FinDonor 10 000 study: A cohort to identify iron depletion and factors affecting it in Finnish blood donors

Short title: Finnish blood donor cohort

Partanen, Jukka*

Niittymäki, Pia

Nikiforow, Nina

Palokangas, Elina

Larjo, Antti

Mattila, Pirkko

Castrén, Johanna

Arvas, Mikko*

Finnish Red Cross Blood Service, Research and Development. Kivihaantie 7, 00310 Helsinki, Finland.

*address correspondence to: mikko.arvas@veripalvelu.fi or jukka.partanen@veripalvelu.fi

Key words:

blood donors, health, cohort, iron, blood donation

Abstract

Background

There is increasing evidence that frequent blood donation depletes the iron stores of some but not all blood donors. To identify risk groups and factors affecting the iron stores in the Finnish donor population the Fin Donor 10 000 project, a prospective study observing blood donor iron stores and genetic and lifestyle factors associated with iron stores, was set up.

Material and methods

2584 blood donors (N= 8003 samples) were recruited in the study alongside the standard donation at three fixed donation sites in the capital region of Finland during 5/2015 – 12/2017. All participants were asked to fill in a pseudonymized questionnaire about their health and lifestyle; 2562 donors (99.1%) completed it. Blood samples were collected from the sample pouch of whole blood collection set, kept in cool temperature and processed centrally. The samples were sent to a clinical laboratory for analysis of whole blood count, CRP, ferritin and sTFR; in addition, genomic DNA was isolated for GWAS studies.

Results

The demographics of the participants, albeit in general similar to the general blood donor population in Finland, indicated some bias toward older and more frequent donors, who may be regarded as regular, strongly-committed donors. The effects of the time lag from the sampling to the analysis and the time of the day when sample was drawn was studied and revealed small but significant time-dependent changes in certain measurements.

Discussion

If these time dependent changes are not adequately corrected they may affect the conclusions drawn from similar studies. The FinDonor cohort now provides us with tools to identify the potential risk groups and genetic and non-genetic factors behind their tendency to iron deficiency.

Introduction

Approximately 200 - 250 mg of iron is drawn in a standard blood donation; the amount accounts for 25% of average tissue iron stores in men and up to 75% in women. There is compelling evidence that a portion of blood donors may become iron depleted or deficient. Deferral rates of presenting donors, due to too low a level of hemoglobin, vary a lot depending on populations and policies from below 10 % to up to 20 %. In addition, for example, in the U.S. 35% of frequent blood donors were found to be iron deficient [1–6]. Adverse health effects, with various symptoms, such as fatigue, pica, restless legs syndrome and cognitive problems, have been linked to iron deficiency and anemia [7], though the effect of iron deficiency without anemia in otherwise healthy individuals is still unclear [8]. The iron removed by blood donation should be replaced by dietary iron. Some blood banks also provide iron replacement for donors. Another tool to ensure correction of iron stores is the minimum time interval between blood donations, which must be balanced between the donor health issues and blood demand. There is evidence [2,4,9] that the current intervals may not be sufficient for iron or Hb recovery, at least to all donors. Genome screening studies [10,11] on various populations, but so far not directly on blood donors, have indicated a number of genes associated with red blood cell indices. The relevance of these results to the response to frequent blood donation is open. Identification of donors who are in the highest risk for developing iron depletion or anaemia as a result of frequent blood donation is warranted as it could give us fact-based tools to correctly steer the recruitment of donors.

To complement the above mentioned studies on iron depletion and cohorts collected elsewhere [12–14], we set up a cohort study focusing primarily on measurement of iron stores and factors affecting to them. The study recruited in three blood donation sites in the Helsinki metropolitan area. Any blood donor in these sites was qualified to participate. Based on our sociological study the Finnish blood donors are willing to donate also for research use [15]; in general, blood donors have been regarded as an important resource for biomedical studies [16]. Consent for laboratory and genetics analyses was asked and the participants were asked to fill in a health and lifestyle questionnaire via internet. We here describe the cohort and give evidence that there are factors in the measurements that may affect the results if not carefully considered.

Materials and methods

Finnish Red Cross Blood Service (FRCBS) is a national blood establishment responsible for the collection, processing and distribution of blood products in Finland. Ethical permit was obtained from the Ethical Board of Helsinki University Hospital, Helsinki, Finland and an inform consent

was obtained from the participants and they were allowed to withdraw from the study at any time.

Sample collection

Fin Donor 10 000 study was started 18th May 2015 and lasted till the end of year 2017. All donors in the participating three donation sites located in the Helsinki metropolitan area (Kivihaka, Sanomatalo, Espoo) were informed about the possibility to join the study with leaflets, posters and Facebook –postings. A permanent deferral was the only criteria for an exclusion from the study.

Each participant was asked to fill up a questionnaire on health and life style issues using an internet portal provided at the donation site. The questions are shown in Supplementary Table 1 and are similar to those used e.g. in the Danish blood donor studies (Pedersen 2012).

Samples of peripheral blood were collected from the diversion pouch (CompoFlow® Quadruple T&B, Fresenius Kabi, Germany). For blood counts and genomic dna extraction samples were collected to K2-EDTA tubes (Terumo Europe NV, Belgium) and for CRP, ferritin and sTfR measurements to Lithium heparin tubes (Terumo Europe NV, Belgium). A cell pellet was stored for genomic DNA preparation. Samples together with the signed informed consent was sent to the FRCBS headquarters, located in Kivihaka, Helsinki, Finland, where the identity of the donor and coding on the tubes were double-checked and added to the research database. Plasma samples were sent for laboratory measurements (blood count, CRP, ferritin and sTfR; see Supplementary Table 2 for further details) in batches twice a day to the clinical laboratory of University of Helsinki Central Hospital. Blood counts were measured with Sysmex XE (Sysmex, Japan). The measurements are accredited by the local authority FINAS and the laboratory participates in national and international quality assurance rounds. The tubes were kept in cool storage between transportations and handling and time points from blood drawing to the analysis were recorded.

Genotyping

The first 760 individuals enrolled to the cohort were genotyped using an Illumina HumanCoreExome-24v1-1_A beadchip (Illumina, USA) to estimate ethnic diversity and relatedness of the participants. Genotyping was performed at the FIMM Technology Centre, Helsinki, Finland. DNA samples from the Finnish cohort were extracted using the QIAamp DNA Blood Mini Kit (Qiagen, Germany). Imputation was carried out as in [17] according to instructions of [18] using plink 1.90b3.29 [19] for quality filtering and IMPUTE2 [20] for the actual imputation with 1000 Genomes Phase 3 [21] as a phased reference. Post-imputation filtering excluded variants

having an IMPUTE2 INFO-field <0.5. After post-imputation filtering 11 880 443 variants were included. PCA (Principal Component Analysis) of the imputed and filtered dataset was carried out with eigensoft version 6.1.4 [22]. The datasets analyzed during the current study are not publicly available due to limitations of ethical permits which do not allow distribution of personal data, including individual genetic and clinical results.

Blood count data analysis

Effect of donation time and sampling-to-analysis delay were analyzed using a generalized linear multiple regression model. Blood count measurement data were used as a response variable and the donation time, sampling-to-analysis delay, sex and age group (ten year bins) were used as explanatory variables. CRP and ferritin were log2-transformed. Graphical examination of the data suggested interactions between the explanatory factors donation time and sampling-to-analysis delay. Hence, first a model that includes an interaction term for them was tested and if the interaction was not found significant a model without an interaction term was fitted to calculate coefficients of the explanatory variables. The significance of p-values was estimated with Bonferroni correction for multiple testing. Effect of repeated measures was tested by fitting mixed effect models using lme4 R-library [23]. Mixed effect models were identical to the no-interaction models with an additional explanatory variable of individuals as a random effect. Qualitatively similar effect sizes were obtained from mixed effect models in comparison to no-interaction models. The data were analyzed using R version 3.4.3 [24] and plotted with the ggplot2 R-library [25].

Results

In total 2584 individual donors (1015 men, 1569 women) gave consent to participate in the study and donated at least one sample. This represent 6 % of all donors who visited the study sites during study. Of them 2562 (1004 men, 1558 women) answered the questionnaire. In total the participants donated 8003 samples. None of the participants has withdrawn his or her consent so far (12/2018). 174 were new or reactivated (previous donation more than 2 years before) donors (see Supplementary Table 3 for further details). Distributions of sex, age and hemoglobin levels of the study population (Group "Study" in Figure 1 and 2) were similar to the entire national donor population (Group "All" in Figure 1 and 2) and to the donor population of the recruitment sites (Group "Study Sites in Figure 1 and 2) during the study period. The age distribution on the study population resembles more the national donor population than the donor population of the study sites (Figure 1A). The study population, however, consisted of a markedly higher proportion of donors with frequent donation activity and fewer new donors (Figure 2).

FinDonor study cohort

At least two blood samples could be collected from 1976 donors (N= 840 men, N= 1136 women, see Supplementary Table 4 for further details). Multiple samples from same donors are valuable for many longitudinal studies, for example, in serum biomarker or metabolomics studies.

DNA samples from 760 participants were subjected to a genome-wide SNP array screen. To demonstrate the genetic homogeneity of the study population we performed principal component analysis (Supplementary Figure 1) that demonstrated that all 760 participants could be classified as Northern Europeans.

For each blood count test the time from sampling of a donor was registered (Figure 3), as well as the time when analysis result was ready and registered to the laboratory information system. The difference between these two time points was used as the sampling-to-analysis delay time (Figure 4). To assess if the trends shown in Figures 3 and 4 were significant a multiple regression model was fitted for each individual blood count measurement including both the donation time, sampling-to-analysis delay, an interaction term between these two variables, and sex and age as explanatory factors (Table 1). The interaction term was used to evaluate if the two time variables can be separated from each other in this cohort. If interaction term was not found to be significant then the co-efficient and p-value of the time variables was calculated from a model without the interaction term. While the co-efficient from a model without the interaction term is the unit change in an hour e.g. Hb drops on average 0.2 g/l per hour when going from 8 a.m. to 8 p.m., there is no such intuitive interpretation for co-efficients from a model with an interaction term. An effect in same direction and very similar size range has been shown previously from finger prick Hb measurements [26].

For ferritin, MCV, MCHC, RDW and sTfR there is significant interaction of donation time (h) and sampling-to-analysis delay time (h) and hence their individual effects cannot be separated. For hematocrit and platelet count there is a separable significant sampling-to-analysis delay time (h) effect and for erythrocyte count, Hb, hematocrit, MCH and leukocyte count there is a separable significant donation time (h) effect.

A single elevated CRP measurement i.e. a value above 10 that points to inflammation was found in 1.8 % of women and 0.3 % of men. Respectively, 14.3 % and 7.6 % had a single measurement between 3 and 10 i.e. low grade inflammation. To investigate if these measurements would point to chronic inflammation we investigated the data graphically (Supplementary Figure 2). No donor was found to have two CRP measurements above 10, hence, they most likely had no chronic inflammation.

Discussion

The present cohort was collected primarily for studies on effects of regular blood donation on donors' health, in particular iron stores and factors regulating them. The cohort complements those collected elsewhere [10-12]. We put special emphasis on careful collection and follow-up of plasma and serum samples to understand possible confounding factors in measurements of iron markers. Modelling blood count and related data is not simple as HCT, MCHC, MCV, RDW and CRP show properties of counts instead of true continuous measurements, hence, defying the assumptions of standard linear regression. To mitigate this we experimented with various transformations, link functions and error distributions in the modelling process to no avail. The results (Figures 3 and 4, Table 1) clearly demonstrate that it is essential to know the detailed conditions of collected samples if accurate results are assumed. There are interesting studies demonstrating large effects related to the time of drawing the samples [27] and this variation also is relevant in everyday blood donation practice [26]. However, for many screening purposes the variation in measurements found in the present study is apparently not critical.

Another factor that is essential to know in cohort studies is how well the study population represents the overall target population, here blood donors. Based on the demographics, Hb distribution and donation activity the participants (Figures 1 and 2) of the Fin Donor 10 000 cohort can be regarded in general to represent Finnish donor pool but there was some bias toward frequent and committed donors. As a key question to be clarified with the cohort is related to effects of frequent donation, this bias may not have serious drawbacks in our future studies.

The prevalence of CRP between 3 and 10 was very similar to what has been reported in the Danish blood donor population (14.4% and 6.1% [28]). The prevalence of CRP > 10 was found to be below 2%. Approximately 7% of the apparently healthy Finnish adults have earlier been reported to have the CRP over 10 [29,30]. As a minor CRP elevation (between 3 and 10) does not directly point to inflammation but rather to various mild tissue stress or injury [31] the Fin Donor 10 000 cohort appears as very inflammation free. This is in accordance to the healthy donor effect [32] i.e. healthier individuals get selected as donors and are able to maintain the habit from years to decades

Data from the cohort will be utilized to analyze the effects of regular blood donation on donors' health. The cohort is also of great value as a population control.

Conclusions

Together with similar cohorts from other populations [12–14] the FinDonor cohort provides us with tools to further identify the potential risk groups and genetic and non-genetic factors behind

their tendency to iron deficiency in blood donation as well as to study other health effects of blood donation. These background facts are needed for ensuring safer and more personalized blood donation in the future.

Acknowledgements

We want to thank the nurses and other staff of the Kivihaka, Sanomatalo and Espoo donation sites for recruiting and for blood donors for participation in the study.

Funding and resources

This study was partially supported by the Finnish Funding Agency for Technology and Innovation (Tekes) to the Salwe GID (Personalized Diagnostics and Care) program (ID 3982/31/2013) and by the VTR funding from the Finnish Government.

Author contributions:

Study concept: JP and JC;

Management and planning of recruitment and sample processing: PM, PN, EP and NN;

Data processing and analysis: AL and MA;

Interpretation of results: PN, EP, JC, MA, JP;

Manuscript preparation: JP and MA.

Disclosure of conflicts of interest:

Nothing to disclose

References

1. Cable, R. G. *et al.* Iron deficiency in blood donors: The REDS-II Donor Iron Status Evaluation (RISE) study. *Transfusion* 2012; **52,** 702–711.

2. Rigas, A. S. *et al.* Predictors of iron levels in 14,737 Danish blood donors: results from the Danish Blood Donor Study. *Transfusion* 2014; **54,** 789–796.

3. Waldvogel-Abramovski, S. *et al.* Iron and transfusion medicine. *Blood Rev.* 2013; **27,** 289–295.

4. Schotten, N. *et al.* The donation interval of 56 days requires extension to 180 days for whole blood donors to recover from changes in iron metabolism. *Blood* 2016; **128,** 2185–2188.

5. Davies, G. *et al.* Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. *Nat. Commun.* 2018; **9,** 2098.

6. Simon, T. L., Garry, P. J. & Hooper, E. M. Iron stores in blood donors. *JAMA* 1981; **245,** 2038–43.

7. Mast, A. E. Putting donor health first in strategies to mitigate donor iron deficiency. *Transfusion* 2017; **57,** 495–498.

8. Cook, R. L. *et al.* Iron deficiency anemia, not iron deficiency, is associated with reduced attention in healthy young women. *Nutrients* 2017; **9,** 1–13.

9. Niittymäki, P. *et al.* Retrospective analysis of capillary hemoglobin recovery in nearly 1 200 000 blood donor returns. *Blood Adv.* 2017; **1,** 961–967.

10. Tanaka, T. *et al.* A genome-wide association analysis of serum iron concentrations. *Blood* 2010; **115,** 94–96.

11. Chambers, J. C. *et al.* Genome-wide association study identifies variants in TMPRSS6 associated with hemoglobin levels. *Nat. Genet.* 2009; **41,** 1170–1172.

12. Di Angelantonio, E. *et al.* Efficiency and safety of varying the frequency of whole blood donation (INTERVAL): a randomised trial of 45 000 donors. *Lancet* 2017; **390,** 2360–2371.

13. Pedersen, O. B. *et al.* The Danish Blood Donor Study: a large, prospective cohort and biobank for medical research. *Vox Sang.* 2012; **102,** 271–271.

14. Zoglmeier, C., Martin, S. & Weinauer, F. The Bavarian Red Cross Blood Donor BioBank: the first successful combination of blood donation and biobanking for medical research. *Transfusion* 2011; **51,** 1121–1122.

15. Raivola, V., Snell, K., Pastila, S., Helén, I. & Partanen, J. Blood donors' preferences for blood donation for biomedical research. *Transfusion* 2018; **58,** 1640–1646.

16. Mitchell, R. Blood banks, biobanks, and the ethics of donation. *Transfusion* 2010; **50,** 1866–1869.

17. Hyvärinen, K. *et al.* Genetic polymorphism related to monocyte-macrophage function is associated with graft-versus-host disease. *Sci. Rep.* 2017; **7,** 1–10.

18. Anderson, C. A. *et al.* Data quality control in genetic case-control association studies. *Nat. Protoc.* 2010; **5,** 1564–1573.

19. Purcell, S. *et al.* PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* 2007; **81,** 559–575.

20. Howie, B. N., Donnelly, P. & Marchini, J. A Flexible and Accurate Genotype Imputation Method for the Next Generation of Genome-Wide Association Studies. *PLoS Genet.* 2009; **5,** e1000529.

21. Gibbs, R. A. *et al.* A global reference for human genetic variation. *Nature* 2015; **526,** 68–74.

22. Patterson, N., Price, A. L. & Reich, D. Population Structure and Eigenanalysis. *PLoS Genet.* 2006; **2,** e190.

23. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting Linear Mixed-Effects Models Using **lme4**. *J. Stat. Softw.* 2015; **67,** 1–48.

24. Team, R. D. C. & R Development Core Team, R. R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput.* 2016; **1,** 409.

25. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer-Verlag New York, 2016).

26. Bäckman, S. *et al.* Season and time of day affect capillary blood hemoglobin level and low hemoglobin deferral in blood donors: Analysis in a national blood bank. *Transfusion* 2016; **56,** 1287–1294.

27. Dopico, X. C. *et al.* Widespread seasonal gene expression reveals annual differences in human immunity and physiology. *Nat. Commun.* 2015; **6,** 7000.

28. Sørensen, C. J. *et al.* Combined oral contraception and obesity are strong predictors of low-grade inflammation in healthy individuals: Results from the Danish Blood Donor Study (DBDS). *PLoS One* 2014; **9,**

29. Palosuo, T., Husman, T., Koistinen, J. & Aho, K. C-reactive protein in population samples. *Acta Med. Scand.* 1986; **220,** 175–9.

30. Chudal, R. *et al.* Gestational maternal C--reactive protein and risk of bipolar disorder among young individuals in a Nationwide Birth Cohort. *J. Affect. Disord.* 2017; **208,** 41–46.

31. Kushner, I., Rzewnicki, D. & Samols, D. What does minor elevation of C-reactive protein signify? *Am. J. Med.* 2006; **119,** 17–28.

32. Atsma, F. & de Vegt, F. The healthy donor effect: a matter of selection bias and confounding. *Transfusion* 2011; **51,** 1883–1885.

TABLES

Table 1 Significance of explanatory variables for each measurement. For each variable a model was fitted with age, sex, sampling-to-analysis delay (h) and donation time (h) as explanatory variables. A model including an interaction term was first fitted ("p-value from a model with interaction"). If the interaction term was not significant, coefficient and p-value is provided for a model without interaction. P-values that are significant after Bonferroni correction of alpha (original alpha 0.05) are marked with an asterix (*).

Measurement	Number of measurements	Coefficient from a model without interaction		p-value from a model without interaction		p-value from a model with interaction		
		Sampling-to-analysis delay (h)	Donation time (h)	Sampling-to-analysis delay (h)	Donation time (h)	Sampling-to-analysis delay (h)	Donation time (h)	Sampling-to-analysis delay : Donation time
log2(CRP mg/l) - C-Reactive Protein	7906	1.000	0.999	7.2E-01	6.2E-01	3.1E-01	2.8E-01	3.2E-01
Eryt (x $10^{12/l}$) -Erythrocyte count	7909	-0.001	-0.011	7.9E-03	9.7E-11 *	1.4E-01	8.9E-01	6.8E-02
log2(Ferrit µg/l) - Ferritin	7902					2.6E-04 *	2.1E-04 *	1.1E-03 *
Hb (g/l) - Haemoglobin	7910	-0.031	-0.178	2.7E-02	7.6E-05 *	2.9E-01	7.8E-01	1.9E-01
HKR (%) - Hematocrit	7919	0.028	-0.083	2.1E-14 *	8.7E-12 *	2.2E-01	1.3E-04 *	3.0E-02
Leuk (x 10^9/l) - Leukocyte count	7909	0.004	0.108	9.3E-02	2.6E-41 *	3.1E-02	1.1E-07 *	4.9E-02
MCH (pg / cell) - Mean corpuscular hemoglobin	7918	0.003	0.028	2.6E-01	1.4E-03 *	3.4E-01	8.0E-01	2.7E-01
MCHC (g/l) - Mean corpuscular hemoglobin concentration	7910					8.4E-06 *	6.9E-15 *	2.8E-12 *
MCV (fl) - Mean corpuscular volume	7914					3.0E-04 *	4.1E-06 *	2.0E-07 *
RDW % - Red cell Distribution Width	7920					7.8E-03	1.3E-04 *	1.5E-03 *
sTfR (mg/l) - soluble Transferrin Receptor	7646					2.1E-04 *	5.8E-05 *	4.8E-05 *
Trom (x 109/l) - Thrombocyte count	7911	0.240	0.430	1.5E-03 *	8.0E-02	9.7E-01	9.8E-01	6.7E-01

FIGURES

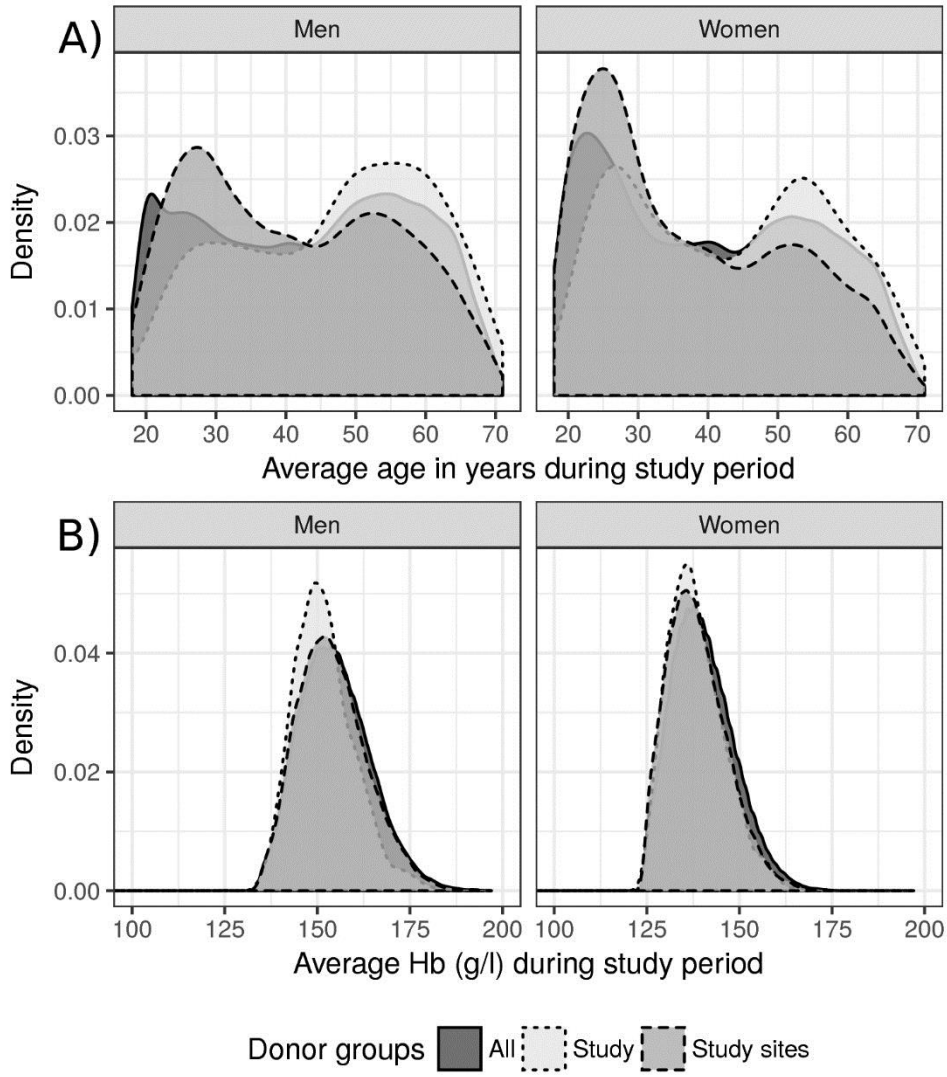

Figure 1: Density function of A) age distributions of all donors ("All"), all donors that donated in the collection sites where study took place ("Study sites") and study participants ("Study") during the study period. Respectively the density distribution of B) hemoglobin. Average over all donation events during the study period is used as the value of age and hemoglobin. Mean of hemoglobin was "All": 140 women and 154 men; "Study sites" 138 and 154 and "Study": 138 and 151.

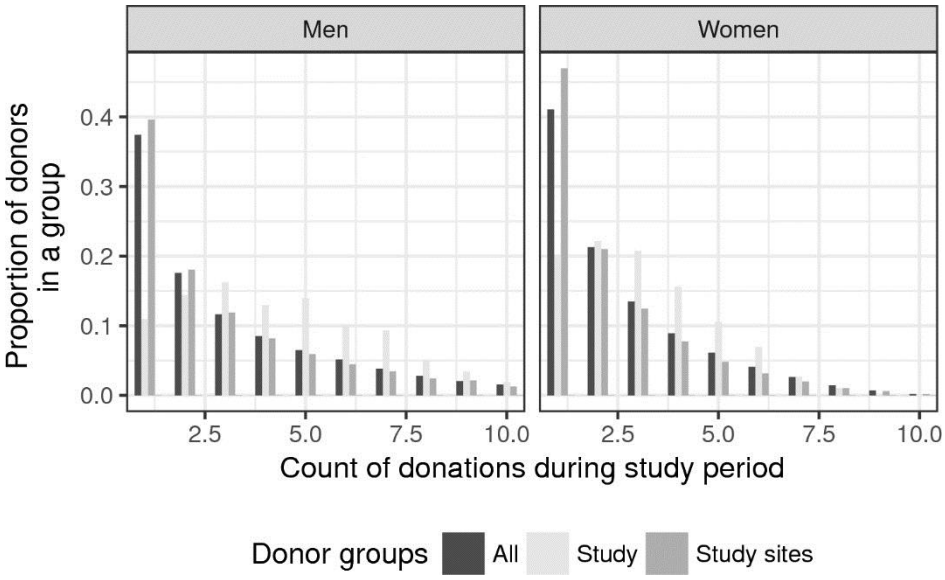

Figure 2: Counts of donations of all donors ("All"), all donors that donated in the collection sites where study took place ("Study sites") and study participants ("Study") during the study period during the study period .

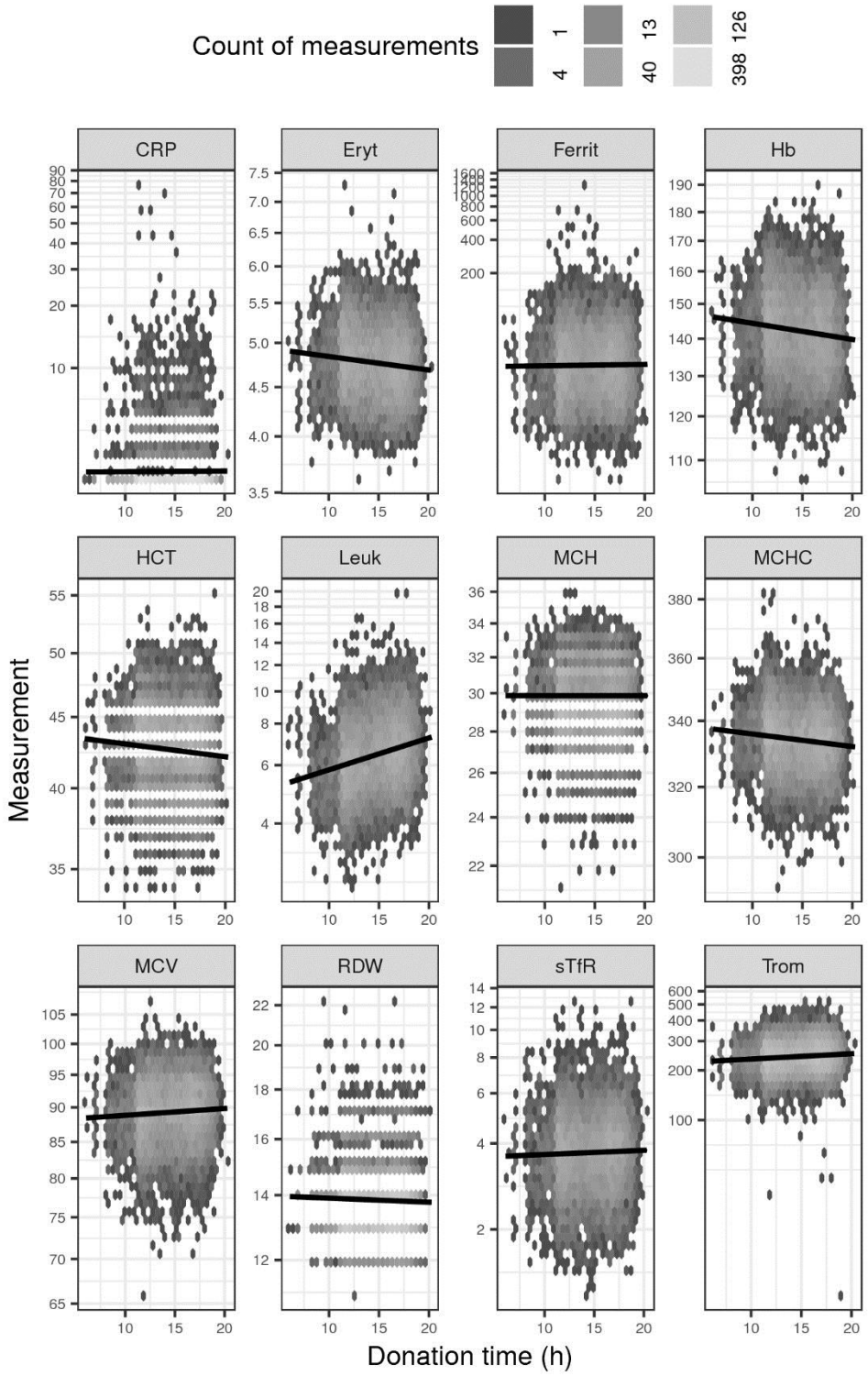

Figure 3: Effect of donation time to measurement values. Data is binned to hexagons and the color of each hexagon shows how many individual measurements are contained in each hexagon. The black line shows a linear regression trend line. See Table 1 for full description of measured variables.

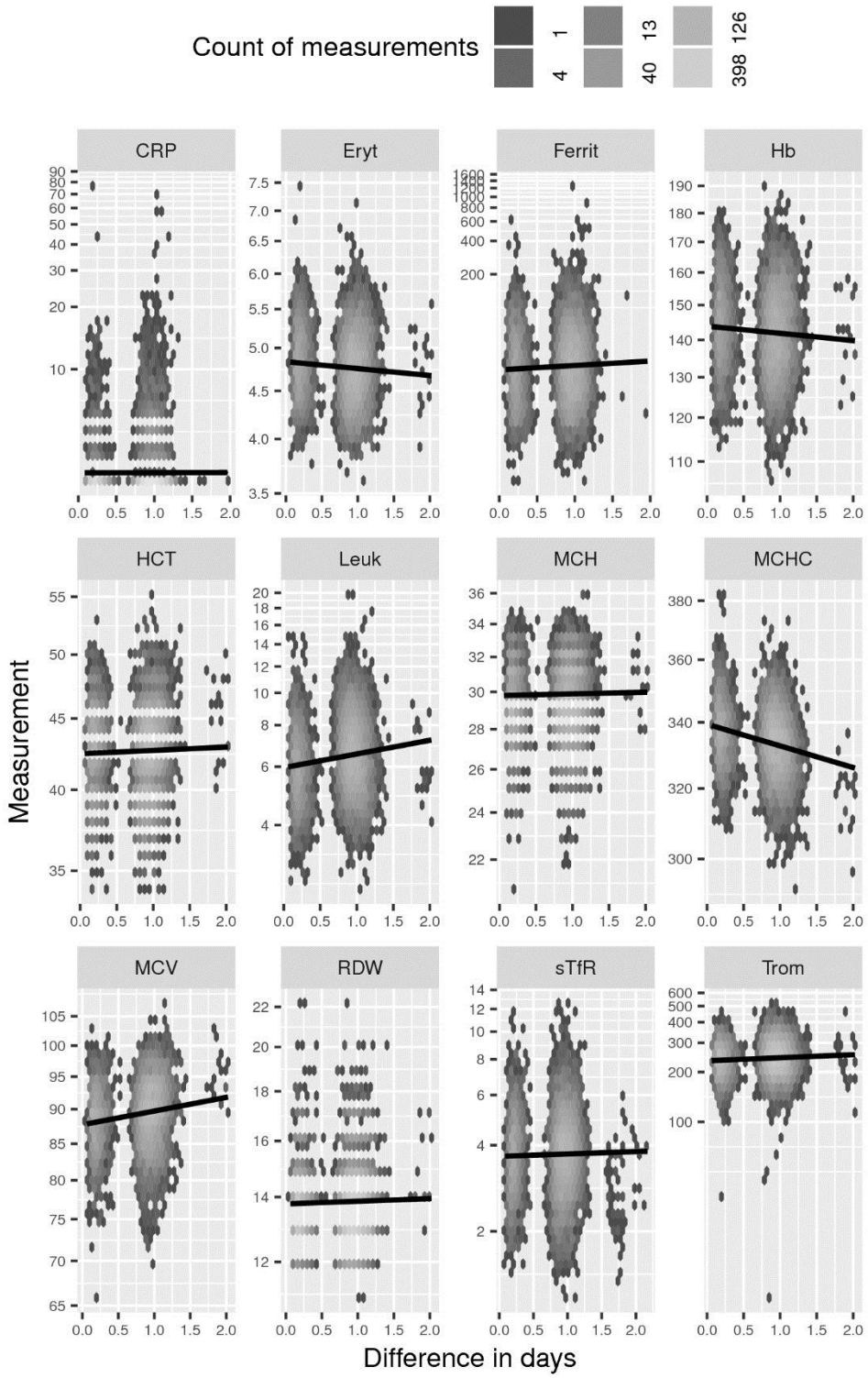

Figure 4 Effect of sampling-to-analysis delay time (d) to measurement values. Data is binned to hexagons and the color of each hexagon shows how many individual measurements are contained in each hexagon. The black line shows a linear regression trend line.