 if false.

Chapter 7

Prediction of hemoglobin levels in whole blood donors: how to model donation history

A. Mireille Baart*†, Yvonne Vergouwe†, Femke Atsma*, Karel G.M. Moons†, Wim L.A.M. de Kort*

* Sanquin Research, dept. Donor Studies, Nijmegen, the Netherlands

† Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands

Abstract

Background: Recently, prediction models for hemoglobin (Hb) deferral risk have been developed. These models consider the previous Hb level plus a change in Hb. Here, we investigated if the performance of models could be improved by considering more information on the history of Hb levels.

Study design and methods: Data of 187,711 Dutch whole blood donors with sequential Hb measurements during two years were used to develop and internally validate three different regression models: two simple linear models with history of Hb levels included as i) Hb at the previous visit plus change in Hb, or ii) mean of all previous Hb levels; and one mixed effect model including measurements of all previous Hb levels.

Results: Thirteen percent of men and 19% of women were deferred because of a low Hb level at least once in two years. The simple linear models and the mixed effect model performed similar, if an estimate of the random intercept of the mixed effect model was used for individual donors to calculate the predicted Hb level. In men, the concordance(c)-statistic ranged from 0.87 to 0.89 and the R^2 from 0.41 to 0.44. In women, the c-statistic ranged from 0.81 to 0.84. Values of R^2 were higher for the linear models than for the mixed effect model, 0.36 and 0.38 vs. 0.29 respectively.

Conclusion: Previous Hb levels could be summarized with one predictor as the mean value of all previous Hb levels. This predictor can be used in an easy to use simple linear regression model.

Introduction

Prior to a blood donation, the iron status of blood donors is assessed, most commonly by measuring hemoglobin (Hb) levels. Donors with low Hb levels are deferred from donation to protect them from developing iron deficiency anemia after a donation and to guarantee that blood units for transfusion meet the required standards for Hb content.[1]

In the past, prediction models for Hb deferral risk have been developed by our own research group.[2] These models predict the risk of Hb deferral for individual donors at the moment of possible invitation for a next donation. The strongest predictor in these models was the Hb level measured at the previous visit. The prediction models performed quite well, but might be improved by adding additional factors.

In the Netherlands, male whole blood donors may donate up to five times per year and female whole blood donors up to three times a year.[3] Consequently, there is a lot of information available in the donor data base. The previously developed prediction models only include the Hb level measured at the previous visit and a change in Hb level between the second last visit and the previous visit as information on previous Hb levels. We hypothesized that including more information about the history of Hb levels might further improve the predictive performance. This information may for example be included as the mean of all previous Hb levels, or measurements of all previous Hb levels could be included using an analysis with repeated measurements.

Additionally, using Hb level as a continuous outcome instead of a dichotomous outcome (Hb deferral yes/no) might improve the predictions even further. Moreover, by using continuous values of Hb level as outcome we avoid the problem of different Hb cutoff values for donation that other countries may use. In an earlier study, we found limited performance when the prediction models for Hb deferral were externally validated in a population with different Hb cutoff values for donation.[4] By using continuous values of Hb level we may overcome this problem.

In the current study we compared the performance of different modeling techniques for the prediction of Hb level as a continuous outcome: simple linear regression analysis versus mixed effect regression analysis using repeated measurements. Furthermore, we compared different predictors that contained information on the donation history in the simple linear regression analysis.

Donors and methods

Donors

Whole blood donors, who visited a blood collection center in the Netherlands in the years 2007 until 2009 were eligible for the study. We defined whole blood donors as blood donors whose donations after 1 January 2005 were all whole blood donations, rather than

plasma donations. All these donors fulfilled the Dutch criteria for donation with relation to the time interval between donations and donation frequency. The minimum time interval between two donations is 56 days, and the maximum number of donations allowed per year is five for men and three for women.
From these donors, data of all visits from 2005 until 2009 were extracted from the donor database. To be included in the study, donors should have visited the blood collection center twice in the two years before their last visit, irrespective of the donor was approved or deferred for donation at this visit. In addition, donors should have donated twice in the four years before their last visit. These criteria were required to collect sufficient information on donation history. Donors with uncertainties in their history, for example whether or not a donation took place, were excluded. Donors with missing values were also excluded. Finally, donors registered as a new donor in the two years preceding the last visit were also excluded.
Finally, a total of 187,711 whole blood donors (96,514 men and 91,197 women) were included in the study. From these donors, data of visits in the two years before their last visit were used for the analyses.

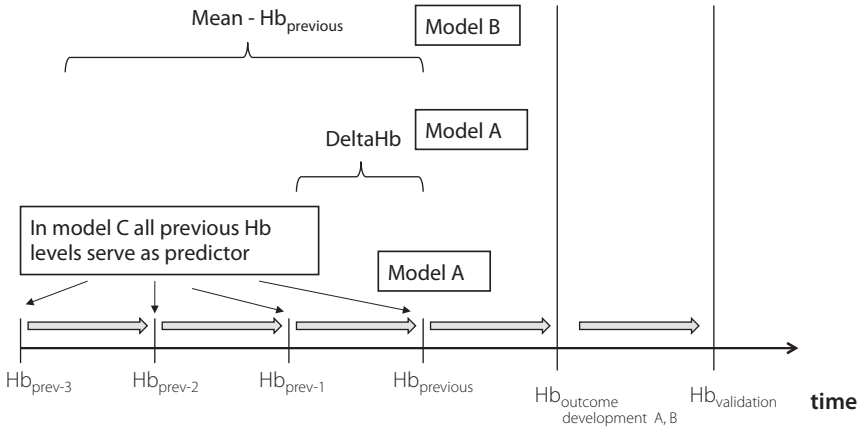
Outcome variable and predictors

We predicted the Hb level measured during donor screening at the next visit to the blood collection center. Hb levels were routinely measured in finger stick capillary samples using a photometer (HemoCue, Angelholm, Sweden).
The same predictors as in the previously developed models for Hb deferral risk[2] were considered: age, seasonality, time since the previous visit, deferral at the previous visit, and total number of whole blood donations in the past two years. Seasonality was defined as the four meteorologic seasons: winter (visits between December 1 and February 29), spring (March 1 to May 31), summer (June 1 to August 31), fall (September 1 - November 30). Deferral at the previous visit was categorized as deferral because of a low Hb level, deferral because of reasons other than a low Hb level, and no deferral. Information on the history of Hb levels was included in different ways in three different models. Model A included the Hb level measured at the previous visit (previous Hb level) plus a change in Hb level between the second last visit and the previous visit (delta Hb). Model B included the mean of all previous Hb levels assessed for a donor in the two years period preceding the last visit. Model C included all previous Hb levels assessed for a donor in the two years period, treated as repeated measurements. An overview of the different predictors on the history of Hb levels is presented in Figure 1.

Statistical analysis

Model A and B were developed with simple linear regression analysis and model C with mixed effect regression analysis. Data from the second last visit and before were used to develop the models (development data). Data from the last visit were used to validate the

Figure 1 Overview of the different predictors on the history of Hb levels used in different model (**A**, **B** and **C**)



In model A, the previous Hb level and delta Hb are used as predictor.
In model B, the mean of all previous Hb levels is used as predictor.
In model C, all previous Hb levels are used as predictor.

models (validation data). The models were developed and validated separately for men and women.

For the development of model A and B, the Hb level measured at the second last visit was used as outcome variable. Values for the predictors were also assessed at the second last visit. For the development of model C, Hb levels measured at the second last visit and before were used as outcome variables. Values of predictors at all those visits were included in the analysis.

The multiple measurements per donor of Model C will be correlated by sharing (partly) the same information and are therefore not independent. Mixed effect models can be used to account for the correlations of observations within the same donor. Mixed effect models typically include both fixed and random effects and have become a primary method for longitudinal data analysis. The coefficients for the fixed effects provide estimates for the predictor effects in the donor cohort. The donor specific random effects account for variance heterogeneity among responses from different individual donors.[5]

In the mixed effect models, all predictors were considered as fixed effects. Per donor, a random intercept was estimated. The presence of a random time effect (age) could not be found. The difference in -2 loglikelihood between a model with and without a random time effect (age) was not statistically significant.

All three models were validated on the last visit of all donors in the dataset. Furthermore, a cross-validation procedure was performed. Hereto, prediction models were developed separately in each of the four Dutch geographical regions (northeast, northwest, southeast or southwest). All models were subsequently validated in the four regions separately. For the validation, the predictive performance of the models was assessed in terms of calibration and discrimination.[6] Calibration is the agreement between predicted and observed Hb levels. Calibration was studied with a simple linear model with Hb level as continuous outcome and the linear predictor as the only covariate.[7] The regression coefficient of the linear predictor (the calibration slope, visualized in a calibration plot) reflects whether the effects of the predictors are estimated correctly, and is ideally 1. We also assessed calibration-in-the-large by fitting a simple linear model with the linear predictor as an offset variable (setting the regression coefficient to 1). The intercept indicates whether predicted Hb values are on average correct, and is ideally 0. Discrimination is the ability of the model to differentiate between deferred and approved donors. Discrimination was determined with the concordance(c)-statistic.[8] A value of 1 indicates perfect discrimination; a value of 0.5 indicates poor discrimination, equivalent to flipping a coin. To calculate the c-statistic, observed Hb levels were dichotomized at the sex-specific cutoff level for donation. Furthermore, the overall performance of the prediction models was assessed with the explained variance (R^2), estimated as (Pearson rho)^2.

In order to validate model C, the model was applied in three different manners. First, the model predictions were calculated based on the fixed effects only (C1). Second, the model predictions were calculated based on the fixed effects plus the random intercept that was estimated during model development for each donor (C2). Third, the model predictions were calculated based on the fixed effects plus an estimate of the random intercept per donor (C3). Hereto, a univariable simple linear model was fitted with the random intercept as dependent variable and the mean of the previous Hb levels as predictor. The random intercept could then be estimated as $\alpha + \beta$ [mean of the previous Hb levels].

Statistical analyses were performed with SPSS, Version 20, SPSS, Inc., Chicago, IL; and R, Version 2.15.2, http://cran.r-project.org/bin/windows/base/old/2.15.2/.

Results

Table 1 presents the characteristics of the whole blood donors at the second last visit (the last visit in the development data). At this visit, mean Hb levels in men were 9.3 mmol/L and in women 8.5 mmol/L. Hb deferral percentages at the second last visit were 3% for men and 6% for women. In the two years period, 13% of men and 19% of women were deferred at least once because of a low Hb level (Table 2).

Univariate effects of the predictors, estimated as one measurement (simple linear regression) or as repeated measurements (mixed effect regression), are presented in Table 3.

Table 1 Donor characteristics at second last visit*

Characteristic	Men n=96,514	Women n=91,197
Age, years	50 (±12)	45 (±13)
Seasonality		
Winter	25,833 (27)	24,153 (26)
Spring	23,620 (24)	23,063 (25)
Summer	23,729 (25)	21,146 (23)
Fall	23,332 (24)	22,835 (25)
Previous Hb level, mmol/L	9.4 (±0.7)	8.5 (±0.6)
Delta Hb, mmol/L	0 (-0.4 – 0.4)	0 (-0.4 – 0.4)
Mean of previous Hb levels, mmol/L	9.4 (±0.5)	8.5 (±0.5)
Time since previous visit, days	98 (77 – 161)	140 (119 – 189)
Deferral at previous visit		
Due to low Hb	2,832 (3)	5,217 (6)
Due to other reason than low Hb	2,858 (3)	3,532 (4)
Number of whole blood donations in past 2 years	5 (3 – 7)	3 (2 – 4)
Hb, mmol/L	9.3 (±0.7)	8.5 (±0.6)
Hb deferral†	3,326 (3)	5,863 (6)

* Data are reported as mean (±SD), n (%) or median (25^{th} – 75^{th} percentile).
† Hb < 8.4 mmol/L for men and < 7.8 mmol/L for women.

The strongest predictive effects were observed for predictors on the history of Hb levels: for previous Hb level as one measurement or as repeated measurements and for the mean of the previous Hb levels.

The predictive strength of previous Hb level was much larger in model A compared to model C (Table 4). The effect of other predictors in the different models was comparable.

In men, the performance of most prediction models was good when the model was applied to predict the Hb level at the last visit in the validation procedure (Table 5). The simple linear models A and B performed similar. The calibration slope and calibration-in-the-large were close to the ideal values of 1 and 0 respectively. Calibration is also shown in Figure 2. The discriminative ability for model A and B was similar. Overall, the performance was slightly better for model B (the R^2 of model B was 0.44 vs. 0.41 for model A).

The performance of model C was disappointing when model predictions were calculated based on the fixed effects only (C1). When a random intercept was used, either as estimated during model development (C2) or based on the mean of the previous Hb levels (C3), the

Table 2 Number of visits and Hb deferrals in the two years period

Number of visits	**2**	**3**	**4**	**5**
Male donors, n (%)	7,428 (8)	11,618 (12)	13,582 (14)	12,092 (13)
Number of visits	**2**	**3**	**4**	**5**
Female donors, n (%)	9,987 (11)	16,760 (18)	19,898 (22)	19,346 (21)
Number of Hb deferrals	**0**	**1**	**2**	**3**
Male donors, n (%)	83,800 (87)	8,585 (9)	2,819 (3)	984 (1)
Female donors, n (%)	73,651 (81)	11,671 (13)	4,089 (4)	1,383 (2)

Table 3A Univariate effects for the prediction of Hb levels in men

Predictor	**Value or coding**	**Simple linear model**	**Mixed effect model**
		Beta* (se)	**Beta* (se)**
Age, years	59 vs. 42†	-0.08 (0.003)	-0.09 (0.002)
Seasonality	winter	0	0
	spring	-0.10 (0.006)	-0.08 (0.002)
	summer	-0.12 (0.006)	-0.11 (0.002)
	fall	-0.01 (0.006)	-0.03 (0.002)
Previous Hb level, mmol/L	9.8 vs. 8.9†	0.52 (0.002)	0.19 (0.001)
Delta Hb, mmol/L	-0.4 vs. 0.4†	0.03 (0.003)	n.a.
Mean of previous Hb levels, mmol/L	9.7 vs. 9.0†	0.55 (0.002)	n.a.
Time since previous visit, days	161vs. 77†	0.08 (0.002)	0.06 (0.001)
Deferral at previous visit	no deferral	0	0
	due to low Hb	-0.55 (0.013)	0.15 (0.004)
	due to other reason than low Hb	0.15 (0.013)	0.10 (0.004)
Number of whole blood donations in past 2 years	7 vs. 3†	-0.17 (0.004)	-0.20 (0.002)
Random intercept	SD	n.a.	0.49

* Beta = regression coefficient, shows the strength and the direction of the variable's influence.
† Interquartile range.
n.a. not applicable.

6	**7**	**8**	**9**	**10**	**≥ 11**
11,350 (12)	10,904 (11)	10,434 (11)	10,699 (11)	7,964 (8)	443 (0.5)
6	**≥ 7**				
23,134 (25)	2,072 (2)				
≥ 4					
326 (0.3)					
403 (0.4)					

Table 3B Univariate effects for the prediction of Hb levels in women

Predictor	Value or coding	Simple linear model	Mixed effect model
		Beta* (se)	Beta* (se)
Age, years	55 vs. 34†	0.16 (0.003)	-0.16 (0.001)
Seasonality	winter	0	0
	Spring	-0.05 (0.006)	-0.04 (0.002)
	Summer	-0.07 (0.006)	-0.06 (0.002)
	Fall	-0.01 (0.006)	-0.01 (0.002)
Previous Hb level, mmol/L	8.9 vs. 8.1†	0.41 (0.002)	0.25 (0.001)
Delta Hb, mmol/L	-0.4 vs. 0.4†	0.02 (0.003)	n.a.
Mean of previous Hb levels, mmol/L	8.8 vs. 8.2†	0.43 (0.002)	n.a.
Time since previous visit, days	189 vs. 119†	0.04 (0.002)	0.03 (0.001)
Deferral at previous visit	no deferral	0	0
	due to low Hb	-0.42 (0.009)	0.12 (0.004)
	due to other reason than low Hb	0.11 (0.010)	0.12 (0.004)
Number of whole blood donations in past 2 years	4 vs. 2†	0.01 (0.003)	-0.06 (0.002)
Random intercept	SD	n.a.	0.43

* Beta = regression coefficient, shows the strength and the direction of the variable's influence.
† Interquartile range.
n.a. not applicable.

Table 4A Multivariable models for the prediction of Hb levels in men

Predictor	Value or coding	Model A Beta* (se)	Model B Beta* (se)	Model C Beta* (se)
Fixed effects				
Age, years	59 vs. 42†	-0.02 (0.003)	-0.01 (0.002)	-0.04 (0.002)
Seasonality	winter	0	0	0
	Spring	-0.12 (0.005)	-0.11 (0.005)	-0.09 (0.002)
	Summer	-0.13 (0.005)	-0.14 (0.005)	-0.10 (0.002)
	Fall	0.00 (0.005)	-0.02 (0.005)	-0.02 (0.002)
Previous Hb level, mmol/L	9.8 vs. 8.9†	0.67 (0.003)	n.a.	0.24 (0.001)
Delta Hb, mmol/L	-0.4 vs. 0.4†	-0.25 (0.003)	n.a.	n.a.
Mean of previous Hb levels, mmol/L	9.7 vs. 9.0†	n.a.	0.57 (0.002)	n.a.
Time since previous visit, days	161 vs. 77†	0.05 (0.002)	0.04 (0.001)	0.04 (0.001)
Deferral at previous visit	no deferral		0	0
	due to low Hb	0.29 (0.011)	0.11 (0.010)	0.31 (0.005)
	due to other reason than low Hb	0.11 (0.010)	0.10 (0.010)	0.08 (0.005)
Number of whole blood donations in past 2 years	7 vs. 3†	-0.01 (0.001)	-0.06 (0.003)	-0.11 (0.002)
Intercept		2.44 (0.032)	1.72 (0.033)	7.09 (0.018)
Random effect				
Random intercept	SD	n.a.	n.a.	0.35

* Beta = regression coefficient, shows the strength and the direction of the variable's influence.

† Interquartile range.

n.a. not applicable.

Table 4B Multivariable models for the prediction of Hb levels in women

Predictor	Value or coding	Model A Beta* (se)	Model B Beta* (se)	Model C Beta* (se)
Fixed effects				
Age, years	55 vs. 34†	0.06 (0.003)	0.06 (0.003)	0.09 (0.002)
Seasonality	Winter		0	0
	Spring	-0.06 (0.005)	-0.06 (0.005)	-0.05 (0.003)
	Summer	-0.08 (0.005).	-0.08 (0.005)	-0.07 (0.003)
	Fall	-0.00 (0.005)	-0.00 (0.005)	0.00 (0.003)
Previous Hb level, mmol/L	8.9 vs. 8.1†	0.56 (0.003)	n.a.	0.44 (0.002)
Delta Hb, mmol/L	-0.4 vs. 0.4†	-0.24 (0.003)	n.a.	n.a.
Mean of previous Hb levels, mmol/L	8.8 vs. 8.2†	n.a.	0.46 (0.002)	n.a.
Time since previous visit, days	189 vs. 119†	0.04 (0.002)	0.04 (0.001)	0.05 (0.001)
Deferral at previous visit	no deferral	0	0	0
	due to low Hb	0.25 (0.009)	0.13 (0.008)	0.32 (0.005)
	due to other reason than low Hb	0.13 (0.009)	0.11 (0.009)	0.19 (0.005)
Number of whole blood donations in past 2 years	4 vs. 2†	0.03 (0.003)	-0.00 (0.003)	0.04 (0.002)
Intercept		2.22 (0.032)	1.78 (0.032)	3.41 (0.018)
Random effect				
Random intercept	SD	n.a.	n.a.	0

* Beta = regression coefficient, shows the strength and the direction of the variable's influence.
† Interquartile range.
n.a. not applicable.

Table 5A Performance of the different prediction models in men at internal validation on the last visit

Model	Calibration-in-the-large	Calibration slope	C-statistic	R^2
A. Simple linear model *Previous Hb level + delta Hb*	-0.02	0.97	0.88	0.41
B. Simple linear model *Mean of previous Hb levels*	-0.03	1.00	0.87	0.44
C1. Mixed effect model *Application without random intercept*	-0.05	1.58	0.81	0.28
C2. Mixed effect model *Application with random intercept*	-0.04	1.01	0.89	0.42
C3. Mixed effect model *Application with estimated random intercept**	-0.04	1.08	0.88	0.44

Model A and B were developed using data from only the second last visit; model C was developed using data from all visits except the last visit (i.e. using repeated measurements).
* The random intercept was estimated as -4.45 + 0.47 x mean of previous Hb levels.

Table 5B Performance of the different prediction models in men at internal validation on the last visit

Model	Calibration-in-the-large	Calibration slope	C-statistic	R^2
A. Simple linear model *Previous Hb level + delta Hb*	-0.02	0.99	0.84	0.36
B. Simple linear model *Mean of previous Hb levels*	-0.03	1.03	0.84	0.38
C1. Mixed effect model *Application without random intercept*	-0.04	0.99	0.81	0.29
C2. Mixed effect model *Application with random intercept*	-0.04	0.99	0.81	0.29
C3. Mixed effect model *Application with estimated random intercept*	-0.04	0.99	0.81	0.29

Model A and B were developed using data from only the second last visit; model C was developed using data from all visits except the last visit (i.e. using repeated measurements).

performance increased substantially. In these latter situations, the performance of model C was similar to the performance of model A and B.

In women, the discriminative ability of model A and B were similar again, with model B slightly better overall performance (the R^2 of model B was 0.38 vs. 0.36 for model A). The random intercept was estimated as 0. Consequently, predictions for C1, C2 and C3 were the same with the same model performance. Compared to model C, the performance of model A and B was slightly better.

The cross-validation in the different geographical regions showed similar results.

Figure 2 Calibration plots for the different models in men at internal validation at the last visit

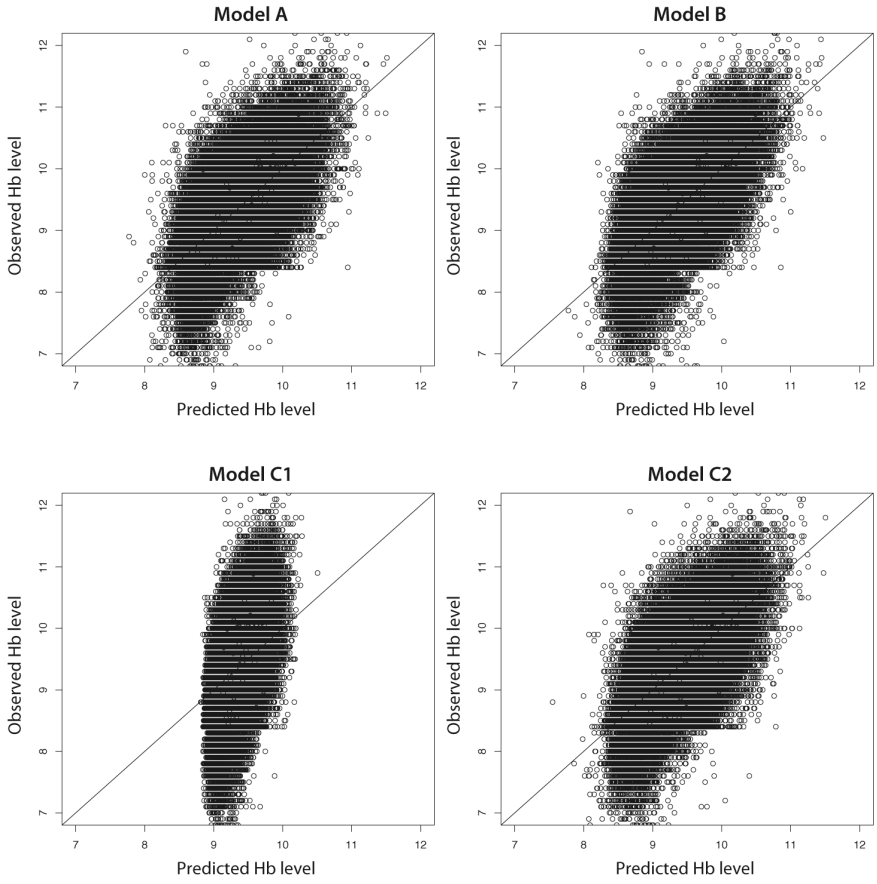
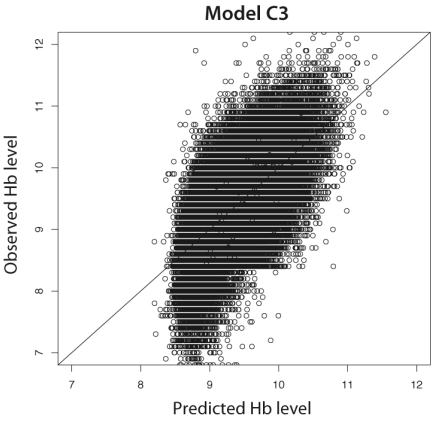


Figure 2 Continued



Calibration plots for the different models in men at internal validation at the last visit: A simple linear model with predictors previous Hb level plus delta Hb, B simple linear model with predictor mean of the previous Hb levels, C1 mixed effect model with model predictions calculated based on the fixed effects only, C2 mixed effect model with model predictions calculated based on fixed effects plus random intercept as estimated during model development, C3 mixed effect model with model predictions calculated based on fixed effects plus random intercept based on the mean of the previous Hb levels.

Discussion

In the present study, we compared the predictive performance of different modeling techniques for the prediction of Hb level as a continuous outcome. We examined different types of regression models and different predictors for the history of Hb levels. Overall, the three models that were studied performed well and their predictive performance was similar. The mixed effect model in which repeated Hb measurements in time were considered did not show better performance than the simple linear models. Moreover, in women the performance of the simple linear models was even better than the performance of the mixed effect model.

A striking difference in predictor effects in the simple linear models versus the mixed effect model was observed for the previous Hb level, both univariately and multivariably. The strength of the previous Hb level was smaller in the mixed effect model. This finding might be explained by the inclusion of a random intercept in the mixed effect model, which contains information on the mean of all measured Hb levels for the individual donors.

The performance of a simple linear model that contained the predictors previous Hb level and delta Hb was similar to the performance of a simple linear model that contained the predictor mean of the previous Hb levels. The overall performance was slightly better when the mean of the previous Hb levels was used (difference in R^2 was 0.03 in men and 0.02 in women). Possible explanations are that the value of the mean of the previous Hb levels is based on more measurements than the previous Hb level and delta Hb. Furthermore, the reliability of Hb measurement in capillary blood, which is the extent to which a measurement in a donor is reproducible over time, might be low;[9] using the mean of measurements decreases the error in measured Hb values due to unreliability.

We showed that the mixed effect model can only be useful in men, if an estimate of the random intercept is available for prediction. Otherwise, important donor information is neglected and the model performs poorly. In our situation, the random intercept could well be estimated with the mean of the previous Hb levels. The additional simple linear model used to estimate the random intercept in situation C3 is hence an important tool to use the mixed model in practice. If this model is being used in other donors than used for the development of the model, the simple linear model gives the random effect estimate per donor given the mean of the previous Hb levels of that donor. In women, the variance in random intercept was 0 for the multivariable model. We can only speculate why the variance is that low.

In men, the performance of the mixed effect model was similar to the performance of the simple linear models when a random intercept was used in the mixed effect model. In women, the simple linear models performed slightly better than the mixed effect model. A possible explanation for the similar performance of the different models is the relatively large error in measured Hb values, leading to random variation.

Results obtained in this study do not provide evidence that one of the three developed models is the best. The simple linear model using the mean of the previous Hb levels may be preferred for three reasons. Thirst, the overall performance assessed with the R^2 was highest for this model. Second, using the mean of previous Hb levels decreases the error in measured Hb values due to unreliability. Finally, for practical reasons one may prefer to use a simple linear model because such model might be more user friendly than a more advanced mixed effect model using repeated measurements.

Compared to our previously developed prediction models using Hb deferral as a dichotomous outcome (logistic prediction models),[2] the performance of the linear models developed in the current study was similar. The c-statistic for the logistic model in men in the development data was 0.89 and ranged from 0.88 to 0.89 in a cross-validation in the four geographical regions. For women, the c-statistic for the logistic model in the development data was 0.84 and ranged from 0.80 to 0.84 in a cross-validation. In another study in which the performance of a logistic prediction model for Hb deferral was compared with the performance of a simple linear regression model, there was also no difference observed in discriminative ability.[10]

For the development of the linear models in the current study we considered the same predictive factors as in the previously developed models for Hb deferral risk. However, there are more factors that are associated with Hb levels, for example ethnicity,[11,12] smoking,[13,14] nutrition,[15] physical activity[16-18] and other iron parameters such as ferritin[19,20] or zinc protoporphyrin.[21,22] It may be worthwhile to study the added value of these factors to the linear models in a subsequent study.

Models that predict Hb levels can be used in the management of the donation program by applying them in the invitation process of blood donors. The predicted Hb level for each individual donor should be compared with the sex specific Hb cutoff level for donation. Donors with predicted Hb levels at or above the cutoff level for donation could be invited for a next donation with preference. For donors with predicted Hb levels below the cutoff level, the invitation could be postponed, or another intervention, such as a dietary advice or iron fortification, could be applied. An advantage of the linear models compared to the logistic prediction models is that they can be used worldwide. Hb criteria for donation may vary between countries, and prediction models that predict the risk of Hb deferral may only be valuable in countries with the same Hb cutoff level for donation as used in the development of the prediction models. This disadvantage of logistic models can be overcome by prediction of continuous values of Hb level.

In conclusion, we compared several modeling strategies for the history of Hb levels in whole blood donors to predict Hb levels. We showed that considering multiple Hb measurements gives similar performance as considering only the previous Hb level plus a change in Hb. Previous Hb levels could be summarized with one predictor as the mean value of all previous Hb levels. This predictor can be used in an easy to use simple linear regression model.

Acknowledgements

The authors thank Ingrid Veldhuizen for her contribution in fund raising. Karel G.M. Moons gratefully acknowledges financial contribution for all his research by the Netherlands Organisation for Scientific Research (project 9120.8004 and 918.10.615).

Reference List

1. European Directorate for the Quality of Medicines and HealthCare. Guide to the preparation, use and quality assurance of blood components, European Committee (partial agreement) on Blood Transfusion (CD-P-TS), Recommendation No. R(95) 15, 15th edition. Strasbourg: Council of Europe; 2009, p 1899.
2. Baart AM, de Kort WL, Atsma F, Moons KG, Vergouwe Y. Development and validation of a prediction model for low hemoglobin deferral in a large cohort of whole blood donors. Transfusion 2012;52:2559-69.
3. The Commission of the European Communities. European Commissian Directive 2004/33/EC. 2004.
4. Baart A.M., Atsma F., E.N.McSweeney, Moons K.G.M., Vergouwe Y., Kort W.L.A.M de. External validation and updating of a Dutch prediction model for low hemoglobin deferral in Irish whole blood donors. 2013. Accepted for publication in Transfusion.
5. Cheng J, Edwards LJ, Maldonado-Molina MM, Komro KA, Muller KE. Real longitudinal data analysis for real people: Building a good enough mixed model. Stat.Med. 2009;29:504-20.
6. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann.Intern.Med. 1999;130:515-24.
7. Cox DR. Two further applications of a model for binary regression. Biometrika 1958;45:562-5.
8. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat.Med. 1996;15:361-87.
9. Morris SS, Ruel MT, Cohen RJ, Dewey KG, de la BB, Hassan MN. Precision, accuracy, and reliability of hemoglobin assessment with use of capillary blood. Am.J.Clin.Nutr. 1999;69:1243-8.
10. Zuithoff NPA, Geerlings MI, Baart A.M., Kort W.L.A.M de, Moons K.G.M., Vergouwe Y. Dichotmising continuous outcomes in multiple regression, a bad idea? 2013. Submitted.
11. Perry GS, Byers T, Yip R, Margen S. Iron nutrition does not account for the hemoglobin differences between blacks and whites. J.Nutr. 1992;122:1417-24.
12. Johnson-Spear MA, Yip R. Hemoglobin difference between black and white women with comparable iron status: justification for race-specific anemia criteria. Am.J.Clin.Nutr. 1994;60:117-21.
13. Nordenberg D, Yip R, Binkin NJ. The effect of cigarette smoking on hemoglobin levels and anemia screening. JAMA 1990;264:1556-9.
14. Skjelbakken T, Dahl IM, Wilsgaard T, Langbakk B, Lochen ML. Changes in haemoglobin levels according to changes in body mass index and smoking habits, a 20-year follow-up of a male cohort: the Tromso Study 1974-1995. Eur.J.Epidemiol. 2006;21:493-9.
15. Zimmermann MB, Hurrell RF. Nutritional iron deficiency. Lancet 2007;370:511-20.
16. Cook JD. The effect of endurance training on iron metabolism. Semin.Hematol. 1994;31:146-54.
17. Beard J, Tobin B. Iron status and exercise. Am.J.Clin.Nutr. 2000;72:594S-7S.
18. Ottomano C, Franchini M. Sports anaemia: facts or fiction? Blood Transfus. 2012;10:252-4.
19. Skikne B, Lynch S, Borek D, Cook J. Iron and blood donation. Clin.Haematol. 1984;13:271-87.
20. Pasricha SR, McQuilten ZK, Keller AJ, Wood EM. Hemoglobin and iron indices in nonanemic premenopausal blood donors predict future deferral from whole blood donation. Transfusion 2011;51:2709-13.
21. Harthoorn-Lasthuizen EJ, Lindemans J, Langenhuijsen MM. Zinc protoporphyrin as screening test in female blood donors. Clin.Chem. 1998;44:800-4.
22. Baart AM, de Kort WL, Moons KG, Atsma F, Vergouwe Y. Zinc protoporphyrin levels have added value in the prediction of low hemoglobin deferral in whole blood donors. 2012. Accepted for publication in Transfusion.

Chapter 8

Summary, clinical implications, and perspectives for future research

Summary

Each year, a relevant proportion of the invited blood donors is eventually deferred from donation because of low Hb levels. Although deferrals are meant to protect donors from developing iron deficiency anemia after a blood donation, they also are demoralizing for donors. As a consequence, the risk of donor lapse is increased, although the donor may actually meet the Hb criterion at the time of the next invitation to donate. Early estimation of the risk of Hb deferral on the next visit to the blood collection center could be helpful in the management of the blood donation program.

Chapter 2 describes the development of a first prediction model for Hb deferral in a sample of 5191 Dutch whole blood donors. From these donors, 143 donors (2.8%) were deferred because of a low Hb level. Eleven candidate predictors were considered in logistic regression models to predict Hb deferral. The performance of the prediction model was studied with the c-statistic. Internal validity was assessed with a bootstrap procedure. Strong predictors of Hb deferral were sex, seasonality, Hb level measured at the previous visit, difference in Hb levels between the previous two visits, time since the previous visit, deferral at the previous visit, and the total number of whole blood donations in the past two years. Internal validation showed a c-statistic of 0.87. The prediction model developed in this chapter provides accurate discrimination between donors with low and appropriate Hb levels.

Chapter 3 describes the development of sex-specific prediction models for Hb deferral in a large cohort consisting of all Dutch whole blood donors who fulfilled the inclusion criteria (n=220,946). The same candidate predictors as described in chapter 2 were considered in logistic regression models. Validity of the prediction models was assessed with a cross-validation. A total of 4568 male donors (4.1%) and 8297 female donors (7.7%) were deferred because of a low Hb level. The strongest predictors of Hb deferral were Hb level measured at the previous visit, age, seasonality, difference in Hb levels between the previous two visits, time since the previous visit, deferral at the previous visit, and the total number of whole blood donations in the past two years, for both men and women. The prediction models had a c-statistic of 0.89 for men and 0.84 for women. Cross-validation showed similar results and good calibration.

External validation of prediction models is a necessary step before implementation in practice. In **chapter 4** the sex-specific Dutch prediction models developed in chapter 3 were externally validated and updated in a cohort of Irish whole blood donors. A total of 45,031 Irish whole blood were included in the validation study. Hb cutoff levels for donation were approximately 0.35 mmol/L lower in Ireland than the Dutch cutoff levels (8.07 vs. 8.40 mmol/L in men; 7.45 vs. 7.80 mmol/L in women). The predictive performance

of the models was assessed with calibration plots, calibration-in-the-large and the c-statistic. The models were updated by revising the strength of the individual predictors in the models. In the Irish donor cohort, 613 men (2.4%) and 1624 women (8.4%) were deferred from donation because of a low Hb level. Validation demonstrated underestimation of predicted risks and lower c-statistics for men and women compared to the Dutch cohort. The strength of most predictive factors, particularly previous Hb level, was lower in Irish donors. The updated models showed a c-statistic of 0.83 (95% CI 0.81 – 0.84) for men and 0.76 (95% CI 0.74 – 0.77) for women. Hence, the performance of the Dutch prediction models for Hb deferral was limited when validated in Irish whole blood donors. Updating the models resulted in different predictor effects. This improved mainly the model calibration; the improvement in discrimination was small.

Blood donors that meet the hemoglobin (Hb) criteria for donation may have undetected sub-clinical iron deficiency. **Chapter 5** describes a study in which the prevalence of sub-clinical iron deficiency was assessed in Dutch whole blood donors that were not deferred for low Hb levels. The prevalence of sub-clinical iron deficiency was estimated by measuring ZPP levels. In addition, prevalence rates based on other iron parameters were assessed for comparison. The study population comprised a sample of 5280 Dutch whole blood donors, who passed the Hb criteria for donation. During donor screening, Hb levels were measured in finger stick capillary samples, and venous blood samples were taken for measurements of ZPP and other iron parameters. These parameters included ferritin, transferrin saturation, soluble transferrin receptor (sTfR), hepcidin, erythrocyte mean corpuscular volume (MCV) and mean cell Hb (MCH). Results showed that with a ZPP cutoff level of ≥100 µmol/mol heme, sub-clinical iron deficiency was present in 6.9% of male donors and in 9.8% of female donors. Based on other iron parameters, iron deficiency was also observed. Prevalence rates ranged from 4.8% (based on transferrin saturation) to 27.4% (based on hepcidin concentration) in men, and from 5.6% (based on sTfR concentration) to 24.7% (based on hepcidin concentration) in women. The latter results confirm the presence of sub-clinical iron deficiency among blood donors that meet the Hb criteria for blood donation. This finding needs attention because these donors are at increased risk of developing iron deficiency affecting Hb formation and other cellular processes.

In **chapter 6,** the added value of ZPP levels to the sex-specific prediction models developed in chapter 3 was studied. For this study, data of 4598 Dutch whole blood donors were used. Information on ZPP levels measured at the previous visit was added to the existing prediction models to estimate the risk of Hb deferral. Models were compared using the following measures: c-statistic, continuous net reclassification improvement (NRI), and clinical net benefit (NB). A total of 76 men (2.9%) and 69 women (3.5%) were deferred because of a low Hb level. Previous ZPP level was associated with risk of Hb deferral (OR for interquartile range of previous ZPP level, men: 2.0 (95% CI 1.7 – 2.3); women:

2.2 (95% CI 1.9 – 2.4)) in a multivariable risk model. Addition of ZPP into the models resulted in an increase of the c-statistic from 0.93 to 0.94 for men, and from 0.80 to 0.85 for women. The added value of ZPP was confirmed by measures of clinical usefulness. NRI for men was 0.42, for women 0.56. At relevant threshold probabilities between 10% and 15%, NB was higher for models considering ZPP. These results show that ZPP measurements obtained at the previous visit may have added value in the risk prediction of Hb deferral in whole blood donors.

The prediction models presented in chapters 2-4 and 6 are logistic regression models with the dichotomous outcome Hb deferral yes/no. In **chapter 7**, sex-specific linear regression models were developed using Hb level as a continuous outcome. The prediction models in the previous chapters consider only the Hb level measured at the previous visit plus a change in Hb as information on previous Hb levels. In this study, more information on the history of Hb levels was considered in order to investigate if this could improve the model performance. Data of 187,711 Dutch whole blood donors with sequential Hb measurements during two years were used to develop and internally validate three different regression models: two simple linear models with history of Hb levels included as i) Hb at the previous visit plus change in Hb, or ii) mean of all previous Hb levels; and one mixed effect model including measurements of all previous Hb levels. Thirteen percent of men and 19% of women were deferred because of a low Hb level at least once in two years. The simple linear models and the mixed effect model performed similar, if an estimate of the random intercept of the mixed effect model was used for individual donors to calculate the predicted Hb level. In men, the c-statistic ranged from 0.87 to 0.89 and the R^2 from 0.41 to 0.44. In women, the c-statistic ranged from 0.81 to 0.84. Values of R^2 were higher for the simple linear models than for the mixed effect model, 0.36 and 0.38 vs. 0.29 respectively. These results show that the previous Hb levels could be summarized with one predictor as the mean value of all previous Hb levels. This predictor can be used in an easy to use simple linear regression model.

Results from studies presented in this thesis show that with a limited number of easy-to-measure characteristics the risk of Hb deferral in whole blood donors can be reliably predicted. The predictions made by carefully developed and validated models might be used in the management of the blood donation program.

Clinical implications

The prediction models developed and validated in this thesis can be helpful in the management of the donation program by applying them in the invitation process of blood donors. The number of donor invitations depends on the available blood stock

level. When the level decreases, more donors need to be invited and vice versa. Prediction models for Hb deferral risk can be used to identify blood donors from the donor base eligible to donate. Donors with a low predicted risk of Hb deferral, or with a predicted Hb level equal to or above the Hb cutoff level for donation, can be invited for a donation with preference. Donors with a high predicted risk should not be invited in order to prevent Hb deferral and subsequent demotivation of the donor.

Application of the risk models for Hb deferral requires a choice of threshold or cutoff value of predicted risk above which donors are classified in the high risk group.[1] Using a low threshold value, i.e. a slightly increased risk is already considered as high (and thus preventively not invited), more deferrals will be prevented (higher sensitivity), but at the cost that a lot of donors who could have donated are not invited for a donation (lower specificity). Overcautious, i.e. very low, threshold values could negatively affect the available blood stock level in case uninvited donors are not replaced by eligible donors. Therefore, the number of deferrals that can be prevented must be balanced against the number of donors with appropriate Hb levels that are unnecessarily not invited for a donation. For our models, we found that using a threshold level of 10% risk, the percentage of Hb deferrals in men was decreased to 2.6% compared to 4.1% in the situation without using the model. The percentage of donors that was unnecessary not invited was 10%. In women, the same threshold level could decrease the percentage of Hb deferrals from 7.7% to 5.6%; the percentage of women that was unnecessary not invited was 20%. We consider these percentages of unnecessary postponements acceptable in relation to the decrease in percentages of Hb deferral as long as the number of collected blood donations remains sufficient.

In contrast to the risk models, application of the models for continuous values of Hb level does not require the choice of a cutoff value to divide donors into low and high risk groups. The predicted Hb levels can just be compared with the sex-specific cutoff level for donation. Now, the variability in predicted Hb levels may cause misclassifications, i.e. donors with low Hb levels are invited (because the predicted Hb level was equal to or above the Hb cutoff level for donation) and donors with appropriate Hb levels are not invited (because the predicted Hb level was below the Hb cutoff level for donation). The discriminative ability of the models for Hb level was studied by comparing the predicted Hb levels with the observed Hb levels dichotomized at the cutoff level for donation. We found comparable discriminative ability for the risk models and the models predicting Hb level.

In conclusion, donors with a low risk of Hb deferral or with a predicted Hb level that meets the criteria for donation can thus be invited for a next donation. For donors with high risks of Hb deferral or with predicted Hb levels that are too low, the invitation could be postponed. This gives the donor more time to recover from the previous donation, and the invitation might be postponed to another season in which Hb levels are generally higher. Another intervention might be a dietary advice in order to increase the uptake of

iron. Finally, the donor may be advised to switch from whole blood donation to plasma donation. Although the donor is lost as a whole blood donor, the donor can still be valuable for the blood bank. Eventually, application of the models may result in a decrease in the number of donor deferrals for low Hb levels. Thus, the prediction models could contribute in improving donor base management and donor health care, and increase donor satisfaction on top of that.

Perspectives for future research

The prediction models presented in this thesis enable an adequate prediction of Hb deferral and continuous values of Hb level in whole blood donors. However, we should be careful to immediately implement these models in practice.

The good discriminative ability of the models might be improved by including extra predictors in the models. Improvement of discrimination may result in less misclassification. Additional research could hence focus on the discovery of new predictors of Hb deferral. As a start, we developed prediction models with data that we could easily obtain from the donor database. However, there may be more factors that are associated with Hb levels which can also be studied as candidate predictors. These factors include for example ethnicity,[2,3] smoking,[4,5] nutrition,[6] and physical activity.[7-9] In addition, genetic factors may influence Hb levels as well.[10] Furthermore, iron parameters that can detect iron depletion or iron deficient erythropoiesis, early stage of iron deficiency, might be predictive for Hb deferral. We indeed found that addition of ZPP levels measured in venous samples on the previous visit might have added value in the prediction of Hb deferral. Our relatively small study sample implies that more research on the added value of ZPP levels is necessary along with research on the possible added value of other iron parameters.

For practical reasons it is often preferred to develop a model with as little as possible predictors, and/or with predictors that can easily be measured or obtained, whilst the predictive performance of the model remains still adequate.[1] It may also be possible that first a simple model is used to roughly divide donors into low, intermediate and high risk categories. Information on new predictors may alter the predicted risk to some extent, which will only be relevant for the intermediate risk group. Low predicted risks will not change that much that a donor moves from the low risk category to the high risk category. Likewise, high predicted risks will not change that much that a donor moves from the high risk category to the low risk category. For donors with an intermediate risk of Hb deferral, an extended model with more (complex) predictors can be used to discriminate better between donors with low and appropriate Hb levels. We have suggested that the model with ZPP level added might be used for this purpose.

Once a final set of predictors is combined in one or two (i.e. a simple and extended) prediction models, these models should be externally validated and if necessary updated

to local circumstances.[11-18] The predictive accuracy of the models should ideally be studied across different countries which use either the same or different Hb criteria for donation, preferably at a different (more recent) point in time before wide implementation in practice. Furthermore, we developed our prediction models in donors who had given at least two whole blood donation in the past. The prediction model should also be validated in newly registered donors and plasma donors.

Finally, when a developed and validated, and if needed updated, prediction model shows adequate predictive accuracy, a so-called impact study is warranted.[17,19-21] In an impact study, the effect of using the model on the number of Hb deferrals, blood stock levels, and costs can be compared with a setting in which the prediction model is not used. Such impact study may be performed with a cluster randomized trial.[18] In such a prospective impact study, blood collections centers (clusters) are randomized. The clusters are divided into clusters where donors are invited as usual (without using the prediction model) and clusters where the prediction model is applied for the selective invitation of blood donors. After a follow-up period, the effect of using the model can be compared with the usual invitation process. Otherwise, decision analytic modelling or cost-effectiveness modelling studies can be used, without prospectively using the model in a new group of blood donors.[18,22] Positive results of an impact study will bring a strong support for implementation of the prediction model in practice.

Reference List

1. Steyerberg EW. Clinical prediction models. New York: Springer; 2009.
2. Perry GS, Byers T, Yip R, Margen S. Iron nutrition does not account for the hemoglobin differences between blacks and whites. J.Nutr. 1992;122:1417-24.
3. Johnson-Spear MA, Yip R. Hemoglobin difference between black and white women with comparable iron status: justification for race-specific anemia criteria. Am.J.Clin.Nutr. 1994;60:117-21.
4. Nordenberg D, Yip R, Binkin NJ. The effect of cigarette smoking on hemoglobin levels and anemia screening. JAMA 1990;264:1556-9.
5. Skjelbakken T, Dahl IM, Wilsgaard T, Langbakk B, Lochen ML. Changes in haemoglobin levels according to changes in body mass index and smoking habits, a 20-year follow-up of a male cohort: the Tromso Study 1974-1995. Eur.J.Epidemiol. 2006;21:493-9.
6. Zimmermann MB, Hurrell RF. Nutritional iron deficiency. Lancet 2007;370:511-20.
7. Cook JD. The effect of endurance training on iron metabolism. Semin.Hematol. 1994;31:146-54.
8. Beard J, Tobin B. Iron status and exercise. Am.J.Clin.Nutr. 2000;72:594S-7S.
9. Ottomano C, Franchini M. Sports anaemia: facts or fiction? Blood Transfus. 2012;10:252-4.
10. Soranzo N, Spector TD, Mangino M, Kuhnel B, Rendon A, Teumer A, Willenborg C, Wright B, Chen L, Li M, et al. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. Nat.Genet. 2009;41:1182-90.
11. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat.Med. 1996;15:361-87.
12. Harrell FE, Jr. Regression modeling strategies: with applications to linear models, logistic regression, and survival analysis. New York: Springer; 2001.
13. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. Ann.Intern.Med. 1999;130:515-24.
14. Altman DG, Vergouwe Y, Royston P, Moons KG. Prognosis and prognostic research: validating a prognostic model. BMJ 2009;338:b605.
15. Steyerberg EW, Borsboom GJ, van Houwelingen HC, Eijkemans MJ, Habbema JD. Validation and updating of predictive logistic regression models: a study on sample size and shrinkage. Stat.Med. 2004;23:2567-86.
16. Janssen KJ, Moons KG, Kalkman CJ, Grobbee DE, Vergouwe Y. Updating methods improved the performance of a clinical prediction model in new patients. J.Clin.Epidemiol. 2008;61:76-86.
17. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. J.Clin.Epidemiol. 2008;61:1085-94.
18. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, Woodward M. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98:691-8.
19. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. Ann.Intern.Med. 2006;144:201-9.
20. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009;338:b606.
21. Moons KG, Kengne AP, Grobbee DE, Royston P, Vergouwe Y, Altman DG, Woodward M. Risk prediction models: II. External validation, model updating, and impact assessment. Heart 2012;98:691-8.
22. Schaafsma JD, van der GY, Rinkel GJ, Buskens E. Decision analysis to complete diagnostic research by closing the gap between test characteristics and cost-effectiveness. J.Clin Epidemiol. 2009;62:1248-52.

Nederlandse samenvatting

Nederlandse samenvatting

Jaarlijks wordt een aanzienlijk deel van de bloeddonors die zijn opgeroepen om te doneren afgekeurd voor een donatie vanwege een te laag hemoglobine (Hb)-gehalte: een Hb-afkeuring. Hoewel afkeuringen bedoeld zijn om donors te beschermen tegen het ontwikkelen van ijzerdeficiënte anemie na een bloeddonatie, hebben ze ook een ontmoedigend effect op donors. Ten gevolge van een afkeuring neemt de kans toe dat een donor stopt met doneren, hoewel de donor bij een volgende oproep voor donatie heel wel aan het Hb-criterium zou kunnen voldoen. Het vroegtijdig schatten van de kans op Hb-afkeuring bij het eerstvolgende bezoek aan een afnamelocatie zou behulpzaam kunnen zijn bij het managen van het bloeddonatieprogramma.

Hoofdstuk 2 beschrijft de ontwikkeling van een eerste predictiemodel voor Hb-afkeuring in een steekproef van 5191 Nederlandse volbloeddonors. Van deze donors werden er 143 (2,8%) afgekeurd vanwege een te laag Hb-gehalte. Elf kandidaat-predictoren werden bekeken in logistische regressiemodellen om Hb-afkeuring te voorspellen. Het voorspellend vermogen van het predictiemodel werd bestudeerd met de *c-statistic*. De interne validiteit werd vastgesteld met een *bootstrap procedure*. Sterke predictoren van Hb-afkeuring waren geslacht, seizoensinvloeden, het Hb-gehalte gemeten tijdens het vorige bezoek, het verschil in Hb-gehalte tussen de twee voorafgaande bezoeken, de tijd sinds het vorige bezoek, afkeuring tijdens het vorige bezoek, en het aantal volbloeddonaties in de afgelopen twee jaar. Interne validatie liet een *c-statistic* zien van 0,78. Het predictiemodel dat in dit hoofdstuk is ontwikkeld voorziet in accurate discriminatie tussen donors met een te laag en een voldoende Hb-gehalte.

Hoofdstuk 3 beschrijft de ontwikkeling van geslachtsspecifieke predictiemodellen voor Hb-afkeuring in een groot cohort bestaande uit alle Nederlandse volbloeddonors die aan de inclusiecriteria voldeden (n=220.946). Dezelfde kandidaat-predictoren als beschreven in hoofdstuk 2 werden bekeken in logistische regressiemodellen. De validiteit van de predictiemodellen werd vastgesteld met een kruis-validatie. In totaal werden 4568 (4,1%) mannelijke donors en 8297 (7,7%) vrouwelijke donors afgekeurd vanwege een te laag Hb-gehalte. De sterkste predictoren van Hb-afkeuring waren, voor zowel mannen als vrouwen, het Hb-gehalte gemeten tijdens het vorige bezoek, leeftijd, seizoensinvloeden, het verschil in Hb-gehalte tussen de twee voorafgaande bezoeken, de tijd sinds het vorige bezoek, afkeuring tijdens het vorige bezoek, en het aantal volbloeddonaties in de afgelopen twee jaar. De predictiemodellen hadden een *c-statistic* van 0,89 voor mannen en 0,84 voor vrouwen. De kruis-validatie liet vergelijkbare resultaten zien en goede calibratie.

Externe validatie van predictiemodellen is een noodzakelijke stap voordat zij geïmplementeerd kunnen worden in de praktijk. In **hoofdstuk 4** worden de geslachtsspecifieke predictiemodellen die in hoofdstuk 3 ontwikkeld zijn extern gevalideerd en aangepast in een cohort van Ierse volbloeddonors. In totaal werden 45.031 Ierse volbloeddonors geïncludeerd in de validatiestudie. Hb-grenswaarden voor donatie waren in Ierland ongeveer 0,35 mmol/L lager dan de Nederlandse grenswaarden (8,07 vs. 8,40 mmol/L voor mannen; 7,45 vs. 7,80 mmol/L voor vrouwen). Het voorspellend vermogen van de modellen werd vastgesteld met calibratie plaatjes, *calibration-in-the-large* en de *c-statistic*. De modellen werden aangepast door de sterkte van de individuele predictoren in de modellen te herzien. In het Ierse donorcohort werden 613 mannen (2,4%) en 1624 vrouwen (8,4%) afgekeurd voor een donatie vanwege een te laag Hb-gehalte. Validatie liet een onderschatting van het voorspelde risico zien, en een lagere *c-statistic* voor mannen en vrouwen in vergelijking met het Nederlandse cohort. De sterkte van de meeste predictieve factoren, en vooral van het Hb-gehalte gemeten tijdens het vorige bezoek, was lager bij de Ierse donors. De aangepaste modellen lieten een *c-statistic* zien van 0,83 (95% CI 0,81 – 0,84) voor mannen en 0,76 (95% CI 0,74 – 0,77) voor vrouwen. Dus, het voorspellend vermogen van de Nederlandse modellen voor Hb-afkeuring was beperkt wanneer zij werden gevalideerd bij Ierse volbloeddonors. Aanpassing van de modellen resulteerde in andere predictoreffecten. Hierdoor werd voornamelijk de calibratie van het model verbeterd; de verbetering in discriminatie was klein.

Bloeddonors die voldoen aan de Hb-criteria voor donatie hebben mogelijk onopgemerkte subklinische ijzerdeficiëntie. **Hoofdstuk 5** beschrijft een studie waarin de prevalentie van subklinische ijzerdeficiëntie werd vastgesteld bij Nederlandse volbloeddonors die niet waren afgekeurd vanwege een te laag Hb-gehalte. De prevalentie van subklinische ijzerdeficiëntie werd geschat door het ZPP-gehalte te meten. Daarnaast werden ter vergelijking prevalenties vastgesteld op basis van andere ijzerparameters. De studiepopulatie bestond uit een steekproef van 5280 Nederlandse volbloeddonors, die voldeden aan de Hb-criteria voor donatie. Het Hb-gehalte werd tijdens de donorkeuring gemeten in een druppel bloed die met een vingerprik was afgenomen. Daarnaast werden ook veneuze bloedmonsters afgenomen voor het meten van ZPP en andere ijzerparameters. Andere parameters waren: ferritine, transferrine saturatie, *soluble transferrin receptor* (sTfR), hepcidine, *erythrocyte mean corpuscular volume* (MCV) en *mean cell Hb* (MCH). De resultaten lieten zien dat er met een ZPP-grenswaarde van ≥100 µmol/mol heem sprake was van subklinische ijzerdeficiëntie bij 6,9% van de mannelijke donors en bij 9,8% van de vrouwelijke donors. Op basis van andere ijzerparameters werd er ook ijzerdeficiëntie waargenomen. De prevalenties liepen bij mannen uiteen van 4,8% (gebaseerd op de transferrine saturatie) tot 27,4% (gebaseerd op de hepcidine concentratie), en bij vrouwen van 5,6% (gebaseerd op de sTfR-concentratie) tot 24,7% (gebaseerd op de hepcidine concentratie). Deze laatstgenoemde resultaten bevestigen het bestaan van subklinische

ijzerdeficiëntie bij volbloeddonors die voldoen aan de Hb-criteria voor bloeddonatie. Deze bevinding verdient aandacht omdat deze donors een verhoogde kans hebben op het ontwikkelen van ijzerdeficiëntie, wat de vorming van Hb en andere cellulaire processen aantast.

In **hoofdstuk 6** werd de toegevoegde waarde bestudeerd van ZPP aan de in hoofdstuk 3 ontwikkelde geslachtsspecifieke predictiemodellen. Voor deze studie werden gegevens van 4598 Nederlandse volbloeddonors gebruikt. Informatie over het ZPP-gehalte gemeten tijdens het vorige bezoek werd toegevoegd aan de bestaande predictiemodellen voor het schatten van de kans op Hb-afkeuring. Modellen werden vergeleken aan de hand van de volgende maten: *c-statistic*, *continuous net reclassification improvement* (NRI), en *clinical net benefit* (NB). In totaal werden 76 mannen (2,9%) en 69 vrouwen (3,5%) afgekeurd vanwege een te laag Hb-gehalte. Het ZPP-gehalte gemeten tijdens het vorige bezoek was geassocieerd met de kans op Hb-afkeuring (OR voor de interkwartiel range van het "vorige ZPP-gehalte" was 2,0 (95% CI 1,7 – 2,3) voor mannen, en 2,2 (95% CI 1,9 – 2,4) voor vrouwen) in een multivariabel predictiemodel. Toevoeging van ZPP aan de modellen resulteerde in een toename van de *c-statistic* van 0,93 tot 0,94 voor mannen, en van 0,80 tot 0,85 voor vrouwen. De toegevoegde waarde van ZPP werd bevestigd door maten van klinisch nut. De NRI voor mannen was 0,42 en voor vrouwen 0,56. De NB was hoger voor modellen met ZPP. Deze resultaten laten zien dat ZPP-gehaltes gemeten tijdens het vorige bezoek van toegevoegde waarde kunnen zijn bij het voorspellen van de kans op Hb-afkeuring bij volbloeddonors.

De predictiemodellen die zijn gepresenteerd in de hoofdstukken 2-4 en 6 zijn logistische regressiemodellen met de dichotome uitkomst Hb-afkeuring ja/nee. In **hoofdstuk 7** werden geslachtsspecifieke lineaire regressiemodellen ontwikkeld met het Hb-gehalte als continue uitkomst. De predictiemodellen in de voorgaande hoofdstukken bekijken als informatie over eerdere Hb-gehaltes alleen het Hb-gehalte gemeten tijdens het vorige bezoek plus een verandering in Hb-gehalte. In deze studie werd meer informatie over de geschiedenis van Hb-gehaltes bekeken met het doel om te onderzoeken of dit het voorspellend vermogen van het model kan verbeteren. Gegevens van 187.711 Nederlandse volbloeddonors met herhaalde Hb-metingen in een periode van twee jaar werden gebruikt voor het ontwikkelen en intern valideren van drie verschillende regressiemodellen: twee eenvoudige lineaire modellen met de geschiedenis van Hb-gehaltes daarin opgenomen als i) Hb-gehalte gemeten tijdens het vorige bezoek plus een verandering in Hb-gehalte, of ii) het gemiddelde van alle eerder gemeten Hb-gehaltes; en één *mixed effect model* met daarin opgenomen alle eerder gemeten Hb-gehaltes. Dertien procent van de mannen en 19% van de vrouwen werd tenminste één keer afgekeurd vanwege een te laag Hb-gehalte in een periode van twee jaar. Het voorspellend vermogen van de eenvoudige lineaire modellen en het *mixed effect model* was vergelijkbaar, wanneer voor

individuele donors een schatting van het *random intercept* van het *mixed effect model* werd gebruikt om het voorspelde Hb-gehalte te berekenen. Voor de modellen voor mannen liep de *c-statistic* uiteen van 0,87 tot 0,89, en de R^2 van 0,41 tot 0,44. Voor de modellen voor vrouwen liep de *c-statistic* uiteen van 0,81 tot 0,84. Waarden voor de R^2 waren hoger voor de eenvoudige lineaire modellen dan voor het *mixed effect model*, respectievelijk 0,36en 0,38 vs. 0,29. Deze resultaten laten zien dat eerdere Hb-gehaltes samengevat kunnen worden in één predictor als het gemiddelde van alle eerder gemeten Hb-gehaltes. Deze predictor kan gebruikt worden in een gemakkelijk te gebruiken eenvoudig lineair regressiemodel.

Resultaten van studies die in dit proefschrift zijn gepresenteerd laten zien dat de kans op Hb-afkeuring bij volbloeddonors met een beperkt aantal gemakkelijk te bepalen karakteristieken op betrouwbare wijze voorspeld kan worden. De voorspellingen die met behulp van zorgvuldig ontwikkelde en gevalideerde modellen gemaakt zijn zouden gebruikt kunnen worden bij het managen van het bloeddonatieprogramma.

Dankwoord

Dankwoord

Op het moment dat ik dit dankwoord schrijf is het einde van mijn promotietraject in zicht. Ik denk kort terug aan de eindfase van mijn studie aan de Wageningen Universiteit, aan een fantastische ervaring bij een van mijn afstudeeronderzoeken tijdens een medische expeditie in Nepal.
Stap voor stap kom ik hoger. Af en toe moet ik even stilstaan om weer op adem te komen. In een vast ritme ben ik al een paar uur onderweg naar de top van de Mera Peak, een berg van 6476m hoogte in de Himalaya. In de afgelopen weken heb ik met de andere expeditieleden genoten van de trektocht naar het basiskamp, het overschrijden van hoge bergpassen en de mooie uitzichten onderweg. We hebben echter ook te maken gehad met hoogteziekte en slecht weer. Afgelopen nacht raasde er een sneeuwstorm over de tenten. Maar in de vroege ochtend is de storm gaan liggen en is het weer opgeklaard: een ideale dag voor een toppoging. En nu ben ik nog maar een paar stappen verwijderd van de top. Stap voor stap klim ik de laatste meters omhoog. En dan kan ik geen volgende stap meer zetten: ik sta op de top!
Mijn promotietraject was ook een soort expeditie. Onderweg heb ik veel geleerd en nieuwe gebieden ontdekt. Er waren hoogtepunten, bijvoorbeeld als analyses mooie resultaten opleverden of als een artikel werd geaccepteerd voor publicatie, maar soms zat het ook weleens tegen. En uiteindelijk is met het voltooien van dit proefschrift ook een nieuwe top bereikt.

In de afgelopen jaren hebben veel mensen bijgedragen aan dit proefschrift. Graag wil ik de volgende personen bedanken.

In de eerste plaats mijn promotor en co-promotoren.
Prof. dr. K.G.M. Moons. Beste Carl, jouw uitgebreide kennis van predictieonderzoek en methodologie hebben een waardevolle bijdrage geleverd aan dit proefschrift. We hebben elkaar niet vaak gesproken, maar je was altijd bereikbaar en je kon snel antwoord geven op vragen, wat erg prettig was.
Dr. Y. Vergouwe. Beste Yvonne, ik heb erg veel aan je gehad als begeleider. Je hebt me veel geleerd over alle facetten van predictieonderzoek en jouw commentaar op mijn artikelen was altijd erg waardevol. Ik ben je dankbaar dat je mij bent blijven begeleiden toen je een aanstelling kreeg bij het Erasmus MC in Rotterdam.
Dr. W.L.A.M. de Kort. Beste Wim, jouw enthousiasme voor dit project werkte erg stimulerend. Je had veel goede ideeën voor verder onderzoek en als we die allemaal hadden uitgevoerd zou dit boekje twee keer zo dik zijn geworden. Jij was niet alleen behulpzaam bij het onderzoek, maar ook kon ik voor andere dingen altijd bij je terecht. Veel dank daarvoor!

Daarnaast hebben verschillende andere mensen mij begeleid. In chronologische volgorde:
Dr. I.J.T. Veldhuizen. Beste Ingrid, jij hebt ervoor gezorgd dat ik een goede start kon maken bij Sanquin.
Dr. C.J.M. Doggen. Beste Carine, al vrij snel nam jij de begeleiding over van Ingrid. Ik vond het erg prettig om met je samen te werken. In de anderhalf jaar dat je mij hebt begeleid heb ik veel van je geleerd. Ook als persoon mocht ik je erg graag. Het kwam dan ook als een donderslag bij heldere hemel toen jij bekend maakte dat je een nieuwe baan had geaccepteerd bij de Universiteit Twente en je mij niet langer kon begeleiden.
Dr. F. Atsma. Beste Femke, toen Carine Sanquin had verlaten nam Wim de begeleiding over, en anderhalf jaar later werd jij ook betrokken bij mijn promotieonderzoek. Bedankt voor je begeleiding in de laatste jaren.

Many thanks to the other co-authors: dr. E.N. McSweeney, dr. P.A.H. van Noord, en ing. E.T. Wiegerinck.

Leden van de leescommissie: prof. dr. Ph.G. de Groot, prof. dr. K.C.B Roes, prof. dr. G.A. Zielhuis, en prof. dr. D.W. Swinkels. Bedankt voor het lezen en beoordelen van mijn manuscript.

Giny Hetterscheit, Mirianne van Gerwen en alle andere medewerkers van het QC lab. Bedankt voor jullie hulp bij het verrichten van de ZPP-bepalingen en het meten van diverse bloedcel-indices.

Dr. N.P.A. Zuithoff. Beste Peter, het was fijn dat ik bij jou terecht kon met vragen over tips en trucs in SPSS. Het is mooi dat jij ook gebruik hebt kunnen maken van mijn dataset voor een van jouw onderzoeken.

Kazem Nasserinejad. You are also working on prediction models for low Hb levels in whole blood donors. It is nice that you could use my data for one of your interesting studies. I wish you good luck to finish your PhD thesis!

Collega's van Sanquin Research, afdeling Donorstudies: Wim de Kort, Ingrid Veldhuizen, Carine Doggen, Femke Atsma, Pieternel Pasker, Paul van Noord, Katja van den Hurk, Anneke Dijkstra, Anne van Dongen, Anne Wevers, Karlijn Peffer, Nienke Schotten, Maurits Hoogerwerf, Elze Wagenmans, Jacqueline Melssen, Karin Habets en Karen van den Toren. En ook alle andere collega's van Sanquin. Jullie hebben direct of indirect bijgedragen aan dit proefschrift, en dankzij jullie was de sfeer op het werk erg gezellig.
Anne van Dongen, we begonnen en eindigden ons promotieonderzoek vrijwel tegelijk, en bovendien waren we al die tijd kamergenoten. Ik vind het dan ook niet meer dan logisch dat jij mijn paranimf bent.

Mijn carpoolgenoten: Gerhard Zielhuis, Linda Garms, Cootje Vermaat, Mariëtte Hooiveld, Judith Vos, Irene Loman en Maurits Hoogerwerf. Dankzij jullie was het op en neer reizen van Wageningen naar Nijmegen erg aangenaam en gezellig. Gerhard, het was leuk om onderweg ook inhoudelijk over mijn promotieonderzoek te praten. Bedankt voor je adviezen en je belangstelling voor mijn project.

Clubgenoten van Wageningse atletiekvereniging Pallas'67 en triatlonvereniging VZC E&P in Veenendaal. De gezamenlijke trainingen hebben me enorm geholpen om na een lange werkdag mijn hoofd leeg te kunnen maken en de volgende dag vol nieuwe energie weer aan het werk te gaan. Kirsten, Frederike, Nina, Ingrid, Petra en Gretha, ook bedankt voor alle gezelligheid buiten de trainingen!

Vrienden en vriendinnen, en in het bijzonder Marleen Newmei, Nicole Voet en Sjouke Venema.
Lieve Marleen, ook al zien we elkaar niet zo vaak, ik vind onze vriendschap erg waardevol. Het is fijn dat ik jou altijd kon bellen en je er altijd voor mij was als ik daar behoefte aan had. Het is jammer dat je waarschijnlijk in India gaat wonen, maar gelukkig hebben ze daar ook telefoon!
Lieve Nicole, ik heb jou in de laatste twee jaar van mijn promotietraject leren kennen en daar ben ik erg blij om. We hebben veel gemeen en het was fijn om de ups en downs van promotieonderzoek met je te kunnen delen. Ik hoop op vele jaren vriendschap!
Lieve Sjouke, het samen sporten, de fijne weekendjes weg en de mooie vakanties in de afgelopen jaren hebben mij erg goed gedaan. Het is erg leuk om samen met jou bergen te bedwingen en grenzen te verleggen. Ik kijk erg uit naar de volgende uitdaging, de volgende expeditie. Nu dit boekje over ijzer af is, is het volbrengen van een Ironman een mooi toekomstig project. Met jouw ervaring kun jij mij daar vast goed bij begeleiden!

En tenslotte mijn ouders en broertje. Pa en ma, bedankt voor de mogelijkheid die jullie me hebben gegeven om mij te ontwikkelen tot de persoon die ik nu ben. Stephan, bedankt voor alle interesse en support. Ik ben blij dat je mijn paranimf wilt zijn.

Curriculum Vitae

Curriculum Vitae

Alexandra Mireille Baart was born on September 8th 1981 in Rotterdam, the Netherlands. She graduated from secondary school at Stedelijk Gymnasium Schiedam in 1999. Afterwards, she started her study Nutrition and Public Health at Wageningen University. During this period, she conducted two research projects at the department of Human and Animal Physiology. The first project was about the effect of gonadal steroids on growth hormone receptor expression and the number of somatostatin containing cells in the periventricular nucleus in the female rat. The second project was about heart rate variability at high altitude. For the latter project, she joined a medical expedition in the Himalayas in Nepal organized by Medex. She obtained her Master of Science degree in 2004. Thereafter, she attended the first year of the Selective Utrecht Medical Master (SUMMA) program at Utrecht University. In 2007, she started working as a PhD student at Sanquin Research, at the department of Donor Studies in Nijmegen, in collaboration with the Julius Center for Health Sciences and Primary Care at the University Medical Center in Utrecht. She combined her PhD research project with the postgraduate master program Clinical Epidemiology at Utrecht University. She obtained her Master of Science degree in Clinical Epidemiology in 2010. Her work as a PhD student has resulted in the studies presented in this thesis.