Bleeding in hemato-oncology patients: beyond the platelet paradigm

Cornelissen, L.L.


5

Chapter 5

Expected individual benefit of prophylactic platelet transfusions in hemato-oncology patients based on bleeding risks

Transfusion. 2021;61:2578–87. https://doi.org/10.1111/trf.16587

Loes L. Cornelissen[1,2,3] | Camila Caram-Deelder[1,3] | Susanna F. Fustolo-Gunnink[1,3,4]
Rolf H.H. Groenwold[3] | Simon J. Stanworth[5,6,7] | Jaap Jan Zwaginga[1,2*]
Johanna G. van der Bom[1,3*]

[1] Jon J van Rood Center for Clinical Transfusion Research, Sanquin/LUMC, Leiden, The Netherlands
[2] Department of Hematology, Leiden University medical Center, Leiden, The Netherlands
[3] Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands
[4] Department of Pediatric Hematology, Emma Children's Hospital, Amsterdam University Medical Center (UMC), University of Amsterdam, The Netherlands
[5] Transfusion Medicine, NHS Blood and Transplant (NHSBT), Oxford, United Kingdom
[6] Department of Haematology, Oxford University Hospitals NHS Foundation Trust, Oxford, United Kingdom
[7] Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom; NIHR Oxford Biomedical Research Centre, Oxford, United Kingdom

* Shared last author

Abstract

Background

Prophylactic platelet transfusions prevent bleeding in hemato-oncology patients, but it is unclear how any benefit varies between patients. Our aim was to assess if patients with different baseline risks for bleeding benefit differently from a prophylactic platelet transfusion strategy.

Study design / methods

Using data from the randomized controlled TOPPS trial (Trial of Platelet Prophylaxis), we developed a prediction model for World Health Organization grade 2, 3 and 4 bleeding risk (defined as at least one bleeding episode in 30 days) and grouped patients in four risk-quartiles based on this predicted baseline risk. Predictors in the model were baseline platelet count, age, diagnosis, disease modifying treatment, disease status, previous stem cell transplantation and the randomization arm.

Results

The model had a c-statistic of 0.58 (95% Confidence Interval (CI) 0.54 to 0.64). There was little variation in predicted risks (quartiles 46%, 47%, and 51%), but prophylactic platelet transfusions gave a risk reduction in all risk quartiles. The absolute risk difference (ARD) was 3.4% (CI -12.2 to 18.9) in the lowest risk quartile (quartile 1), 7.4% (95% CI -8.4 to 23.3) in quartile 2, 6.8% (95% CI -9.1 to 22.9) in quartile 3 and 12.8% (CI -3.1 to 28.7) in the highest risk quartile (quartile 4).

Conclusion

In our study, generally accepted bleeding risk predictors had limited predictive power (expressed by the low c-statistic), and, given the wide confidence intervals of predicted ARD, could not aid in identifying subgroups of patients who might benefit more (or less) from prophylactic platelet transfusion.

Introduction

Patients with hematological malignancies often develop thrombocytopenia as a direct consequence of their disease and/or treatment regime. Thrombocytopenia is weakly associated with bleeding, varying from skin bleeds to major bleeding in organs, among others cerebral hemorrhage.[1]

Current guidelines recommend to administer prophylactic platelet transfusions to patients with hemato-oncological disorders at a platelet count threshold of <10x10^9/L to prevent bleeding.[2-6] Guidelines also recommend to consider giving prophylactic transfusions at higher platelet count thresholds if patients have an expected higher bleeding risk, or to withhold prophylactic transfusions if the bleeding risk is relatively low, for example in autologous stem cell recipients.[2,3] The quantification of bleeding risks, however, is not standardized, resulting in considerable variation in transfusion strategies in clinical practice.[2-7]

In the randomized controlled TOPPS trial (Trial of Platelet Prophylaxis), it was found that prophylactic platelet transfusions reduce bleedings with a World Health Organization (WHO) bleeding grade of 2, 3 or 4, compared to no-prophylactic platelet transfusions (i.e. therapeutic).[8,9] What is more, this beneficial effect differed between subgroups of patients with the least effect for patients receiving autologous stem cell transplantation (SCT).[10] Other clinical variables, like fever and sex, also seemed to influence the effect of prophylactic platelet transfusion on bleeding in this trial.[11]

Overall, there remains limited quantitative evidence on how prophylactic platelet transfusions reduce the bleeding risk differently in patients with likely divers a priori bleeding risks. Where trial results give a quantification of the effect of transfusion strategies for the 'average' patient in the trial population, in practice this 'average' patient does not exist. Average effects from a trial do not necessarily apply to individual patients, in whom the actual treatment effect may differ (heterogeneity of treatment effect).[12,13] Traditionally, heterogeneity of treatment effects is investigated by comparing subgroups of patients based on a single variable. However, combining multiple patient characteristics might enable a better personalized prediction of the effect of prophylactic platelet transfusions. For example, one can imagine that a female patient with acute leukemia who has a platelet count of 45x10^9/L before treatment receiving intensive cytoreductive chemotherapy will benefit more from a prophylactic platelet transfusion strategy than a male who receives an autologous SCT to treat lymphoma with a platelet count of 155x10^9/L at the day of admission. All these, and other, clinical factors when combined can contribute to a bleeding risk, and patients with different bleeding risks may benefit differently from platelet transfusions. To know this at the start of an intensive treatment regime, such as a SCT or chemotherapy, could potentially lead to more personalized prophylactic platelet transfusion strategies.

We therefore aimed to quantify effects of a prophylactic platelet transfusion strategy compared with a therapeutic platelet transfusion strategy on the occurrence of WHO grade 2, 3 or 4 bleeding stratified by predicted baseline bleeding risks of patients with hemato-oncological diseases.

Methods

For this study, we used the data of the TOPPS trial. The design was described previously.[8,14] In short, 600 hemato-oncological patients were randomized in a prophylactic arm receiving platelet transfusions based on a threshold of $10x10^9$/L, and a therapeutic (or no-prophylaxis) arm receiving platelet transfusions in case of active bleeding. The primary outcome was the occurrence of WHO grade 2, 3 or 4 bleedings. The dataset for the analysis consisted of all 598 patients who were also included in the analysis of the TOPPS trial, of whom 47% (279 patients) developed at least one WHO grade 2, 3 or 4 bleeding during 30-day follow-up. Since we used previously collected data of one of the largest datasets for this subject, and larger trials are not likely performed in the future, no formal sample size calculation was performed for this post-hoc analysis of RCT data.

Predictors of bleeding risk

We developed a model to predict the risk of WHO grade 2, 3 or 4 bleeding within 30 days after randomization. To improve the stability of this model, we selected a limited number of baseline characteristics for inclusion in the model. The selection was made based on i) prior research that showed associations between the variables and the outcome, ii) the completeness of the data, iii) expert opinion. Selected variables were age at randomization, platelet count on day of randomization, sex, diagnosis (acute leukemia versus other), disease modifying treatment (chemotherapy/allogeneic SCT versus autologous SCT), disease status (new diagnosis versus relapsed disease), the presence of a SCT in medical history, and the randomization arm.[8,11,15-18] The randomization arm was added because ignoring treatments that affect the outcome in the prediction model can lead to an inaccurate predicted probability.[15,19] Thus, adding the randomization arm improves the prediction of the treatment effect in a heterogeneity of treatment effect analysis. Although proof of interactions cannot be obtained with the present sample size, based on clinical reasoning, interaction terms were included for the likely deemed interactions between prophylactic platelet transfusions and diagnosis, as well as for prophylaxis and treatment.

Missing data

Missing values were imputed. Given the low numbers of missing values (in total six subjects had one missing variable, Table 1), we imputed the modal value for missing categorical values. For the continuous variable platelet count, the subsequent value within three days of the randomization date of the same patient was used. If the value was unknown for these days, we imputed the median observed value of the other patients. To check robustness of the findings we performed sensitivity analyses in the subjects without missing values (n=592).

Development of bleeding risk prediction model

We developed a logistic regression model to predict the risk of WHO grade 2, 3 or 4 bleeding within 30 days after randomization.[20] To correct for optimism, we performed 'shrinkage' of all regression coefficients using penalized Ridge regression. The goal of this is to attempt to create a model that is better applicable to external datasets. Shrinkage in this respect diminishes the effect of all variables, which are likely overoptimistic in the original dataset.[21] The linearity assumption was visually checked for continuous variables; no quadratic terms or splines were deemed necessary.

After development of the model, we calculated the individual predicted 30-day risk of bleeding. For this step, to calculate the risk in absence of prophylactic transfusions for the complete population, we assumed a therapeutic platelet transfusion strategy for all patients, irrespective of their actual treatment allocation. This was necessary to be able to compare the risk with and without prophylaxis for the heterogeneity of treatment effect analysis described below, and enabled usage of the complete dataset for more power. All steps below were executed for a model without shrinkage (binary logistic regression) and for the penalized model (Ridge regression). Below, the results of the penalized model are presented; results for the crude model are presented in the supplementary material.

Model predictive performance

Performance of the model was expressed via the discriminative ability of the model (c-statistic), and as a visualization of the comparison between the predicted probability against the observed risk of bleeding (calibration plot).

Heterogeneity of treatment effect analysis

To assess the heterogeneity of treatment effects, patients were stratified in four quartiles by their predicted baseline risk. Within the quartiles, we examined heterogeneity of the effect of prophylactic versus therapeutic transfusions by estimating the odds ratio (OR) and the absolute risk difference (ARD) with 95% confidence intervals (95% CI) between the predicted number of bleedings with and without prophylactic transfu-

sions. These confidence intervals are to be considered as a measure of precision only. They were not used for formal statistical testing, given the application of Ridge penalization, and the exploratory nature of this study.

Results

The baseline characteristics of participants in the TOPPS-trial are presented in Table 1. A minority of patients was diagnosed with acute leukemia (19%) and most patients received an autologous SCT (70%). Relapsed disease occurred in approximately 1/3 of patients, and 8% had a bone marrow transplantation in the past. 65% of patients were men, the median age was 58 years and the median platelet count at day of inclusion

Table 1. Baseline characteristics of randomized patients comparing characteristics for patients based on the occurrence of WHO grade 2, 3 or 4 bleeding

	Total cohort (n=598)	No WHO grade 2, 3 or 4 bleeding (n=319)	WHO grade 2, 3 or 4 bleeding (n=279)	p-value¥
Age at inclusion (years)†	58 (49 - 63)	57 (49 - 63)	59 (51 - 64)	0.1044
Platelet count day inclusion ($x10^9$/L)†	41 (30 - 50)	41 (31 - 51)	40 (29 - 50)	0.3391
Male sex (%)	387 (65%)	223 (70%)	164 (59%)	0.005
Diagnosis (%)				0.421
Lymphoma/myeloma/other	482 (81%)	261 (82%)	221 (79%)	
Acute leukemia	116 (19%)	58 (18%)	58 (21%)	
Disease modifying treatment (%)				0.726
Autologous SCT	420 (70%)	226 (71%)	194 (70%)	
Chemotherapy/allogeneic SCT	178 (30%)	93 (29%)	85 (30%)	
Disease status (%)				0.407
New diagnosis	397 (66%)	207 (65%)	190 (68%)	
Relapsed disease	201 (34%)	112 (35%)	89 (32%)	
Stem cell transplantation in history (%)	45 (8%)	26 (8%)	19 (7%)	0.535
Randomization arm (%)				0.070
Therapeutic arm	300 (50%)	149 (47%)	151 (54%)	
Prophylactic arm	298 (50%)	170 (53%)	128 (46%)	

In total, 6 values were missing and imputed (one value per patient): platelet count was imputed for 3 patients, disease status for 1 patient and SCT in history for 2 patients.
† Median (interquartile range); ¥ p-value refers to Kruskal-Wallis equality-of-populations rank test when median is reported and Pearson's chi-squared for equality of proportions
Abbreviations: WHO= World Health Organization

Table 2. Multivariable analysis for primary outcome of WHO grade 2, 3 or 4 bleeding: odds ratios and 95% confidence intervals (CI), for both the crude model as the model after Ridge penalization.

	Crude model OR (95% CI)	Odds ratio penalized model*
Age at inclusion	1.01 (0.99 to 1.02)	1.00 (0.99 to 1.02)
Platelet count on day inclusion	1.00 (0.99 to 1.01)	1.00 (0.99 to 1.01)
Female sex (ref = male)	1.65 (1.17 to 2.33)	1.27 (0.90 to 1.80)
Diagnosis acute leukemia (ref=lymphoma/myeloma/other)	0.92 (0.37 to 2.31)	1.00 (0.40 to 2.49)
Disease modifying treatment chemotherapy or allogeneic SCT (ref=Autologous SCT)	0.74 (0.34 to 1.61)	0.96 (0.44 to 2.09)
Disease status- relapsed disease (ref= new diagnosis)	0.96 (0.66 to 1.38)	0.96 (0.67 to 1.39)
SCT in history (ref = no)	0.82 (0.42 to 1.60)	0.92 (0.47 to 1.80)
Randomization arm (ref = therapeutic)	0.93 (0.63 to 1.38)	0.81 (0.55 to 1.19)
Interaction term randomization arm and diagnosis	1.45 (0.40 to 5.20)	1.22 (0.34 to 4.40)
Interaction term randomization arm and disease modifying treatment	1.72 (0.57 to 5.19)	1.16 (0.38 to 3.49)

*Ridge penalization method, confidence intervals are only to be interpreted as an indication of precision, not as a statistical test
Abbreviations: ref = reference category, WHO = Word Health Organization, SCT=stem cell transplantation

was 41x10^9/L. Women had more bleeding events (55%, compared to 42% of men) and as reported earlier,[8] the incidence of WHO bleeding grade 2, 3 or 4 was higher in the therapeutic arm (50%) compared with the prophylactic arm (43%). Results from table 1 were not used for variable selection for our prediction model (variable selection was pre-specified), but are only descriptive.

Table 2 shows the odds ratio (OR) for WHO grade 2, 3 or 4 bleeding for all selected variables in the multivariable model, with accompanying 95% CI's. After correcting for optimism via penalization, the point estimates of most variables were approximating an OR of 1. The complete model with intercept and all regression coefficients is presented in the supplementary material, as is the crude model before penalization.

The c-statistic of the model after penalization and internal validation was 0.58 (95% CI 0.54 to 0.63), indicating that when two random patients with different bleeding outcomes are chosen, in 58% the predicted bleeding risk was lower in the patient without bleeding compared to the patient with a bleeding event.[22] The calibration plot of our model is presented in Figure 1, the slope of the plot was 2.04 (0.76 to 3.32) with an intercept of -0.06 (-0.22 to 0.10). A good calibration would have a slope approximating 1. However, due to shrinkage of the prediction model, predicted probabilities were shrunken towards the group average and consequently the model appears to be underfitted (i.e. calibration slope >1) as is expected after penalization.

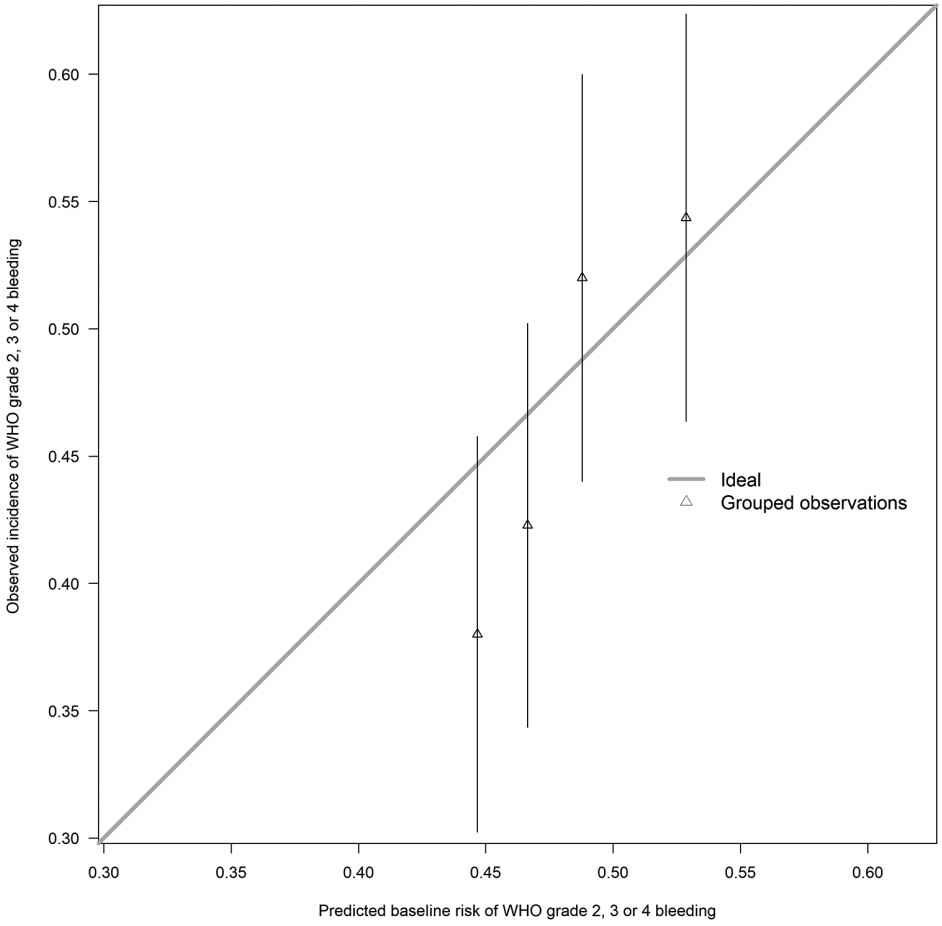

Figure 1
The triangles in this calibration plot of the predictions of WHO grade 2, 3 and 4 bleedings indicate the predicted probabilities and observed frequencies for all four risk quartiles (based on assumption of a therapeutic transfusion strategy). The diagonal line represents ideal calibration, when observed and predicted probabilities are identical. The calibration slope is 2.04 (0.76 to 3.32) with an intercept of -0.06 (-0.22 to 0.10). The c-statistic is 0.58 (0.53 to 0.62).
Abbreviations: WHO = Word Health Organization

Figure 2 shows the distribution of predicted baseline risk; all risks varied between 41% and 55%. Based on quartiles, four bleeding risk groups were defined: <46% (risk quartile 1), 46-47% (risk quartile 2), 47-51% (risk quartile 3) and >51% (risk quartile 4).

Figure 3 presents incidence rates of WHO grade 2, 3 or 4 bleedings, the OR's and risk differences when comparing the prophylactic strategy versus the therapeutic-only strategy for all patients. In all quartiles of baseline risk, the observed incidence of bleeding was higher if patients received therapeutic platelet transfusions (panel A). In panel B the

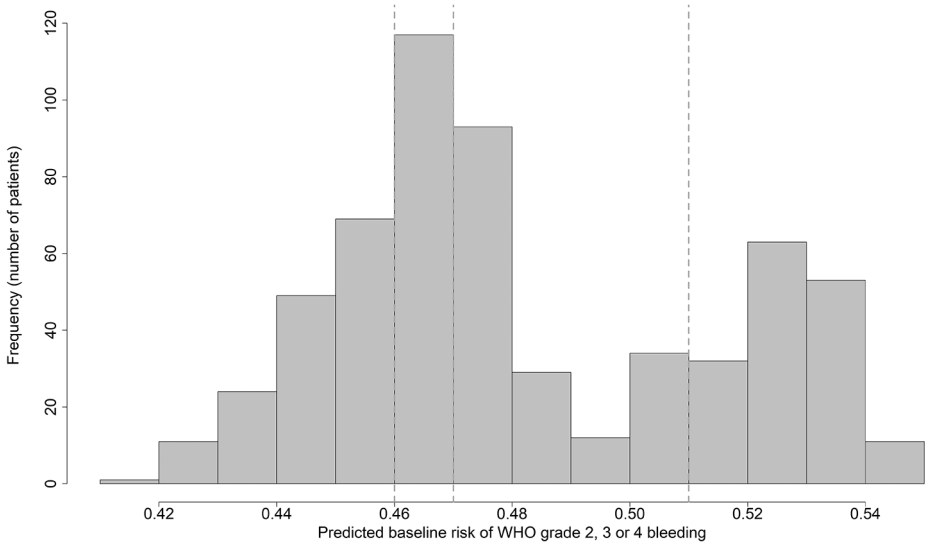

Figure 2
Predicted absolute risk of WHO grade 2, 3 or 4 bleeding (based on assumption of a therapeutic transfusion strategy) is represented as the absolute risk of outcome on the x-axis and the frequency of each absolute risk category (0.41-0.42, 0.42-0.43, etc.) in the trial population on the y-axis. The dotted lines represent the cut-off for the four quartiles of predicted risk on bleeding.
Abbreviations: WHO = Word Health Organization

OR's per quartile are displayed along with the overall odds ratio of the trial. For all quartiles, the OR is < 1, indicating a general benefit of prophylactic transfusions. The first risk quartile has an OR closer to 1, namely 0.87 (95% CI 0.45 to 1.68) compared to the overall OR (overall OR 0.74, 95% CI 0.54 to 1.03). In the fourth risk quartile the OR is more extreme compared to the overall OR, namely 0.59 (95% CI 0.31 to 1.14). The absolute risk difference (ARD, panel C) hence was most pronounced in the highest bleeding risk quartile (12.8%, 95% CI -3.1 to 28.7). This could indicate that patients in the highest risk quartile might benefit most from the prophylactic platelet transfusions, but given the wide confidence intervals this conclusion cannot be drawn on these current data. The ARDs in the other risk quartiles were 3.4% (95% CI -12.2 to 18.9), 7.4% (95% CI -8.4 to 23.3), and 6.8% (95% CI -9.1 to 22.9) respectively for risk quartiles 1, 2 and 3.

As a sensitivity analysis, we performed a complete case analysis using information about the 592 subjects with complete information. Results were comparable to those of the analysis of all 598 subjects (see supplementary material).

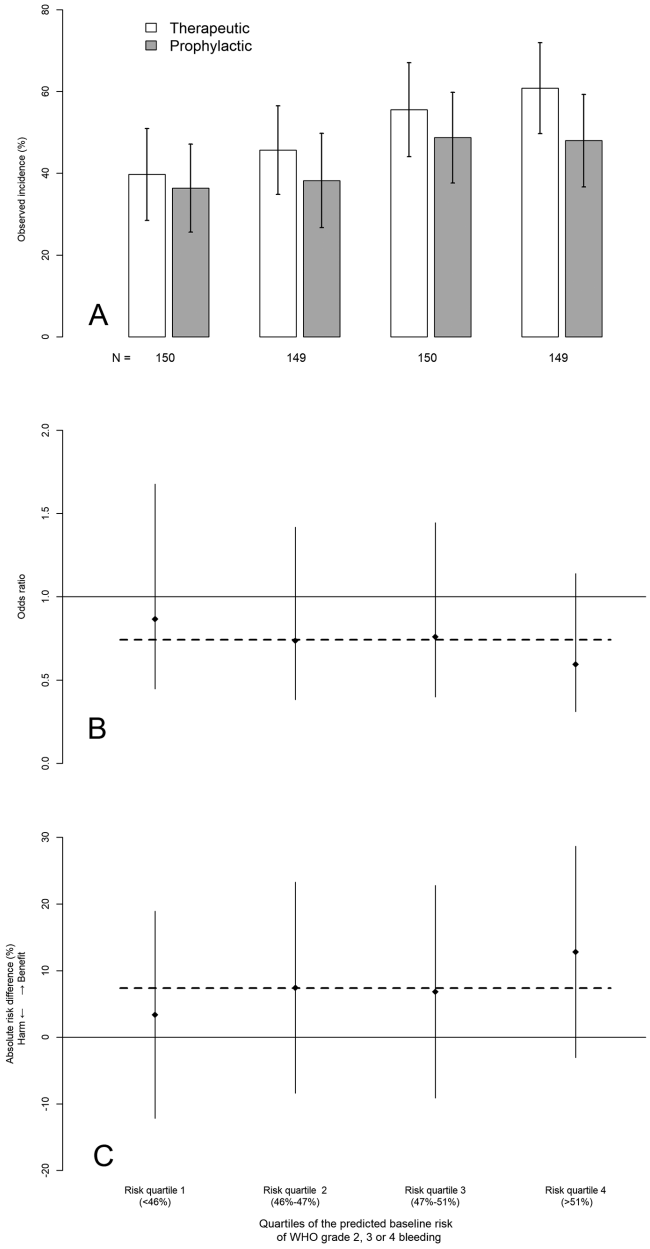

Figure 3
Event rates (panel A), odds ratios (panel B) and absolute risk differences (panel C) are presented for all four risk quartiles, comparing a prophylactically and therapeutically platelet transfusion strategy with respect to WHO grade 2, 3 or 4 bleeding. Vertical lines represent 95% confidence intervals, horizontal dotted lines represent overall trial results. A positive absolute risk reduction represents the risk decrease for a prophylactic platelet transfusion strategy as compared to a therapeutic platelet transfusion strategy.
Abbreviations: WHO = Word Health Organization

Discussion

In this post-hoc analysis of the TOPPS trial, we aimed to assess if patients with different baseline risks for WHO grade 2, 3 or 4 bleeding might benefit differently from prophylactic platelet transfusions. We found that a combination of generally accepted predictors of bleeding risk did not have much predictive power, as indicated by the low c-statistic and the small variation in risks across the risk quartiles. Although the absolute risk difference was most substantial in patients with the highest baseline bleeding risks, these differences were not statistically significant. Based on these analyses we cannot at baseline identify subgroups of patients who benefit more or less than the average effect found in the TOPPS trial.

Originally, we expected that a combination of variables could predict bleeding risk accurately, and that patients with higher bleeding risk would show a larger benefit of prophylactic transfusion. This hypothesis was based on analyses suggesting that several baseline characteristics are associated with the outcome of bleeding in hemato-oncological patients in single variable subgroup analysis.[2,10,11] From this analysis also a limited benefit for prophylactic platelet transfusions was shown for patients receiving an autologous SCT as compared to those patients receiving chemotherapy or an allogeneic SCT.[10,23,24].

However, our combined analysis of the earlier suggested baseline risk factors for bleeding in our study, was not strongly related with bleeding. Looking at e.g. platelet count, we included the baseline value since our aim was to predict bleeding at baseline. We hypothesized that a 'low platelet count at baseline' might be predictive of 'low platelet counts during admission', the latter known to be associated with bleeding risk.[8,11,23,25] More specific, both disease modifying treatments and diagnosis in the context of all other baseline risk factors, did not relevantly influence the predicted bleeding risk. Our bleeding risk prediction model therefore has a poor discriminative ability.[26] This is reflected in the low c-statistic, but also in the small range of predicted risks, namely between 41% and 55%. There are several possible explanations for this lack of predictive power reported in our analysis.

First, a potential explanation for the poor discriminative ability of the model is that our baseline characteristics contained mostly dichotomized variables. Incorporating more continuous baseline variables leads to more variation in predicted risks, but besides age and platelet count, no other continuous variables at baseline were selected beforehand to be likely predictors.

A second possible explanation may be that the sample size was not sufficient to capture the differences to actually identify the nuances in predictive values.

Thirdly, an important explanation could be that, although the included variables were shown to be associated with bleeding in isolation, bleeding is obviously influenced

by additional factors than baseline demographics alone. Instead, bleeding risk might be much better predicted by combining the baseline characteristics with characteristics that vary during treatment. Examples of such time-varying variables, which were not included in our model, are nadirs and averages of low platelet counts during admission, transfusion yields, but also a range of clinical factors such as concurrent infections and mucositis. In addition to clinical variables, biomarkers of platelet function, coagulation or endothelial function that reflect hemostasis could add to the predictive performance of the model. Such biomarkers could be baseline values as well (either inherited or acquired), or time varying during treatment. Thus, further research of time-varying variations should preferably also focus on biomarkers for hemostasis, as potential predictors for bleeding in our patients.[27-30]

However, adding such dynamic characteristics was not part of the present research question while time-dependent modelling likely needs even larger data sets than even that of the TOPPS trial. When such datasets become available in the future, the predictive performance and the clinical applicability of such time-varying bleeding prediction models, that require more frequent re-evaluation of bleeding risk compared to a baseline bleeding risk model, needs to be shown.

In our model, all patients regardless of the predicted bleeding risk benefited from the prophylactic transfusions. The absolute risk differences varied between 3.4% for quartile 1 (patients with the lowest predicted risk) and 12.8% for patients in quartile 4. Although beforehand a larger benefit in the highest risk groups was expected, with the small range of predicted bleeding risk and the wide confidence intervals and based on the included baseline characteristics solely, we cannot conclude that the benefit for patients truly differs between the risk quartiles. Our findings, despite of the limitations of our risk prediction model, can be of importance for clinicians to realize that in our study of almost 600 participants even a combination of baseline risk factors could not distinguish between subgroups with different prophylaxis effects. Of course if in future better bleeding risk discrimination becomes possible, the benefit of prophylactic platelet transfusions needs to be differentially assessed again.

There are some additional limitations that should be considered in our analysis. Firstly, in both our current as well as the original subgroup analysis of the TOPPS RCT – as in any study –, unmeasured confounding of the subgroup effect is possible, meaning that an observed subgroup effect cannot be causally attributed to the subgroup.[31] The odds ratios we present in Table 2 only serve a prediction purpose, and should not be mistaken as evidence for a real causal (in this regard a weak protective) effect of the variable on the risk of bleeding. In that regard, it is also important to clarify that in a shrunken prediction model, the wide confidence interval of the variable 'randomization arm' as presented in table 2 does not mean that the original results of the TOPPS trial should be viewed differently.

A further intrinsic limitation of our study is that our predictive model was developed and tested in the same dataset. Although necessary because qualitative good and large datasets are not easily available, this can lead to an over-optimistic model.[32,33] We tried to minimize this 'overfitting' by applying Ridge penalization. This technique shrinks the regression coefficients towards zero, which aims to result in a more reliable model when applied to other datasets. This strategy to (partly) correct the optimism of the model, comes at the cost of having predicted risks that are too close to the group average risk. Indeed, there was overfitting of the data in the original logistic regression model, and substantial shrinkage was needed. Earlier studies suggest that the more shrinkage is needed, the harder it will be to estimate the amount of shrinkage that is required.[21] What is more, Ridge regression confidence intervals do not have their usual interpretation and are solely reported to show the spreading of the results. All in all, similar as the crude model, the predictive performance of the penalized model remained poor (respectively 0.59, 95% CI 0.55 to 0.64 and 0.58, 95% CI 0.53 to 0.62). Therefore, we conclude that independent of additional penalization, baseline risk factors are suboptimal for predicting relevant bleeding.

The decision to divide patients in quartiles based on their predicted risk was made since such a risk categorization is described in literature before.[20] Looking at more than four groups, moreover, is likely increasingly impractical for clinical practice. Furthermore, more numerous categories would negatively affect the power of analyses leading to probably no additional information from such.

Strengths of this study are that this study is the first to investigate if the beneficial effect of prophylactic platelet transfusions in hemato-oncological patients differs in patients with varying baseline bleeding risks, the latter based on a combination of readily available patient characteristics. Also, a strong suit of our analysis is that instead of a subgroup analysis based on a single variable, we considered many characteristics that likely influence each other, which can lead a more accurate prediction of personalized treatment effects.[12,13] This allows for a more comprehensive evaluation of bleeding risk prediction in this population. In addition, with this technique, besides the odds ratio, we were able to estimate absolute risk differences, which is described to be of greater clinical relevance compared to a relative scale.[12] Another strength is the fact that we predefined all included variables and analysis, instead of statistical selection procedures, to avoid overfitting.[34] Lastly, a major asset of our study is that it is performed in a high quality RCT dataset. Indeed, with 598 patients the TOPPs study is one of the largest studies investigating platelet prophylaxis in this patient population.[8]

In summary, baseline risk factors have low discriminative ability to predict bleeding. With the limitations of the poor prediction of our model leading to uncertainty of our conclusions, patients in all risk groups seemed to benefit from a prophylactic platelet transfusion strategy. While patients in a higher risk group seem to benefit more, we

could not provide statistical evidence for this. Future models that incorporate dynamic (time-dependent) clinical characteristics and biomarkers of hemostasis and endothelial disruption may support better prediction of bleeding, and influence the expected individual benefit for patients with different bleeding risk in time. However, so far and based on this study, we are unable to identify patients with more or less benefit of prophylaxis. Therefore prophylactic platelet transfusions should remain a standard practice for most hemato-oncological patients who receive intensive therapy, although recognizing that many patients continue to experience bleeding events despite prophylaxis.

Acknowledgements

We would like to thank Cara Hudson (from NHSBT Clinical Trials Unit) for assisting with the data management. Also, we would like to thank all the centers that participated in the TOPPS trial.

Funding

The main trial was funded by the National Health Service Blood and Transplant Research and Development Committee and the Australian Red Cross Blood Service.

Conflict of interest

JJZ is in a scientific advisory council of Sanofi and received a speaker's fee. LLC, CCD, SFF-G, RHHG, SJS, and JGvdB have disclosed no conflicts of interest.

Ethics and approval

The original study protocol was approved by independent ethics committees in the United Kingdom and Australia. The current statistical analysis plan was approved by the science committee of the department of Clinical Epidemiology of the Leiden University Medical Center, in Leiden, the Netherlands. The main trial was registered on Controlled-Trials.com number, ISRCTN08758735.

Authors contributions

LLC designed the study, interpreted the results, and wrote the manuscript; CCD designed the study, interpreted results, analyzed the data and revised the manuscript; SFFG advised on the study design and revised the manuscript; RHHG advised on study design/statistical procedures, reviewed the statistical procedures and revised the manuscript; SS designed the TOPPS study and revised the manuscript; JJZ designed the study, interpreted results and revised the manuscript; JGB designed the study, interpreted results and revised the manuscript.

References

1 Slichter, S. J.; Kaufman, R. M.; Assmann, S. F.; McCullough, J.; Triulzi, D. J.; Strauss, R. G.; Gernsheimer, T. B.; Ness, P. M.; Brecher, M. E.; Josephson, C. D.; Konkle, B. A.; Woodson, R. D.; Ortel, T. L.; Hillyer, C. D.; Skerrett, D. L.; McCrae, K. R.; Sloan, S. R.; Uhl, L.; George, J. N.; Aquino, V. M.; Manno, C. S.; McFarland, J. G.; Hess, J. R.; Leissinger, C.; Granger, S. Dose of prophylactic platelet transfusions and prevention of hemorrhage. *N Engl J Med* **2010**, *362*, 600-613.

2 Estcourt, L. J.; Birchall, J.; Allard, S.; Bassey, S. J.; Hersey, P.; Kerr, J. P.; Mumford, A. D.; Stanworth, S. J.; Tinegate, H.; British Committee for Standards in, H. Guidelines for the use of platelet transfusions. *Br J Haematol* **2017**, *176*, 365-394.

3 Schiffer, C. A.; Bohlke, K.; Delaney, M.; Hume, H.; Magdalinski, A. J.; McCullough, J. J.; Omel, J. L.; Rainey, J. M.; Rebulla, P.; Rowley, S. D.; Troner, M. B.; Anderson, K. C. Platelet Transfusion for Patients With Cancer: American Society of Clinical Oncology Clinical Practice Guideline Update. *J Clin Oncol* **2018**, *36*, 283-299.

4 Kaufman, R. M.; Djulbegovic, B.; Gernsheimer, T.; Kleinman, S.; Tinmouth, A. T.; Capocelli, K. E.; Cipolle, M. D.; Cohn, C. S.; Fung, M. K.; Grossman, B. J.; Mintz, P. D.; O'Malley, B. A.; Sesok-Pizzini, D. A.; Shander, A.; Stack, G. E.; Webert, K. E.; Weinstein, R.; Welch, B. G.; Whitman, G. J.; Wong, E. C.; Tobian, A. A.; Aabb. Platelet transfusion: a clinical practice guideline from the AABB. *Ann Intern Med* **2015**, *162*, 205-213.

5 Padhi, S.; Kemmis-Betty, S.; Rajesh, S.; Hill, J.; Murphy, M. F.; Guideline Development, G. Blood transfusion: summary of NICE guidance. *BMJ* **2015**, *351*, h5832.

6 Bloedtransfusiebeleid. (accessed 26-05-2020.

7 Kreuger, A. L.; Middelburg, R. A.; Zwaginga, J. J.; Bom, J. G.; Kerkhoffs, J. L. H. Clinical practice of platelet transfusions in haemato-oncology. *Vox Sanguinis* **2015**, *109*, 91-94.

8 Stanworth, S. J.; Estcourt, L. J.; Powter, G.; Kahan, B. C.; Dyer, C.; Choo, L.; Bakrania, L.; Llewelyn, C.; Littlewood, T.; Soutar, R.; Norfolk, D.; Copplestone, A.; Smith, N.; Kerr, P.; Jones, G.; Raj, K.; Westerman, D. A.; Szer, J.; Jackson, N.; Bardy, P. G.; Plews, D.; Lyons, S.; Bielby, L.; Wood, E. M.; Murphy, M. F.; Investigators, T. A no-prophylaxis platelet-transfusion strategy for hematologic cancers. *N Engl J Med* **2013**, *368*, 1771-1780.

9 Miller, A. B.; Hoogstraten, B.; Staquet, M.; Winkler, A. Reporting results of cancer treatment. *Cancer* **1981**, *47*, 207-214.

10 Stanworth, S. J.; Estcourt, L. J.; Llewelyn, C. A.; Murphy, M. F.; Wood, E. M.; Investigators, T. S. Impact of prophylactic platelet transfusions on bleeding events in patients with hematologic malignancies: a subgroup analysis of a randomized trial. *Transfusion* **2014**, *54*, 2385-2393.

11 Stanworth, S. J.; Hudson, C. L.; Estcourt, L. J.; Johnson, R. J.; Wood, E. M.; Invest, T. S. Risk of bleeding and use of platelet transfusions in patients with hematologic malignancies: recurrent event analysis. *Haematologica* **2015**, *100*, 740-747.

12 Dahabreh, I. J.; Hayward, R.; Kent, D. M. Using group data to treat individuals: understanding heterogeneous treatment effects in the age of precision medicine and patient-centred evidence. *Int J Epidemiol* **2016**, *45*, 2184-2193.

13 Kent, D. M.; Steyerberg, E.; van Klaveren, D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *BMJ* **2018**, *363*, k4245.

14 Stanworth, S. J.; Dyer, C.; Choo, L.; Bakrania, L.; Copplestone, A.; Llewelyn, C.; Norfolk, D.; Powter, G.; Littlewood, T.; Wood, E. M.; Murphy, M. F.; Group, T. S. Do all patients with hematologic malignancies and severe thrombocytopenia need prophylactic platelet transfusions? Background, rationale, and design of a clinical trial (trial of platelet prophylaxis) to assess the effectiveness of prophylactic platelet transfusions. *Transfus Med Rev* **2010**, *24*, 163-171.

15 Groenwold, R. H.; Moons, K. G.; Pajouheshnia, R.; Altman, D. G.; Collins, G. S.; Debray, T. P.; Reitsma, J. B.; Riley, R. D.; Peelen, L. M. Explicit inclusion of treatment in prognostic modeling was recommended in observational and randomized settings. *J Clin Epidemiol* **2016**, *78*, 90-100.

16 Kim, H.; Lee, J. H.; Choi, S. J.; Kim, W. K.; Lee, J. S.; Lee, K. H. Analysis of fatal intracranial hemorrhage in 792 acute leukemia patients. *Haematologica* **2004**, *89*, 622-624.

17 Zumberg, M. S.; del Rosario, M. L.; Nejame, C. F.; Pollock, B. H.; Garzarella, L.; Kao, K. J.; Lottenberg, R.; Wingard, J. R. A prospective randomized trial of prophylactic platelet transfusion and bleeding incidence in hematopoietic stem cell transplant recipients: 10,000/L versus 20,000/microL trigger. *Biol Blood Marrow Transplant* **2002**, *8*, 569-576.

18 Friedmann, A. M.; Sengul, H.; Lehmann, H.; Schwartz, C.; Goodman, S. Do basic laboratory tests or clinical observations predict bleeding in thrombocytopenic oncology patients? A reevaluation of prophylactic platelet transfusions. *Transfus Med Rev* **2002**, *16*, 34-45.

19 Pajouheshnia, R.; Groenwold, R. H. H.; Peelen, L. M.; Reitsma, J. B.; Moons, K. G. M. When and how to use data from randomised trials to develop or validate prognostic models. *BMJ* **2019**, *365*, l2154.

20 Fustolo-Gunnink, S. F.; Fijnvandraat, K.; van Klaveren, D.; Stanworth, S. J.; Curley, A.; Onland, W.; Steyerberg, E. W.; de Kort, E.; d'Haens, E. J.; Hulzebos, C. V.; Huisman, E. J.; de Boode, W. P.; Lopriore, E.; van der

Bom, J. G.; PlaNeT; collaborators, M. Preterm neonates benefit from low prophylactic platelet transfusion threshold despite varying risk of bleeding or death. *Blood* **2019**, *134*, 2354-2360.
21 Van Calster, B.; van Smeden, M.; De Cock, B.; Steyerberg, E. W. Regression shrinkage methods for clinical prediction models do not guarantee improved performance: Simulation study. *Stat Methods Med Res* **2020**, *29*, 3166-3178.
22 Caetano, S. J.; Sonpavde, G.; Pond, G. R. C-statistic: A brief explanation of its construction, interpretation and limitations. *Eur J Cancer* **2018**, *90*, 130-132.
23 Wandt, H.; Schaefer-Eckart, K.; Wendelin, K.; Pilz, B.; Wilhelm, M.; Thalheimer, M.; Mahlknecht, U.; Ho, A.; Schaich, M.; Kramer, M.; Kaufmann, M.; Leimer, L.; Schwerdtfeger, R.; Conradi, R.; Dolken, G.; Klenner, A.; Hanel, M.; Herbst, R.; Junghanss, C.; Ehninger, G.; Study Alliance, L. Therapeutic platelet transfusion versus routine prophylactic transfusion in patients with haematological malignancies: an open-label, multicentre, randomised study. *Lancet* **2012**, *380*, 1309-1316.
24 Kumar, A.; Mhaskar, R.; Grossman, B. J.; Kaufman, R. M.; Tobian, A. A.; Kleinman, S.; Gernsheimer, T.; Tinmouth, A. T.; Djulbegovic, B.; Panel, A. P. T. G. Platelet transfusion: a systematic review of the clinical evidence. *Transfusion* **2015**, *55*, 1116-1127; quiz 1115.
25 Cornelissen, L. L., Kreuger, A.L., Caram-Deelder, C. et al. Thrombocytopenia and the effect of platelet transfusions on the occurrence of intracranial hemorrhage in patients with acute leukemia – a nested case-control study. *Ann Hematol* **2020**.
26 Steyerberg, E. W.; Harrell, F. E., Jr.; Borsboom, G. J.; Eijkemans, M. J.; Vergouwe, Y.; Habbema, J. D. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol* **2001**, *54*, 774-781.
27 Larsen, A. M.; Leinoe, E. B.; Johansson, P. I.; Birgens, H.; Ostrowski, S. R. High syndecan-1 levels in acute myeloid leukemia are associated with bleeding, thrombocytopathy, endothelial cell damage, and leukocytosis. *Leuk Res* **2013**, *37*, 777-783.
28 Leinoe, E. B.; Hoffmann, M. H.; Kjaersgaard, E.; Nielsen, J. D.; Bergmann, O. J.; Klausen, T. W.; Johnsen, H. E. Prediction of haemorrhage in the early stage of acute myeloid leukaemia by flow cytometric analysis of platelet function. *Br J Haematol* **2005**, *128*, 526-532.
29 Batman, B.; van Bladel, E. R.; van Hamersveld, M.; Pasker-de Jong, P. C. M.; Korporaal, S. J. A.; Urbanus, R. T.; Roest, M.; Boven, L. A.; Fijnheer, R. Agonist-induced platelet reactivity correlates with bleeding in haemato-oncological patients. *Vox Sang* **2017**.
30 Vinholt, P. J.; Knudsen, G. H.; Sperling, S.; Frederiksen, H.; Nielsen, C. Platelet function tests predict bleeding in patients with acute myeloid leukemia and thrombocytopenia. *Am J Hematol* **2019**, *94*, 891–901.
31 Groenwold, R. H.; Donders, A. R.; van der Heijden, G. J.; Hoes, A. W.; Rovers, M. M. Confounding of subgroup analyses in randomized data. *Arch Intern Med* **2009**, *169*, 1532-1534.
32 van Klaveren, D.; Balan, T. A.; Steyerberg, E. W.; Kent, D. M. Models with interactions overestimated heterogeneity of treatment effects and were prone to treatment mistargeting. *J Clin Epidemiol* **2019**, *114*, 72-83.
33 Babyak, M. A. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. *Psychosom. Med.* **2004**, *66*, 411-421.
34 Steyerberg, E. W.; Vergouwe, Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *Eur Heart J* **2014**, *35*, 1925-1931.

Supplementary material

Complete models

Crude model

Logit(P (WHO bleeding grade 2, 3 or 4)) = -0.5824† + (0.0081 * age of inclusion) + (0.0031 * platelet count on day of inclusion) + (0.5022 * sex) + (-0.0794 * diagnosis) + (-0.0708 * randomisation arm‡) + (0.3703 * diagnosis * randomisation arm) + (0.5429 * disease modifying treatment * randomisation arm)

Penalized model

Logit(P (WHO bleeding grade 2, 3 or 4)) = -0.2631† + (0.0035 * age of inclusion) + (0.0013 * platelet count on day of inclusion) + (0.2398 * sex) + (-0.0025 * diagnosis) + (-0.0417 * disease modifying treatment) + (-0.0398 * disease status) + (-0.0788 * SCT in history) + (-0.2138 * randomisation arm‡) + (0.2014*diagnosis * randomisation arm) + (0.1466 * disease modifying treatment * randomisation arm)

Add the following numbers in formula

Age of inclusion:	Age in years
Platelet count on day of inclusion:	Platelet count, ...x10^9/L
Sex:	Female =1, Male =0
Diagnosis:	Acute leukemia =1, Lymphoma, Myeloma or Other =0
Disease modifying treatment:	Chemotherapy or allogeneic SCT =1, Autologous SCT =0
Disease status:	Relapsed disease =1, New diagnosis =0
SCT in history:	Yes =1, No =0
Randomization arm‡:	Prophylactic =1, Therapeutic =0

†The intercept of the models represents the risk for patients who would have the value zero for all variables in the model, even for age and platelet count. It therefore is not applicable for any individual patient but could be seen as a baseline risk to which can be altered in both directions based on the true values of the other variables.
‡ The randomization arm was added because ignoring treatments that affect the outcome in the prediction model can lead to an inaccurate predicted probability.(1) Since the original TOPPS paper found that the therapeutic transfusion arm was on average inferior to the prophylactic transfusion arm, the predicted risk of bleeding could be lower than the 'true' risk when not taking the randomization arm into account.
Abbreviations: WHO = Word Health Organization, SCT = stem cell transplantation

1. Groenwold RH, Moons KG, Pajouheshnia R, Altman DG, Collins GS, Debray TP, et al. Explicit inclusion of treatment in prognostic modeling was recommended in observational and randomized settings. J Clin Epidemiol. 2016;78:90-100.

Sensitivity Analyses

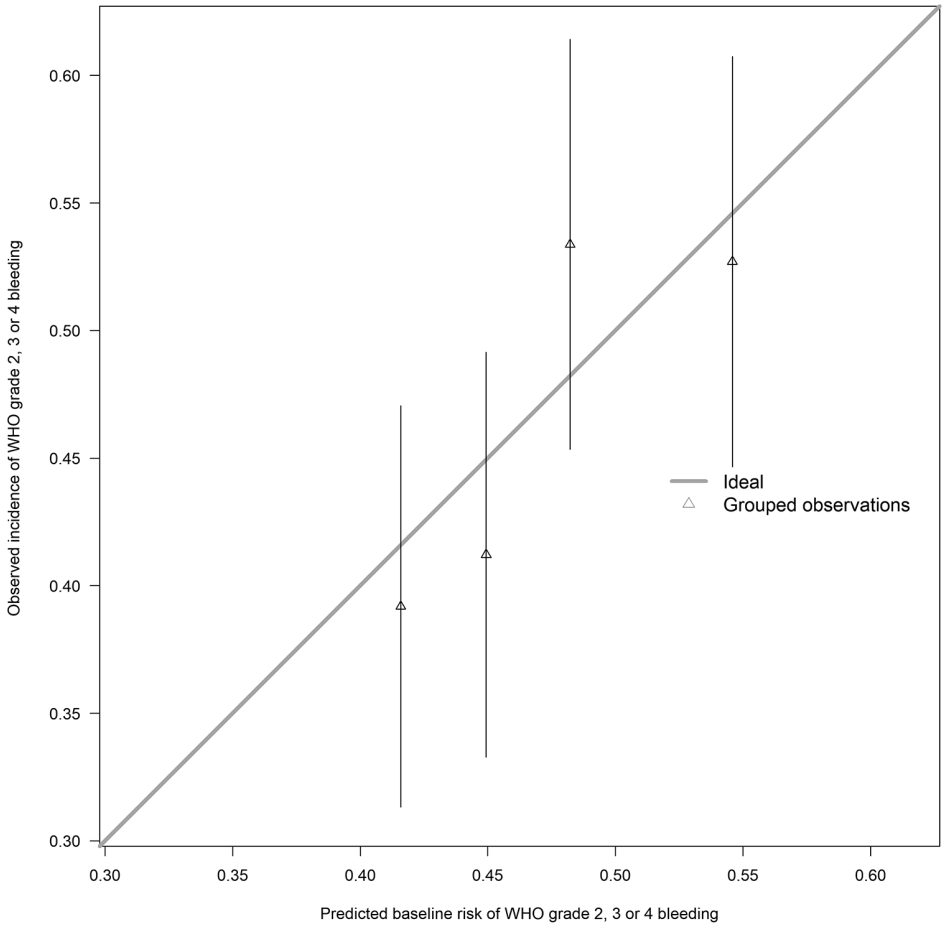

Figure S1. Calibration plot of predictions of WHO grade 2, 3 or 4 bleeding in complete case analysis (n=592)

Validity of predictions of WHO grade 2, 3 or 4 bleeding in complete case analysis. The triangles indicate the predicted probabilities and observed frequencies for all four risk quartiles (based on assumption of a therapeutic transfusion strategy). The diagonal line represents ideal calibration, when observed and predicted probabilities are identical. The calibration slope is 1.25 (0.43 to 2.06) with an intercept of -0.03 (-0.19 to 0.13). The c-statistic is 0.57 (0.53 to 0.62).

Figure S1 is comparable to figure 1, meaning that imputing baseline values when missing (n=6) did not influence our results in a relevant matter.

Abbreviations: WHO = Word Health Organization

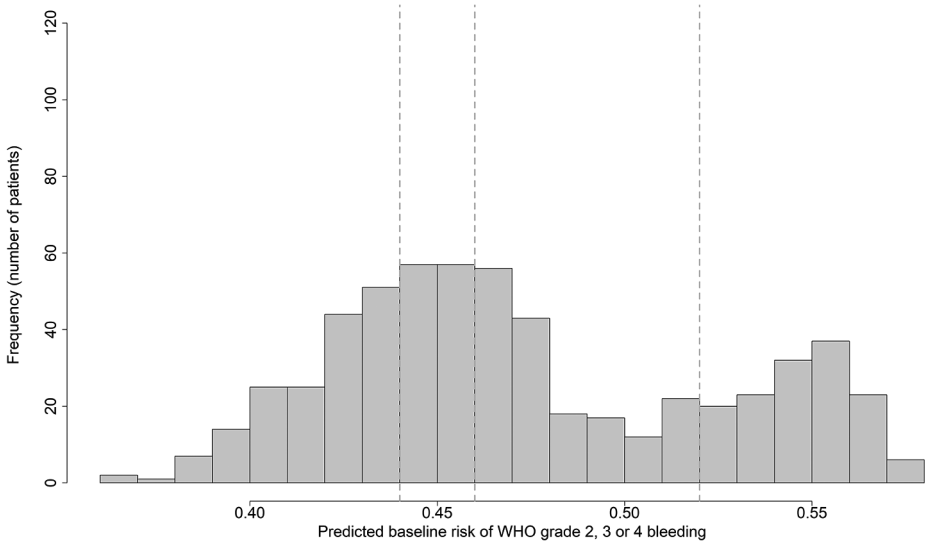

Figure S2. Histogram of predicted absolute risk of WHO grade 2, 3 or 4 bleeding in complete case analysis (n=592)

Predicted absolute risk of WHO grade 2, 3 or 4 bleeding in complete case analysis (based on assumption of a therapeutic transfusion strategy) is represented as the absolute risk of outcome on the x-axis and the frequency of each absolute risk category (0.36-0.37, 0.37-0.38 etc.) in the trial population on the y-axis. The dotted lines represent the cut-off for the four quartiles of predicted risk on bleeding.

Figure S2 is comparable to figure 2, meaning that imputing baseline values when missing (n=6) did not influence our results in a relevant matter.

Abbreviations: WHO = Word Health Organization

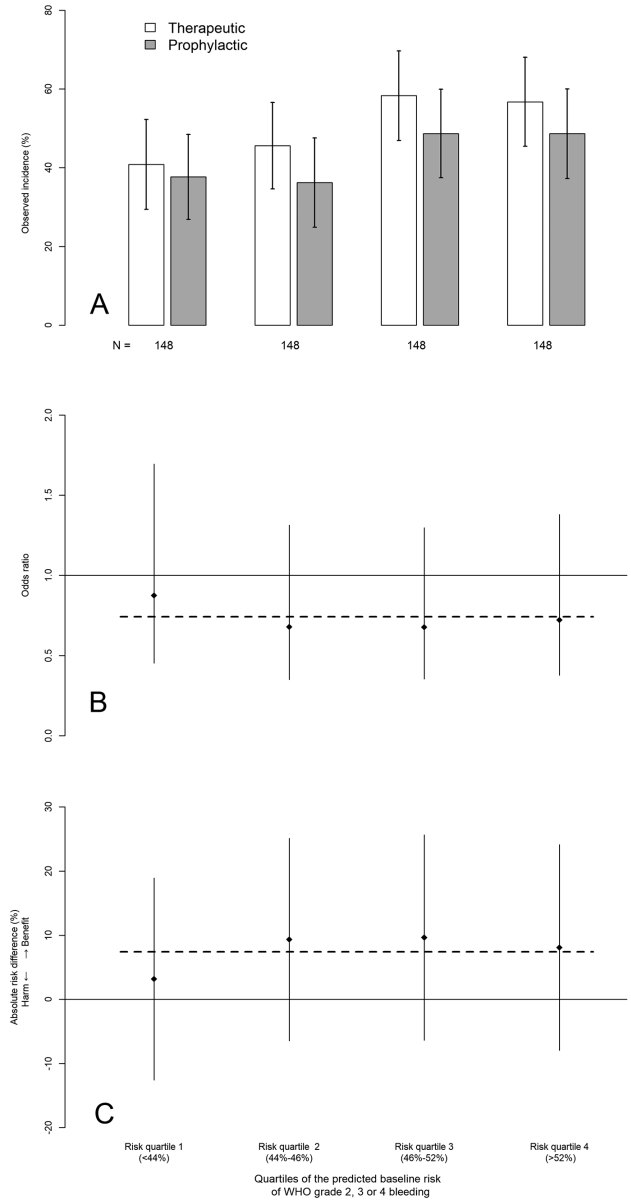

Figure S3. Observed risks, odds ratios and absolute risk differences between a prophylactically and therapeutically platelet transfusion strategy with respect to WHO grade 2, 3 or 4 bleeding in complete case analysis (n=592).

Event rates (panel A), odds ratios (panel B) and absolute risk differences (panel C) for the complete case analysis are presented for all four risk categories, vertical lines represent 95% confidence intervals, horizontal dotted lines represent overall trial results. A positive absolute risk reduction represents the risk decrease for a prophylactic platelet transfusion strategy as compared to a therapeutic platelet transfusion strategy. Figure S3 is comparable to figure 3, meaning that imputing baseline values when missing (n=6) did not influence our results in a relevant matter.

Abbreviations: WHO = Word Health Organization

Results crude model

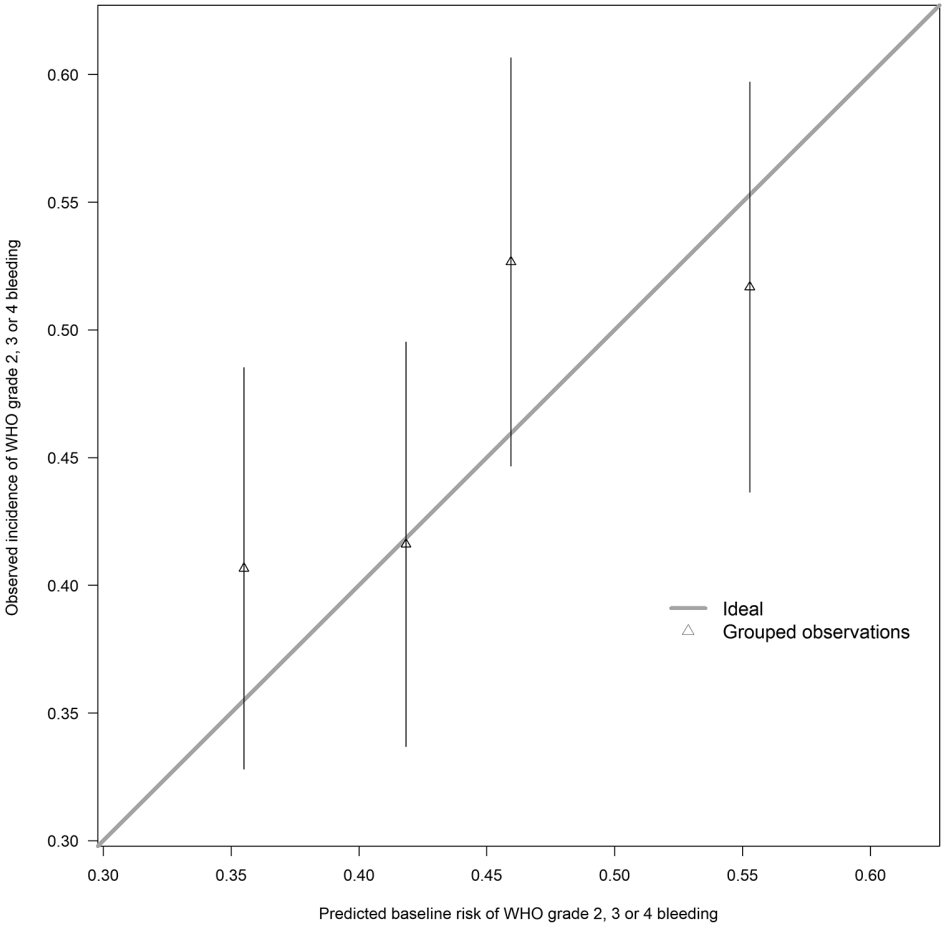

Figure S4. Calibration plot of predictions of WHO grade 2, 3 or 4 bleeding of crude prediction model
Validity of predictions of WHO grade 2, 3 or 4 bleeding when applying a crude prediction model, i.e. without correction for optimism. The triangles indicate the predicted probabilities and observed frequencies for all four risk groups (based on assumption of a therapeutic transfusion strategy). The diagonal line represents ideal calibration, when observed and predicted probabilities are identical. The calibration slope is 0.69 (0.16 to 1.22) with an intercept of 0.08 (-0.08 to 0.25). The c-statistic is 0.56 (0.52 to 0.61).
Abbreviations: WHO = Word Health Organization

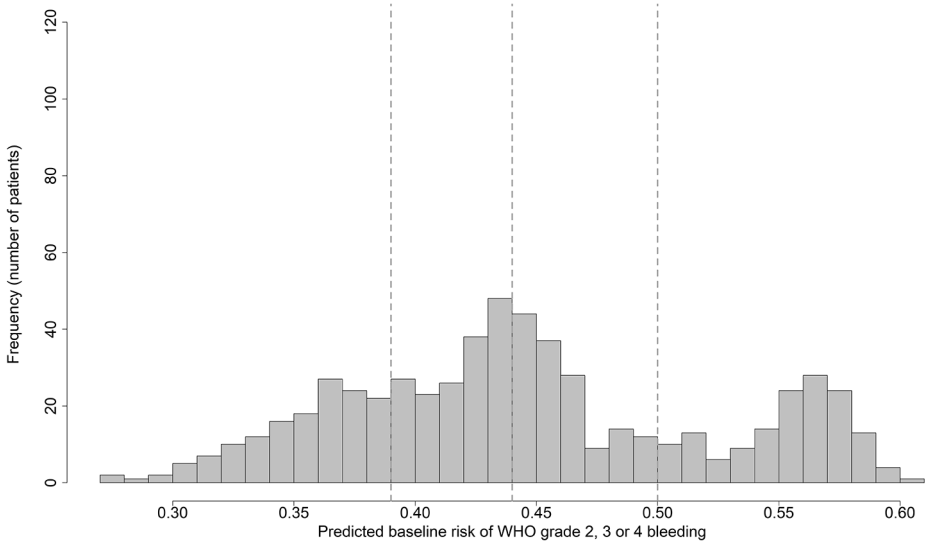

Figure S5. Histogram of predicted absolute risk of WHO grade 2, 3 or 4 bleeding of crude prediction model

Predicted absolute risk of WHO grade 2, 3 or 4 bleeding based on the crude prediction model without correction for optimism (based on assumption of a therapeutic transfusion strategy) is represented as the absolute risk of outcome on the x-axis and the frequency of each absolute risk category (0.27-0.28, 0.28-0.29 etc.) in the trial population on the y-axis. The dotted lines represent the cut-off for the four quartiles of predicted risk on bleeding.

Abbreviations: WHO = Word Health Organization

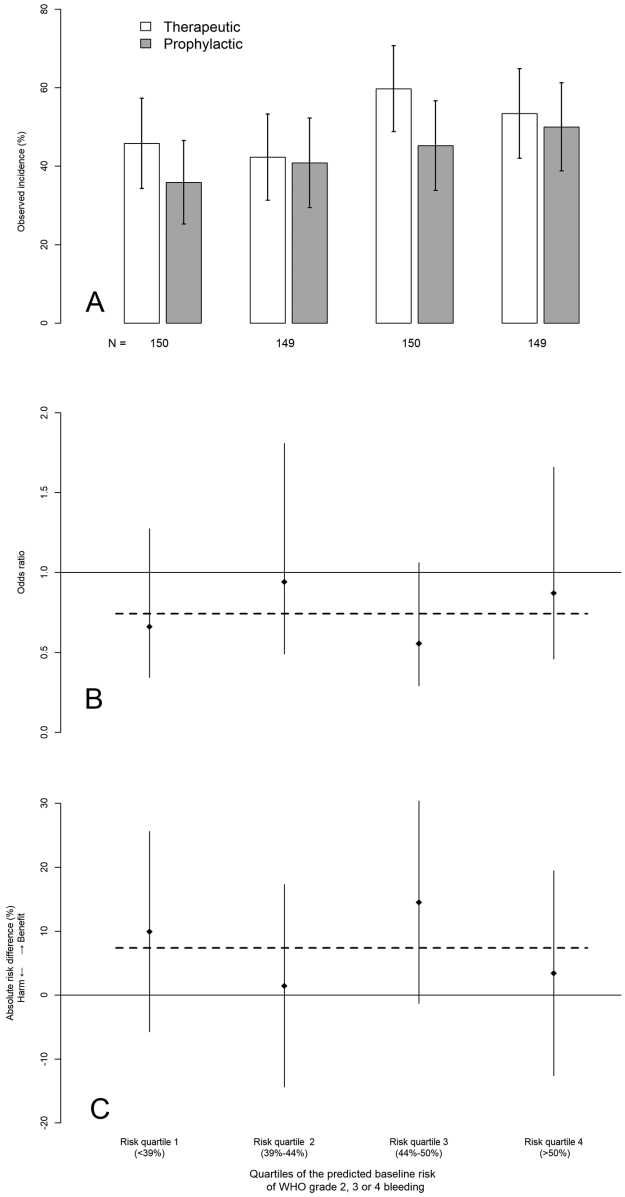

Figure S6. Observed risks, odds ratios and absolute risk differences between a prophylactically and therapeutically platelet transfusion strategy with respect to WHO grade 2, 3 or 4 bleeding for crude prediction model

Event rates (panel A), odds ratios (panel B) and absolute risk differences (panel C) based on the crude prediction model, without correction for optimism, are presented for all four risk categories, vertical lines represent 95% confidence intervals, horizontal dotted lines represent overall trial results. A positive absolute risk reduction represents the risk decrease for a prophylactic platelet transfusion strategy as compared to a therapeutic platelet transfusion strategy.

Abbreviations: WHO = Word Health Organization