Individualized prognosis in childhood immune thrombocytopenia

Schmidt, D.E.

10

BIOLOGICAL STRATIFICATION OF CLINICAL DISEASE COURSES IN CHILDHOOD IMMUNE THROMBOCYTOPENIA.

Journal of Thrombosis and Haemostasis, 2021.

Received: 31 July 2020 | Accepted: 22 December 2020

DOI: 10.1111/jth.15232

ORIGINAL ARTICLE

Biological stratification of clinical disease courses in childhood immune thrombocytopenia

David E. Schmidt[1,2] | Katja M. J. Heitink-Pollé[3] | Bart Mertens[4] | Leendert Porcelijn[5] | Rick Kapur[1,2] | C. Ellen van der Schoot[1,2] | Gestur Vidarsson[1,2] | Johanna G. van der Bom[6,7] | Marrie C. A. Bruin[3,8] | Masja de Haas[5,9]

[1]Sanquin Research, Department of Experimental Immunohematology, Amsterdam, The Netherlands

[2]Landsteiner Laboratory, Amsterdam UMC, University of Amsterdam, Amsterdam, The Netherlands

[3]Department of Pediatric Hematology, Wilhelmina Children's Hospital, University Medical Center Utrecht, Utrecht, The Netherlands

[4]Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands

[5]Department of Immunohematology Diagnostics, Sanquin Diagnostic Services, Amsterdam, The Netherlands

[6]Sanquin Research, Center for Clinical Transfusion Research, Leiden, The Netherlands

[7]Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

[8]Princess Maxima Pediatric Oncology Center, Utrecht, Netherlands

[9]Department of Immunohematology and Blood Transfusion, Leiden University Medical Center, Leiden, The Netherlands

Correspondence
Masja de Haas, Department of Immunohematology Diagnostics, Sanquin Diagnostic Services, PO Box 9190, 1006 AD Amsterdam.
Email: m.dehaas@sanquin.nl

Funding information
Studienstiftung des Deutschen Volkes; Landsteiner Foundation for Blood Transfusion Research

Abstract

Background: In childhood immune thrombocytopenia (ITP), an autoimmune bleeding disorder, there is a need for better prediction of individual disease courses and treatment outcomes.

Objective: To predict the response to intravenous immunoglobulins (IVIg) and ITP disease course using genetic and immune markers.

Methods: Children aged younger than 7 years with newly diagnosed ITP (N = 147) from the Treatment With or Without IVIG for Kids with ITP study were included, which randomized children to an IVIg or observation group. A total of 46 variables were available: clinical characteristics, targeted genotyping, lymphocyte immune phenotyping, and platelet autoantibodies.

Results: In the treatment arm, 48/80 children (60%) showed a complete response (platelets ≥100 × 10^9/L) that lasted for at least 1 month (complete sustained response [CSR]) and 32 exhibited no or a temporary response (absence of a sustained response [ASR]). For a biological risk score, five variables were selected by regularized logistic regression that predicted ASR vs CSR: (1) hemoglobin; (2) platelet count; (3) genetic polymorphisms of Fc-receptor (FcγR) IIc; (4) the presence of immunoglobulin G (IgG)

Manuscript handled by: Matthew T. Rondina

Final Decision: 22 December 2020

anti-platelet antibodies; and (5) preceding vaccination. The ASR sensitivity was 0.91 (95% confidence interval, 0.80-1.00) and specificity was 0.67 (95% confidence interval, 0.53-0.80). In the 67 patients of the observation arm, this biological score was also associated with recovery during 1 year of follow-up. The addition of the biological score to a predefined clinical score further improved the discrimination of favorable ITP disease courses.

Conclusions: The prediction of disease courses and IVIg treatment responses in ITP is improved by using both clinical and biological stratification.

KEYWORDS
immune thrombocytopenia, intravenous immunoglobulins, molecular epidemiology, pediatrics, prognosis

1 | INTRODUCTION

The extensive clinical and molecular heterogeneity of autoimmune disorders leads to complex disease classifications and challenges for the accurate clinical diagnosis, prognostication, and selection of efficacious therapies.[1] With the detailed, high-resolution molecular data that are obtained with current methods, the biological stratification of patients may become possible.[1] Using such data, the identification of clinically relevant subgroups could be achieved by the combination of several biomarkers and genetic variants, which would have limited discriminative value when considered individually. [1,2]

Childhood immune thrombocytopenia (ITP) is a rare autoimmune bleeding disorder with an annual incidence of approximately 5 per 100 000 children.[3] ITP is characterized by thrombocytopenia with a platelet count below 100×10^9/L.[4] Although two-thirds of children spontaneously recover from the disease within 3 months of the diagnosis, the remaining patients exhibit prolonged thrombocytopenia, and some eventually develop chronic disease.[5,6] The quality of life is markedly reduced in patients with prolonged and chronic ITP.[7,8] A key clinical challenge is the early identification of a patient's likely disease course (i.e., platelet count and bleeding over time), as well as the expected response to treatment. Specifically, an indication of a patient's likely disease course may help clinicians in the counselling of patients and families, as well as guide patient monitoring and treatment decisions. The early identification of patients likely to have persistent or chronic ITP would enable the clinicians to target monitoring and perform additional diagnostic testing (e.g., screening for immune deficiencies, other autoimmune diseases, genetic causes), as well as consider adjunctive or alternative forms of treatments. In childhood ITP, treatment is only initiated in emergency situations or on a case-by-case basis with a significant perceived disease burden.[9] In these situations, knowledge of the expected response to intravenous immunoglobulin (IVIg) is critical for managing bleeding and preventing unnecessary hospitalizations, side effects, and substantial costs.[7] However, there are currently no sufficient tools that can be used to predict the disease course of ITP.

Essentials

- Molecular variables may aid in the prediction of the immune thrombocytopenia (ITP) prognosis.
- We developed a biological risk score for recovery after Intravenous Immunoglobulin treatment.
- The biological score also associated with spontaneous recovery and bleeding during observation.
- The biological score improved discrimination of ITP prognosis in addition to a clinical score.

In our recent randomized controlled trial, we observed that although IVIg treatment promoted platelet recovery over the short term, it had no effect on the long-term recovery from ITP (i.e., the rate of chronic disease) at 1 year postdiagnosis.[10] It is well known that the administration of IVIg leads to a rapid resolution of thrombocytopenia in the majority of children with ITP,[4,11] yet a proportion of patients does not respond to treatment. Variable platelet response rates to IVIg of 50% to 95% have been reported in the literature, with differences between patient cohorts and response definitions.[12-16] Moreover, it has been reported that IVIg nonresponders are more likely to develop chronic ITP (i.e., have a lower likelihood of spontaneous recovery).[12,17]

The design of prediction models in childhood ITP is challenging. First, the broad clinical diagnosis of ITP may lead to a wide range of diverse causes of thrombocytopenia (e.g., autoinflammatory and post-infectious syndromes, early signs of autoimmune diseases, immune deficiencies, or, probably more rarely, genetic thrombocytopenia).[18-20] Second, there may be heterogeneity in the underlying pathophysiology of ITP (e.g., autoantibodies specific to platelet GPIb/IX,[21] or CD8-mediated cellular antiplatelet immunity[22]). Third, a further challenge in childhood ITP is the association of age at presentation with a skewing of the disease prognosis, where both adolescents and adults display a different prognosis compared with younger children.[23-25] Age is also associated with differences in cellular subsets and other molecular

markers, which needs to be considered (Schmidt et al., manuscript submitted; https://doi.org/10.1101/2020.06.09.20125385).

In the present study, our objective was to develop a biological risk score to identify at the diagnosis children with newly diagnosed ITP who show no response versus a complete recovery after IVIg. In line with the literature, given the relationship between IVIg response and the ITP disease course, we further tested if the factors associated with a favorable IVIg response were also associated with spontaneous recovery from ITP. Finally, we investigated if these biological factors could improve the discrimination of children with a favorable disease course, as stratified by a score using clinical characteristics alone.

2 | METHODS

2.1 | Study participants

This was a secondary analysis of the Treatment With or Without IVIG for Kids With ITP (TIKI) study (Heitink-Pollé et al[10]; Trial ID NTR1563, www.trialregister.nl). Patients were recruited into an open-label multicenter randomized phase 3 clinical trial. In this trial, 200 children with newly diagnosed ITP and mild bleeding symptoms were allocated to a single dose 0.8 g/kg IVIg or careful observation. The patients were eligible for inclusion if they had a platelet count ≤20 × 10^9 L^{-1} and an age of 3 months to 16 years. Patients were excluded in cases of severe or life-threatening bleeding at diagnosis, or treatment with immunosuppressive therapy in the previous 3 months. This study was conducted in accordance with the second declaration of Helsinki and was approved by the institutional ethical review board of the University Medical Centre Utrecht. Parents and patients aged 12 years or older provided written informed consent before inclusion in the study. All analyses were performed on deidentified anonymous data. All biological data used as predictors were obtained at the time of diagnosis, before IVIg was administered (details in Supplementary Material).

In this study, we only included the children aged younger than age 7 years (N = 147), which reflected 75% of all children originally included in the trial, for two reasons. First, the immune system is undergoing marked changes during child development.[26,27] Children with ITP presenting at various ages also display different clinical and biological baseline characteristics.[23,25,28] (Schmidt et al., manuscript submitted; https://doi.org/10.1101/2020.06.09.20125385) Second, the clinical outcomes in ITP changes with age (i.e., the probability of a transient vs persistent or chronic ITP disease course decreases gradually already from 5 years of age) (Schmidt et al.).

2.2 | Patient data and samples

The clinical follow-up, a standardized bleeding score (modified Buchanan score[29,30]), and complete blood counts were obtained at the recruiting centers at diagnosis and after designated intervals.

2.3 | Clinical outcome definition

A complete sustained response (CSR) after IVIg was defined according to the International Working Group criteria,[4] with a platelet count >100 × 10^9/L measured at two independent occasions, which in our study were 1 and 4 weeks after IVIg. Other patients were classified as having an absence of a sustained response (ASR). A partial response was defined as a platelet count >30 × 10^9/L with a two-fold count increase from baseline, and a complete response as a platelet count >100 × 10^9/L. A temporary response was defined as an initial complete response, which was then lost at the 1-month follow-up.

2.4 | Statistical analysis

Details are provided in the Supplementary Methods. Missing variables were imputed by multiple imputation using chained equations.[31] The age-associated changes in immune variables were first removed from the data by constructing a linear model for each numerical variable. For the biological risk score, a regularized logistic regression model was built on the patients in the IVIg cohort to predict the outcome of a complete versus absent sustained response to IVIg, with the elastic net by *glmnet*.[32] Model features were selected by the inclusion frequency of a variable. The estimates of the coefficients were pooled according to Rubin's Rules.[33] To assess the predictive performance of the biological risk score, cross-validated predictions were determined and performance was assessed using receiver operating characteristics.[34]

2.5 | Data sharing and code availability

Data may be requested for academic collaborations from the corresponding author. All code used for this study is available upon enquiry from the corresponding author.

3 | RESULTS

3.1 | Lack of a sustained response to IVIg was a common phenomenon

In this study, 147 children with newly diagnosed childhood ITP aged younger than age 7 years were included (N = 80 from the IVIg arm; N = 67 from observation; Figure 1A). The median age of the study population was 3 years, and approximately 60% had a preceding infection (Table 1). A total of 48/80 patients allocated to IVIg (60%) responded with an increased platelet count and sustained the response at the 1-month follow-up (Figure 1B; details and bleeding symptoms in Figure S1), marking them as CSRs. Importantly, all patients with CSR remained in clinical remission during the full 1-year follow-up. The other 40% exhibited either a partial or complete

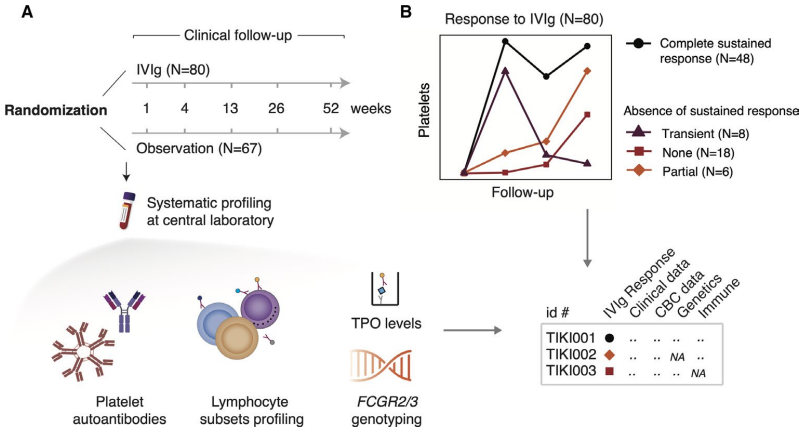


FIGURE 1 Study design. (A) After the diagnosis, the patients were randomized to a single dose of IVIg or a careful observation group and followed for 1 year. Patients with newly diagnosed ITP were systematically characterized, including the clinical data, targeted genotyping, and immune phenotyping as well as autoantibody profiling. (B) The response to IVIg correlated with platelet counts during 3 months' follow-up. Based on the platelet counts, we categorized patients as either IVIg responders and nonresponders (for details and bleeding symptoms, see Figure S2). The complete sustained responders (CSR) showed a platelet count >100 × 10^9 /L 1 week after treatment and the response was sustained 1 month later. If an initial complete response was not maintained at the 1-month follow-up, patients were classified as having a transient response. Partial responders (PR) showed an increase to >30 × 10^9 /L with a two-fold increase of baseline platelet count 1 week after IVIg. Nonresponders (NR) showed neither a complete nor a partial response. PR, NR, and transient responders were grouped together as absence of a sustained response (ASR). ITP, immune thrombocytopenia; IVIg, intravenous immunoglobulin

TABLE 1 Baseline characteristics of the study population

	IVIg	Observation
N	80	67
Female, n (%)	34 (42.5)	30 (44.8)
Age, y	3.09 [2.28, 4.87]	3.04 [1.72, 4.51]
Age, n (%)		
0–1	3 (3.8)	5 (7.5)
>1–3.5	46 (57.5)	35 (52.2)
>3.5–7	31 (38.8)	27 (40.3)
Platelet count, ×10^9 L^{-1}	6.0 [3.0, 8.3]	6.0 [3.0, 10.0]
Duration of symptoms, d	3.0 [2.0, 5.3]	3.0 [1.0, 7.0]
Preceding infection, n (%)	50 (62.5)	39 (60.0)
Preceding vaccination, n (%)	3 (3.8)	2 (3.1)
Buchanan score, n (%)		
0	1 (1.2)	0 (0.0)
1	12 (15.0)	13 (19.7)
2	32 (40.0)	27 (40.9)
3	35 (43.8)	26 (39.4)

Note: Continuous variables are median [interquartile range].
IVIg, intravenous immunoglobulin.

lack of a response to IVIg (Figure 1B), or a transient response and recurrence of thrombocytopenia. Together, these patients were classified as showing an ASR. Moreover, these platelet responses were correlated with bleeding symptoms as expected (Figure S1). Together, these data substantiate that a lack of a sustained response to IVIg is common amongst patients with newly diagnosed ITP.

3.2 | Development of a genetic and immunological score to predict the response to IVIg

To identify the predictors of a response to IVIg treatment at the time of diagnosis, we collected detailed clinical data, performed peripheral immune phenotyping and platelet antigen-specific autoantibody tests, and genotyped the candidate genes in the Fc-γ receptor locus *FCGR2/3* associated with the development of antibody-mediated autoimmune diseases, which may affect the efficacy of IVIg[35-37] (Figure 1; a full list of all the 46 considered variables is listed in Table S1).

To build a biological risk score, we used regularized logistic regression to predict the treatment responses to IVIg (i.e., CSR vs ASR). The predictive ability was evaluated by cross-validation and showed optimal discrimination with five or eight predictors (Figure 2A, arrows). In out-of-sample validation, patients with CSR and ASR who were not included during the modelling were discriminated with a receiver-operating characteristic area under the curve (AUC) of 0.70, where an AUC of 0.50 indicates no discrimination and 1.00 perfect discrimination. Given a similar cross-validated receiver-operating characteristic AUC, the simpler multivariate model with five predictors was chosen for further analysis (Table 2; the eight-predictor model is shown in

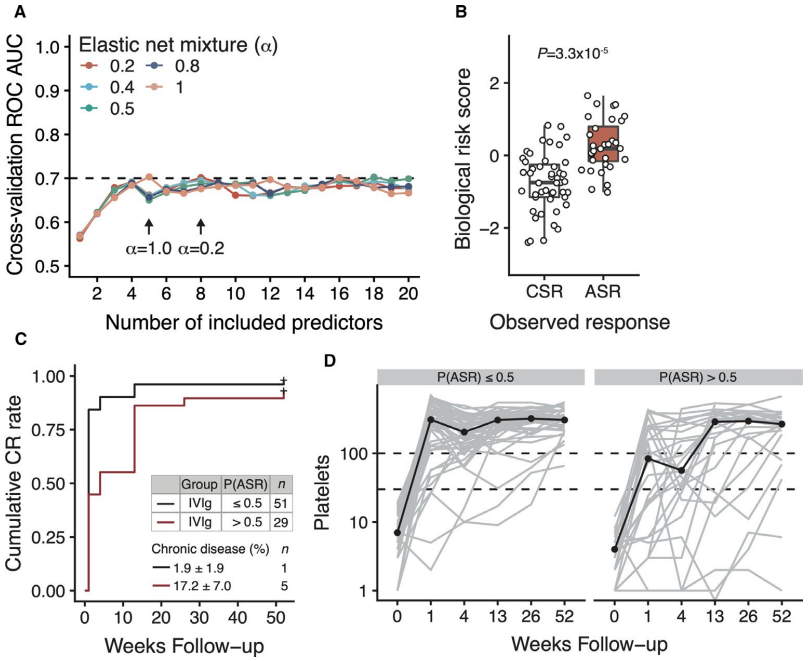


FIGURE 2 Prognosis after IVIg-treatment is predicted by the biological risk score. (A) Using cross-validation, the number of variables and corresponding elastic net mixture (alpha) values were optimized (out-of-sample). The alpha parameter controls the mixture of lasso and ridge penalties in the elastic net. Two parameter combinations yielded optimal discrimination (arrows). The detailed procedure is outlined in the Supplementary Methods. (B) The biological risk score predictions (score on log odds scale) displayed discrimination of IVIg responders (CSR; N = 48) and nonresponders (ASR; N = 32). (C) A Kaplan-Meier plot of a complete response according to the biological risk score. The score was split at 0, yielding posterior probabilities for ASR (P[ASR]) of ≤0.5 and >0.5, respectively. Statistics are provided in the main text. (D) Individual patient platelet count trajectories (gray lines) and medians (black line) show an enrichment of the long-term thrombocytopenic trajectories in patients who had P(ASR) >0.5 (N = 29) vs ≤0.5 (N = 51). Patients with a score above this cutoff showed a reduced response rate to IVIg. Moreover, a slower recovery rate was observed during the 1-year follow-up, and an enrichment of patients who develop chronic ITP. ASR, absence of a sustained response; CSR, complete sustained responder; ITP, immune thrombocytopenia; IVIg, intravenous immunoglobulin

Table S4). There was a clear separation in the predicted scores for the response between CSR and ASR (Figure 2B). The receiver-operating characteristics of the model score showed an AUC of 0.84 (fitted predictions; 95% confidence interval [CI], 0.75-0.93; Figure S2A). The different subgroups of the ASR—partial responders, patients with a transient response who exhibited recurrence of thrombocytopenia, and nonresponders—were all similarly discriminated from patients with CSR (Figure S2 B). This indicates that they were relatively homogeneous regarding the model parameters.

When the biological score was split at a posterior probability for ASR of 50% (score of 0, P[ASR] >0.5), there was a clear enrichment of patients with ASR in the group predicted to show ASR (Figure 2C). Moreover, a predicted ASR was associated with a reduced frequency of complete response over the follow-up with a hazard ratio (HR) of 0.63 (95% CI, 0.50-0.80; Figure 2C), where an HR of 1.00 would indicate no discrimination of disease courses. The same effect was apparent when we assessed the platelet count over time (Figure 2D). Patients with chronic disease at 12 months were enriched in the group predicted to have ASR at the time of diagnosis (Figure 2C,D). Finally, the rate of the patients with prolonged or chronic disease trajectories was increased from low to high scores (Figure S2 C).

We investigated the discriminative nature of the scores at various thresholds. At an optimal cutoff of −0.54, chosen by the Youden's J value, which optimizes total sensitivity and specificity, patients with ASR were detected with a high sensitivity of 0.91 (95% CI, 0.81-1.00) and a specificity of 0.67 (95% CI, 0.53-0.80) (Table S5). This yielded a positive predictive value of 0.64 (95% CI, 0.50-0.78). Importantly, most of the patients predicted to have a CSR also showed a CSR (i.e., the score had a negative predictive value for ASR of 0.91; 95% CI, 0.82-1.00). Discrimination at other score thresholds is shown in Table S6. Taken together, from the variables available shortly after the diagnosis, we identified a biological and immune profile that can discriminate patients with CSR and ASR in this cohort.

TABLE 2 The biological risk score for prediction of the IVIg treatment response

	Biological Risk Score
Odds ratio (95% CI*)	
Hemoglobin, mmol/L	0.54 (0.41-0.70)
Platelet count, $\times 10^9$/L	0.70 (0.59-0.82)
FCGR2C-ORF present	0.44 (0.32-0.59)
MAIPA IgG positive	0.47 (0.30-0.74)
Preceding vaccination	3.11 (1.87-5.15)
ROC AUC (95% CI)	0.84 (0.75-0.93)
Nagelkerke R^2	0.428
Brier Score	0.163

Note: Odds ratios (OR) are provided for the absence of sustained response (ASR; logistic regression with outcome ASR vs. complete sustained response). OR for continuous predictors are standardized coefficients (i.e., the OR for the effect of a change in one standard deviation on the scale of the predictor). All continuous variables were adjusted for age as outlined in the Supplementary Methods. The calculation of an individual's score is shown in the Supplementary Methods. *The 95% confidence intervals (CI) for the coefficients are reported to illustrate the model variability using bootstrapping, but they are biased because of the use of penalized regression.

AUC, area under the curve; IgG, immunoglobulin G; IVIg, intravenous immunoglobulin; MAIPA, monoclonal antibody immobilization of platelet antigens; ORF, open reading frame; ROC, receiver-operating characteristic

3.3 | Biological predictors and treatment responses to IVIg

We next investigated the distribution of the identified variables between patients with CSR and ASR. Aggregated individual patient data are presented in Figure 3A (additional predictors of eight-parameter model in Figure S3). Patients with ASR showed lower hemoglobin levels and a lower platelet count at the time of diagnosis (Figure 3B). Moreover, the patients with ASR displayed a lower frequency of circulating IgG antiplatelet glycoprotein antibodies and the immune-stimulatory open reading frame variant of *FCGR2C*. When a preceding vaccination was present within 28 days of the diagnosis, none of the patients showed CSR (N = 3; Figure 3B). When the variables were analyzed together, a dimensional reduction with a principal component analysis (Figure S4) displayed a good agreement between the observed IVIg responses and the predicted response.

3.4 | Predictors of the IVIg treatment responses were associated with spontaneous ITP recovery

Next, we evaluated whether the biological risk score was also associated with the disease course for the patients who were allocated to the observation group. Because of the randomization, this tested the model's relationship with disease courses independent of IVIg in

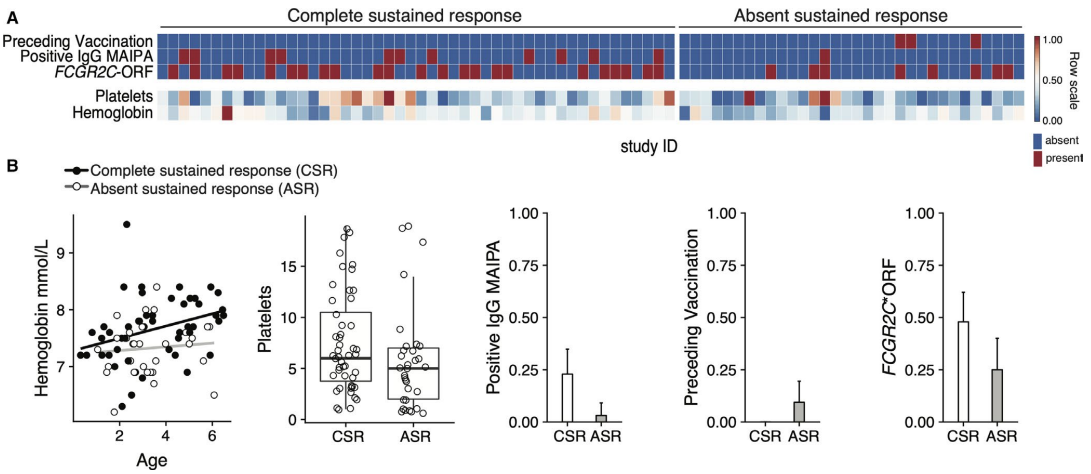


FIGURE 3 Variables associated with the IVIg response as identified by regularized regression. (A) An overview of the predictors is provided as a heatmap for each subject with complete sustained response (CSR; N = 48) and absence of a sustained response (ASR; N = 32). Age-corrected values were used as the input and scaled per row (see legend). (B) Regularized regression identified variables associated with a lack of response to IVIg (ASR), including lower hemoglobin levels (Welch *t* test; $p = 0.008$) and higher platelet counts (Wilcoxon rank-sum test; $p = 0.051$). Statistical testing was performed for the variable residual on age (see Supplementary Methods). Furthermore, patients with ASR displayed an absence of anti-platelet glycoprotein antibodies (IgG MAIPA; Fisher's exact test, $p = 0.022$), the presence of a preceding vaccination ($p = 0.060$), and presence of an open reading frame variant in *FCGR2C* ($p = 0.060$). ASR, absence of a sustained response; ITP, immune thrombocytopenia; IgG, immunoglobulin G; IVIg, intravenous immunoglobulin; MAIPA, monoclonal antibody immobilization of platelet antigens

patients similar to the ones allocated to IVIg treatment. The score discriminated favorable disease courses (Figure 4A; Table 3), and a higher score was associated with reduced complete recovery rates with an HR of 0.69 (95% CI, 0.52-0.91). In the 37 patients who were predicted to have CSR, approximately 95% recovered by the 12-month follow-up (n/N; 35/37), compared with the 9% overall chronic disease rate (6/67). Moreover, the patients predicted to have CSR displayed an enrichment of a faster recovery of platelet counts over time (Figure 4B). For both the observation cohort and the IVIg-treated cohort, the model score differentiated bleeding severity during the full 1-year follow-up (Table S7). These data indicate that the same biological and immune characteristics are associated with the spontaneous recovery from ITP and CSR following treatment with IVIg.

3.5 | Addition of the biological score to clinical data improved the prediction of ITP disease courses

Because the biological risk score identified favorable ITP disease courses, we were next interested in comparing the biological score to a previously proposed clinical score by Edslev et al.[38] In short, this score uses the six clinical characteristics of age, sex, bleeding symptoms, preceding infection, platelet count, and symptom duration to categorize patients into low, intermediate, and high risk of spontaneous recovery groups 3 months after the diagnosis of ITP. We assessed this effect in the observation cohort of the TIKI study because these data were not included during the derivation of the biological risk score. We observed that the biological risk score improved the discrimination of disease courses, as a multivariate Cox proportional hazard model for complete recovery from ITP showed that both the Edslev score and the biological risk score were independently associated with recovery, with an HR of 1.25 (95% CI, 1.12-1.40) and 0.69 (95% CI, 0.52-0.92), respectively. This is illustrated by plotting the predicted ASR and CSR categories of the risk groups (Figure 5). We observed a clear discrimination of patients with a high chance of spontaneous recovery, particularly in patients with intermediate risk by the Edslev score (Figure 5; Table S8). In summary, in ITP patients aged younger than 7 years, the use the biological risk score in addition to clinical characteristics improved the stratification of the disease course and response to IVIg.

4 | DISCUSSION

4.1 | Main results

In this study, we show how multimodal clinical, genetic, and immune data can be combined with longitudinal clinical data to stratify patients with a complex hematological disease for the clinical prediction of disease courses and treatment outcomes. We

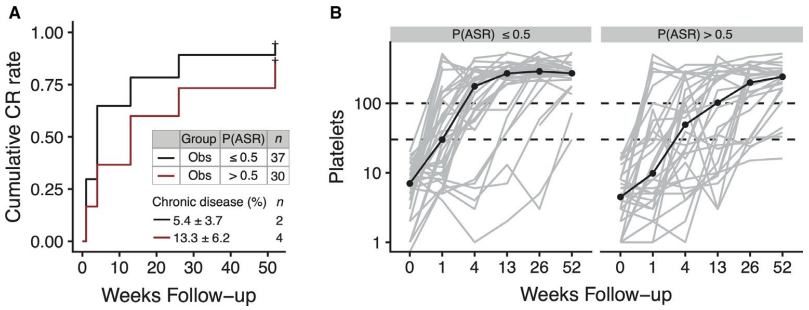


FIGURE 4 Validation of the biological risk score with disease courses in patients randomized to the observation group. (A) Kaplan-Meier plot of the recovery stratified by the biological risk score. Similar to the patients treated with IVIg, the score stratified patients with a higher rate of complete recovery (CR) and enriched patients with chronic disease with a posterior probability of P(ASR) >0.5. Statistics are provided in the main text. (B) This was also observed when the platelet count was assessed over time, as stratified by the predicted response P(ASR) >0.5 (N = 30) *vs.* ≤0.5 (N = 37). The individual patient platelet count trajectories (gray lines) and medians (black line) are shown. IVIg, intravenous immunoglobulin

TABLE 3 Proportion of patients exhibiting complete recovery during 1-year follow-up, stratified by the biological risk score

	IVIg			Observation		
	Pretest	P(ASR) ≤0.5	P(ASR) >0.5	Pretest	P(ASR) ≤0.5	P(ASR) >0.5
Month 3	0.86 (69/80)	0.94	0.71	0.68 (45/67)	0.81	0.53
Month 6	0.88 (70/80)	0.92	0.79	0.82 (55/67)	0.89	0.73
Month 12	0.92 (74/80)	0.98	0.82	0.91 (61/67)	0.95	0.87

Note: Data are the proportion (n/N) of patients exhibiting a complete recovery at the indicated timepoint. P(ASR) given by the biological risk score.

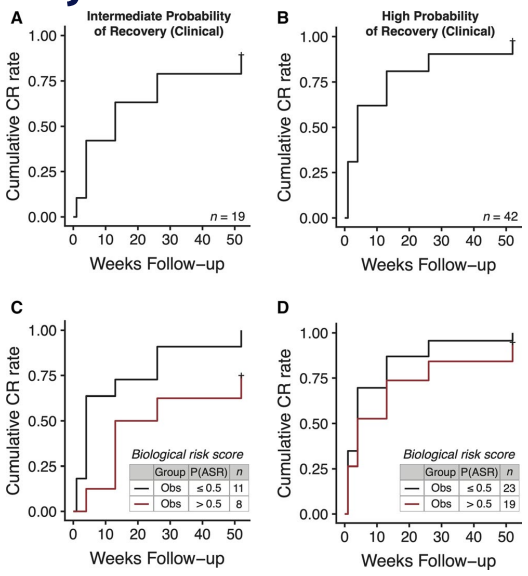


FIGURE 5 Prediction of the disease course using the biological risk score provides enhanced prognostic information for a prediction score based on the clinical characteristics alone. Patients randomized to the observation group were sorted in prognostic strata according to the NOPHO score (Edslev et al. 2007[38]). This Bayesian score uses six clinical characteristics to indicate a low, intermediate, and high risk of spontaneous recovery 3 months after the diagnosis. In particular, the score can accurately enrich patients with a low risk of spontaneous recovery, whereas intermediate- and high-risk groups display little change between the prior and posterior probability.[38] A key criterion of the score (age less than 10 years) was fulfilled for all patients. Therefore, the discriminative capacity is already reduced in this group of patients younger than seven years of age. Only six patients were found to have a low risk and were not included here. Differentiation of disease courses by the clinical prediction score. Kaplan-Meier plot for the complete recovery in patients with an (A) intermediate NOPHO score (N = 19) and (B) high NOPHO score (N = 42). (C) Patients in the intermediate risk group stratified by the biological risk score showed further discrimination into favorable and unfavorable disease courses, with an enrichment of prolonged and chronic disease. (D) Stratification of a high NOPHO risk by the biological risk score, also showed further discriminatio, although less pronounced. Statistics are provided in the main text

profiled ITP patients at the time of diagnosis on both immunological and genetic parameters. Various datasets were integrated to identify the variables that predict the ITP disease course. This approach relates to systems medicine, and recent studies in the field of systems immunology[39] have illustrated the strength of such combined laboratory and computational work for the identification of predictors of clinically important outcomes from a multitude of possible variables, for example, in the context of vaccine responses[40] or leukemia relapse.[41] However, this approach has never been applied to ITP.

A key finding in this study was that the patients who were predicted by biological parameters to respond to IVIg (i.e., with an increase in platelet counts that lasted for at least 1 month) would otherwise show a self-limiting disease course. This finding is in line with the fact that IVIg leads to a short-term resolution of thrombocytopenia; however, it has no effect on the long-term recovery rates.[10] A biological risk score of five variables was able to identify the disease course in ITP: hemoglobin; platelet count; *FCGR2C* genotype; the presence of circulating anti-platelet glycoprotein-specific IgG antibodies; and a preceding vaccination. These factors may be implicated in the pathophysiology of ITP. By testing the association of this immune and genetic profile in the observation group of the TIKI study, we ascertained that a favorable disease course in observed patients was also associated with this biological profile.

4.2 | Findings of others

What determines the disease course in ITP? Although a multitude of predictors have been proposed by previous studies, they are often not tested for the discriminative potential of prospective disease courses, nor integrated into a multivariate prediction.[42] In a systematic review, the factors of male gender, age ≤11 years, platelet count ≤20 × 10^9/L, abrupt disease onset, mucosal bleeding, the absence of antinuclear antibodies, a preceding infection, and leukocytosis at diagnosis have all been described as potential predictors of an early resolution of thrombocytopenia.[43] With the exception of the platelet count, none of these predictors were selected by our regression procedure. There could be several reasons for this effect: (1) they could correlate with predictors that were identified; (2) they have insufficient discriminatory capacity between disease courses; (3) they correlate with age (effects of which were removed in our study); or (4) they are relevant for a subgroup of patients older than 7 years of age. Importantly, our study did not test for causal associations, and variables were only selected variables for their ability to predict the disease course.

At large, it remains unknown why some children do not respond to IVIg. The major hypotheses for the working mechanisms of IVIg are the enhanced catabolism of pathogenic IgG by the saturation of the neonatal Fc Receptor and the blocking of myeloid Fc-receptors (FcγR) that clear IgG-sensitized platelets; however, the exact mechanism remains unresolved.[44-46] On the one hand, genetic diversity might determine the response to IVIg, including the FcγRIIB-p.232 T/T genotype, which has been suggested to be associated with a failure to respond to IVIg in mouse studies and limited human data.[10,37] However, these variants are rare and insufficient to explain the large proportion of non-responders. Retrospective case series have suggested that IVIg-resistant patients have higher rates of leukocyte count ≤7 × 10^9/L,[47] age above 2 years, a platelet count <9 × 10^9/L,[48] and a higher expression of IFN-γ in peripheral blood mononuclear cells.[49] Although the absence of leukocytosis could be associated with a noninflammatory phenotype, potentially pointing to the absence of autoinflammatory or postinfectious reactive autoimmunity, this is in contrast with the increased IFN-γ expression.

Thus, these findings remain difficult to unify in a conceptual model that explains the ITP disease course.

The association between an absence of IgG platelet autoantibodies and nonresponsiveness to IVIg therapy is in line with the findings of adult ITP treated with rituximab in ITP.[50] Circulating IgG autoantibodies in our patient population were rare. Of note, an ideal predictor of disease courses would be present in only a subset of patients (i.e., have a low or moderate sensitivity, such as glycoprotein-specific autoantibodies[51-53]). In contrast to other studies, IgG anti-GP Ib/IX autoantibodies in our study were not predictive of a lack of response to IVIg.[21,35,36] Interestingly, some recently discovered pathogenic effector functions of antiplatelet antibodies may not be influenced by IVIg treatment, such as engagement with the myeloid Fc-γ receptor, platelet sialic acid cleavage, or complement activation.[54,55]

4.3 | Strengths and limitations

A key strength of this study is the testing and thereby validation of the model in patients that were randomized to the observation group within the same trial. Our statistical analysis has several advantages because regularized regression can identify a limited number of predictors in an unbiased, data-driven manner among a wide range of candidate predictors, whereas overfitting is limited. Nonetheless, a variable with a low frequency in the overall population, like preceding vaccination, may represent a spurious association; this can only be assessed during external validation. The generalizability of our study is limited by the employed inclusion criteria (i.e., children with platelet count <20 × 10^9 /L without severe or life-threatening bleeding). In addition, a maximum age criterion of 7 years was used for the present study, which represents 75% of the original study population and a majority of children that are clinically diagnosed with ITP. This restriction on an age younger than 7 years was introduced because of the known heterogeneity in the clinical course and characteristics among older age groups within childhood ITP [25,42](Schmidt et al., manuscript submitted; https://doi.org/10.1101/2020.06.09.20125385). If a greater number of patients older than age 7 years had been available, it would have been possible to perform an analysis stratified by age.

Finally, the predictive ability may increase if we had analyzed other genetic variants associated with thrombocytopathies and immune deficiencies, or performed a deeper immune phenotyping and cytokine analysis. Unfortunately, these analyses were not available and future studies may take this into consideration.

4.4 | Implications

Given the limited sample size, our data should not be seen as evidence-based diagnostic or therapeutic recommendations. When a beneficial response was predicted, patients almost exclusively (91%) exhibited CSR and a substantially higher spontaneous recovery rate in the observation cohort. In such cases, our data suggest that IVIg treatment could be omitted. However, these patients would likely respond if IVIg was administered in emergency situations. Moreover, the patients could be managed with fewer repeated clinical assessments and laboratory investigations, which may be considered to be positive for some patients. At the same time, it is valuable to reconsider the benefit of IVIg in cases in which a full response can be expected. Clearly, the time to recovery can be shortened with the administration of IVIg and bleeding episodes can be prevented. In the TIKI trial,[10] patients in the observation cohort were hospitalized at approximately the same rate as the patients in the IVIg cohort for bleeding events and treatment side effects, respectively. If IVIg can be targeted toward the patients who are expected to achieve the greatest benefit (i.e., the patients predicted to exhibit a response to treatment), the benefit-to-side effect ratio may shift to become more favorable.

Some clinicians regard a failure to respond to IVIg as a reason to reconsider the diagnosis of ITP.[56] Patients who are predicted to display ASR could be targeted for additional diagnostic testing or the early screening of both autoimmune diseases and immune deficiencies. Overall, the aim of reducing ITP-specific treatments in patients with alternative causes of thrombocytopenia is critical to reduce unnecessary side effects.

5 | CONCLUSIONS

We were able to show that the ITP patients who were predicted to respond to IVIg by biological parameters would show self-limiting disease courses without treatment. The biological risk score thus allowed for the accurate prediction of transient disease courses, which may be of assistance for treatment decisions and aid in counselling of patients and their parents. Future studies should validate our findings in an unrelated cohort, investigate additional immune parameters, the optimal timepoints for additional diagnostic testing, and potentially investigate alternative management strategies for patients predicted to show unfavorable disease courses.

ACKNOWLEDGMENTS

We are particularly grateful to Laurens Bosman (University Medical Center Utrecht) for discussions regarding statistical modelling. This study would not have been possible without the dedicated teams at the central laboratory of this study; of the Laboratory of Platelet Serology, Sanquin Diagnostics, in particular Elly Huiskes; and the team of the Immunocytology Laboratory, Sanquin Diagnostics, in particular Christa Homburg. The study was supported by a research grant from the Landsteiner Foundation for Blood Transfusion Research (LSBR), a doctoral stipend to D.E.S. by the Studienstiftung des Deutschen Volkes, and Wilhelmina Children's Hospital Utrecht.

CONFLICT OF INTEREST

The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

David E. Schmidt codesigned the study, analyzed data, and wrote the manuscript. Katja M.J. Heitink-Pollé and Marrie C.A. Bruin collected data, provided input on the study design, and interpreted data. Bart Mertens supervised statistical analyses and codesigned the study. Leendert Porcelijn, C. Ellen van der Schoot, Gestur Vidarsson, Rick Kapur, and Johanna G. van der Bom discussed the study design and interpreted data. Masja de Haas supervised the study. All authors revised and approved the final version of the manuscript.

ORCID

David E. Schmidt https://orcid.org/0000-0001-5555-376X
Gestur Vidarsson https://orcid.org/0000-0001-5621-003X

REFERENCES

1. Barturen G, Beretta L, Cervera R, Van Vollenhoven R, Alarcón-Riquelme ME. Moving towards a molecular taxonomy of autoimmune rheumatic diseases. *Nat Rev Rheumatol*. 2018;14(2):75–93.
2. Cho JH, Gregersen PK. Genomics and the multifactorial nature of human autoimmune disease. Feero WG, Guttmacher AE, editors. *N Engl J Med*. 2011;365(17):1612-1623.
3. Terrell DR, Beebe LA, Vesely SK, Neas BR, Segal JB, George JN. The incidence of immune thrombocytopenic purpura in children and adults: a critical review of published reports. *Am J Hematol*. 2010;85(3):174-180.
4. Rodeghiero F, Stasi R, Gernsheimer T, et al. Standardization of terminology, definitions and outcome criteria in immune thrombocytopenic purpura of adults and children: report from an international working group. *Am Soc Hematol*. 2009;113(11):2386-2393.
5. Rosthøj S, Hedlund-Treutiger I, Rajantie J, et al. Duration and morbidity of newly diagnosed idiopathic thrombocytopenic purpura in children: a prospective Nordic study of an unselected cohort. *J Pediatr*. 2003;143(3):302-307.
6. Kühne T, Buchanan GR, Zimmerman S, et al. A prospective comparative study of 2540 infants and children with newly diagnosed idiopathic thrombocytopenic purpura (ITP) from the Intercontinental Childhood ITP Study Group. *J Pediatr*. 2003;143(5):605-608.
7. Heitink-Polle KMJ, Haverman L, Annink KV, Schep SJ, de Haas M, Bruin MCA. Health-related quality of life in children with newly diagnosed immune thrombocytopenia. *Haematologica*. 2014;99(9):1525-1531.
8. Flores A, Klaassen RJ, Buchanan GR, Neunert CE. Patterns and influences in health-related quality of life in children with immune thrombocytopenia: a study from the Dallas ITP Cohort. *Pediatr Blood Cancer*. 2017;64(8):e26405.
9. Neunert C, Terrell DR, Arnold DM, et al. American Society of Hematology 2019 guidelines for immune thrombocytopenia. *Blood Adv*. 2019;3(23):3829-3866.
10. Heitink-Pollé KMJ, Uiterwaal CSPM, Porcelijn L, et al. Intravenous immunoglobulin vs observation in childhood immune thrombocytopenia: a randomized controlled trial. *Blood*. 2018;132(9):883-891.
11. Neunert C, Lim W, Crowther M, et al. The American Society of Hematology 2011 evidence-based practice guideline for immune thrombocytopenia. *Blood*. 2011;117:4190-4207.
12. Imbach P, Wagner HP, Berchtold W, et al. Intravenous immunoglobulin versus oral corticosteroids in acute immune thrombocytopenic purpura in childhood. *Lancet*. 1985;2(8453):464-468.
13. Blanchette VS, Luke B, Andrew M, et al. A prospective, randomized trial of high-dose intravenous immune globulin G therapy, oral prednisone therapy, and no therapy in childhood acute immune thrombocytopenic purpura. *J Pediatr*. 1993;123(6):989-995.
14. Blanchette V, Imbach P, Andrew M, et al. Randomised trial of intravenous immunoglobulin G, intravenous anti-D, and oral prednisone in childhood acute immune thrombocytopenic purpura. *Lancet*. 1994;344(8924):703-707.
15. Duru F, Fisgin T, Yarali N, Kara A. Clinical course of children with immune thrombocytopenic purpura treated with intravenous immunoglobulin G or megadose methylprednisolone or observed without therapy. *Pediatr Hematol Oncol*. 2002;19(4):219-225.
16. Fujisawa K, Iyori H, Ohkawa H, et al. A prospective, randomized trial of conventional, dose-accelerated corticosteroids and intravenous immunoglobulin in children with newly diagnosed idiopathic thrombocytopenic purpura. *Int J Hematol*. 2000;72(3):376-383.
17. Choi HS, Ji MH, Kim SJ, Ahn HS. Platelet count recovery after intravenous immunoglobulin predicts a favorable outcome in children with immune thrombocytopenia. *Blood Res*. 2016;51(2):95-97.
18. Perez Botero J, Im R, Chen D, Patnaik MM. Grey platelet syndrome misdiagnosed as ITP. *Br J Haematol*. 2016;173(5):662.
19. Downes K, Megy K, Duarte D, et al. Diagnostic high-throughput sequencing of 2,396 patients with bleeding, thrombotic and platelet disorders. *Blood*. 2019;134(23):2082-2091.
20. Podjasek JC, Abraham RS. Autoimmune cytopenias in common variable immunodeficiency. *Front Immunol*. 2012;3:189.
21. Peng J, Ma S-H, Liu J, et al. Association of autoantibody specificity and response to intravenous immunoglobulin G therapy in immune thrombocytopenia: a multicenter cohort study. *J Thromb Haemost*. 2014;12(4):497-504.
22. Zufferey A, Kapur R, Semple JW. Pathogenesis and therapeutic mechanisms in immune thrombocytopenia (ITP). *J Clin Med*. 2017;6(2):16.
23. Kuhne T, Berchtold W, Michaels LA, et al. Newly diagnosed immune thrombocytopenia in children and adults: a comparative prospective observational registry of the Intercontinental Cooperative Immune Thrombocytopenia Study Group. *Haematologica*. 2011;96(12):1831-1837.
24. Despotovic JM, Grimes AB. Pediatric ITP: is it different from adult ITP? *Hematology Am Soc Hematol Educ Program*. 2018;2018(1):405-411.
25. Lowe EJ, Buchanan GR. Idiopathic thrombocytopenic purpura diagnosed during the second decade of life. *J Pediatr*. 2002;141(2):253-258.
26. van den Heuvel D, Jansen MAE, Nasserinejad K, et al. Effects of nongenetic factors on immune cell dynamics in early childhood: the Generation R Study. *J Allergy Clin Immunol*. 2017;139(6):1923-1934.e17.
27. Carr EJ, Dooley J, Garcia-Perez JE, et al. The cellular composition of the human immune system is shaped by age and cohabitation. *Nat Immunol*. 2016;17(4):461-468.
28. Despotovic JM, Grimes AB. Pediatric ITP: is it different from adult ITP? *Hematology*. 2018;2018(1):405-411.
29. Buchanan GR, Adix L. Grading of hemorrhage in children with idiopathic thrombocytopenic purpura. *J Pediatr*. 2002;141(5):683-688.
30. Bennett CM, Rogers ZR, Kinnamon DD, et al. Prospective phase 1/2 study of rituximab in childhood and adolescent chronic immune thrombocytopenic purpura. *Blood*. 2006;107(7):2639-2642.
31. van Buuren S, Groothuis-Oudshoorn K. Mice: Multivariate Imputation by Chained Equations in R. *J Stat Software*. 2011;45(3):1–67.
32. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Software*. 2010;33(1):1-22.
33. Carpenter JR, Kenward MG. *Multiple Imputation and its Application*. Hoboken, NJ: John Wiley & Sons, Ltd; 2013.
34. Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12(1):77.

35. Go RS, Johnston KL, Bruden KC. The association between platelet autoantibody specificity and response to intravenous immunoglobulin G in the treatment of patients with immune thrombocytopenia. *Haematologica*. 2007;92(2):283-284.
36. Fu L, Cheng Z, Gu H, Wu R. Platelet-specific antibodies and differences in their expression in childhood immune thrombocytopenic purpura predicts clinical prognosis. *Pediatr Invest*. 2018;2(4):230-235.
37. Samuelsson A, Towers TL, Ravetch JV. Anti-inflammatory activity of IVIG mediated through the inhibitory Fc receptor. *Science*. 2001;291(5503):484-486.
38. Edslev PW, Rosthøj S, Treutiger I, et al. A clinical score predicting a brief and uneventful course of newly diagnosed idiopathic thrombocytopenic purpura in children. *Br J Haemat*. 2007;138(4):513-516.
39. Davis MM, Tato CM, Furman D. Systems immunology: just getting started. *Nat Immunol*. 2017;18(7):725-732.
40. Tsang JS, Schwartzberg PL, Kotliarov Y, et al. Global analyses of human immune variation reveal baseline predictors of postvaccination responses. *Cell*. 2014;157(2):499-513.
41. Good Z, Sarno J, Jager A, et al. Single-cell developmental classification of B cell precursor acute lymphoblastic leukemia at diagnosis reveals predictors of relapse. *Nat Med*. 2018;24(4):474-483.
42. Schmidt DE, Heitink-Pollé KMJ, Bruin MCA, de Haas M. Intravenous immunoglobulins (IVIg) in childhood immune thrombocytopenia: Towards personalized medicine – A narrative review. *Annal Blood*. 2021; E-pub. https://doi.org/10.21037/aob-20-59
43. Heitink-Pollé KMJ, Nijsten J, Boonacker CWB, de Haas M, Bruin MCA. Clinical and laboratory predictors of chronic immune thrombocytopenia in children: a systematic review and meta-analysis. *Blood*. 2014;124(22):3295-3307.
44. Bleeker WK, Teeling JL, Hack CE. Accelerated autoantibody clearance by intravenous immunoglobulin therapy: studies in experimental models to determine the magnitude and time course of the effect. *Blood*. 2001;98(10):3136-3142.
45. Schwab I, Nimmerjahn F. Intravenous immunoglobulin therapy: how does IgG modulate the immune system? *Nat Rev Immunol*. 2013;13(3):176-189.
46. Debré M, Bonnet MC, Fridman WH, et al. Infusion of Fc gamma fragments for treatment of children with acute immune thrombocytopenic purpura. *Lancet*. 1993;342(8877):945-949.
47. Morimoto Y, Yoshida N, Kawashima N, Matsumoto K, Kato K. Identification of predictive factors for response to intravenous immunoglobulin treatment in children with immune thrombocytopenia. *Int J Hematol*. 2014;99(5):597-602.
48. Higashide Y, Hori T, Yoto Y, et al. Predictive factors of response to IVIG in pediatric immune thrombocytopenic purpura. *Pediatr Int*. 2018;60(4):357-361.
49. Mouzaki A, Theodoropoulou M, Gianakopoulos I, Vlaha V, Kyrtsonis M-C, Maniatis A. Expression patterns of Th1 and Th2 cytokine genes in childhood idiopathic thrombocytopenic purpura (ITP) at presentation and their modulation by intravenous immunoglobulin G (IVIg) treatment: their role in prognosis. *Blood*. 2002;100(5):1774-1779.
50. Porcelijn L, Huiskes E, Schipperus M, van der Holt B, de Haas M, Zwaginga JJ. Lack of detectable platelet autoantibodies is correlated with non-responsiveness to rituximab treatment in ITP patients. *Blood*. 2017;129(25):3389-3391.
51. Schmidt DE, Heitink-Pollé KMJ, Porcelijn L, et al. Anti-platelet antibodies in childhood immune thrombocytopenia: prevalence and prognostic implications. *J Thromb Haemost*. 2020;18(5):1210-1220.
52. Schmidt DE, Lakerveld AJ, Heitink-Pollé KMJ, et al. Anti-platelet antibody immunoassays in childhood immune thrombocytopenia: a systematic review. *Vox Sang*. 2020;20(133):257.
53. Al-Samkari H, Rosovsky RP, Karp Leaf RS, et al. A modern reassessment of glycoprotein-specific direct platelet autoantibody testing in immune thrombocytopenia. *Blood Adv*. 2020;4(1):9-18.
54. Marini I, Zlamal J, Faul C, et al. Autoantibody-mediated desialylation impairs human thrombopoiesis and platelet life span. *Haematologica*. 2019;106(1):196-207.
55. Najaoui A, Bakchoul T, Stoy J, et al. Autoantibody-mediated complement activation on platelets is a common finding in patients with immune thrombocytopenic purpura (ITP). *Eur J Haematol*. 2012;88(2):167-174.
56. Cines DB, Bussel JB. How I treat idiopathic thrombocytopenic purpura (ITP). *Blood*. 2005;106(7):2244-2251.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Schmidt DE, Heitink-Pollé KM, Mertens B, et al. Biological stratification of clinical disease courses in childhood immune thrombocytopenia. *J Thromb Haemost*. 2021;19:1071–1081. https://doi.org/10.1111/jth.15232

Supplementary Methods

Samples

At diagnosis and each follow-up visit, the patient blood samples were transferred to a central laboratory facility (Laboratory for Platelet and Leukocyte Serology; Sanquin Diagnostics, Amsterdam, The Netherlands). The serum and plasma samples were stored at -20°C. DNA was isolated from the whole blood or buffy coat (QIAamp DNA blood mini kit, Qiagen Benelux, Venlo, The Netherlands). Heparin-anticoagulated blood was immediately processed for immune phenotyping.

Genotyping of IgG Fc-receptors

Genotyping of the myeloid Fc-γ receptor locus was performed by multiplex ligation dependent probe amplification (MLPA), as previously described [1,2].

Circulating platelet autoantibodies

Platelet autoantibody testing was performed in our national reference laboratory (Laboratory for Platelet and Leukocyte Serology; Sanquin Diagnostics) by evaluation of reactivity of patient sera with preserved healthy donor platelets. Circulating platelet autoantibodies were detected by platelet immunofluorescence (PIFT) [3]. Glycoprotein-specific autoantibodies to platelet GP IIb/IIIa, GP Ib/IX and GP V were detected by monoclonal antibody immobilization of platelet antigen technique (MAIPA) [4,5]. Low- and high titer controls were used as internal controls. Of note, the sensitivity of MAIPA for circulating glycoprotein-specific antibodies in childhood ITP is low-to-moderate [5,6].

Plasma Thrombopoietin levels

A solid-phase sandwich ELISA for measurement of plasma TPO concentrations (AU/ml) was performed as previously described [7]. One AU equals 9 pg of recombinant Tpo (Research Diagnostics, Flanders, NJ, USA).

Peripheral blood immune phenotyping and analysis of flow cytometry data

Heparin-anticoagulated tubes were transported on the day of venipuncture to our central testing laboratory and immune phenotyping was performed (Sanquin Immunodiagnostics).

Absolute and relative cell counts for total T, CD4, CD8, B and NK cells stained by CD3 FITC, CD16 PE, CD56 PE, CD45 PerCP-Cy5.5, CD4 PE-Cy7, CD19 APC and CD8 APC-Cy7 (BD Biosciences; Vianen, The Netherlands; #644611) were performed in TruCount tubes (BD Biosciences #340334) using 25µL heparin-anticoagulated whole blood. Regulatory T cell staining was performed using CD4 FITC, CD25 PE-Cy7, CD127-Alexa Fluor 647 (BD 560249) and CD3 PerCP (BD Biosciences #345766). T-cell subsets were stained by CD3 PerCP (BD Biosciences #345766), CD4 APC (BD Biosciences #345771), CD45RO PE (BD Biosciences #347967) and CD27 FITC (Sanquin Reagents; Amsterdam, The Netherlands; M1764). T cell subsets were defined as previously described[8,9]: CD4 Treg, $CD4^+$ $CD25^+$ $CD127^{lo}$ cells; naive CD4 T, $CD4^+$ $CD45RO^-$ $CD27^+$; CD4 central memory T, $CD4^+$ $CD45RO^+$ $CD27^+$; CD4 effector memory T, $CD4^+$ $CD45RO^+$ $CD27^-$ (Supplementary Figure 5). For evaluation of cell numbers and frequencies, FCS files were imported in flowJo or BD FACSDiva and analyzed with a standardized workflow. Each sample was separately visualized and inspected.

Statistical analysis

Statistical analyses were performed in R version 3.5.1 (http://www.r-project.org). Samples were included by availability from the clinical trial; no formal sample size calculation was performed. Continuous variables were tested with a Welch's t-test or a Mann-Whitney *U* test. Frequencies were compared with a Fisher's exact test in case of cells with <5 observations, or a Chi-square test. A heatmap was built using age-corrected residuals that were scaled separately for each continuous variable. Kaplan-Meier curves were constructed using the package *survival* and hazard ratios of the score calculated by a Cox-proportional hazard model, given for a one-unit increase of the score. A two-tailed p-value <0.05 was considered as statistically significant.

Molecular risk score to predict the treatment response to IVIg

The statistical procedure to build the prediction model is summarized in Supplementary Figure 6. Missing variables (Supplementary Table 2) were imputed by multiple imputation using chained equations [10] for a total number of 100 imputed datasets. In the IVIg group, the endpoint was imputed for one patient who exhibited a missing datapoint at one week due to a clotted EDTA tube. The age-associated changes in immune variables were first removed from the data by constructing a linear model for each numerical variable and using the model

residuals as features (see Supplementary Table 3). To prevent spurious associations, the eosinophil and basophil counts were excluded due to near-zero variance. The categorical variables were dummy-coded; continuous variables were centered at the mean and standardized during the model procedure. A classifier was built on the patients in the IVIg cohort to predict the binary outcome of a complete *vs.* absent sustained response to IVIg, using regularized logistic regression with the elastic net by *glmnet.* [11] Elastic net is a penalized regression technique that favors sparse models and prevents overfitting through the shrinkage of regression coefficients. Elastic net also allows for complex predictor associations with the outcome, including multicollinearity. Within each imputed dataset 100-fold bootstrapping was performed. Shrinkage parameters were selected by ten-fold cross-validation. Model features were selected by the inclusion frequency of a variable.

Feature selection using elastic net

The elastic net penalty provides feature selection by controlling the gap of the lasso (α=1.0) and ridge penalty (α=0.0). With $\alpha > 0$, some coefficients are set to zero. We used this feature to assign coefficients an importance metric, the *inclusion frequency*, defined as the number of times a coefficient had a nonzero estimate out of all built models. Coefficients were subsequently ranked by their inclusion frequency across all built models. By out-of-sample cross-validation, we chose optimal number of ranked coefficients to be included for discrimination at a corresponding α parameter, as illustrated in Figure 2A in the main text of this manuscript. In case that a factor with >2 categories was included amongst the best ranked coefficients, the coefficients of all categories of the factor were included for the final model (e.g. for a diallelic SNP with alleles *a* and *b*, both heterozygous *a/b* and homozygous *b/b* factor levels, as compared to the reference *a/a*).

Determination of model coefficients, standard error and 95% confidence intervals

After feature selection, Rubin's rules were applied to derive coefficients from the *B* bootstrapped models built in *K* imputed datasets. The procedure is detailed in chapter 2 of *Multiple Imputation and its Application*; First Edition, James R. Carpenter and Michael G. Kenward; 2013 John Wiley & Sons. Under the assumption of *missing at random*, the multiple imputation estimator of each coefficient from *K* imputed datasets is $\hat{\beta}_{MI} = \frac{1}{K}\sum_{k=1}^{K}\hat{\beta}_k$ where each imputed dataset has the point estimate $\hat{\beta}_k$ and variance $\hat{\sigma}_k^2$. The variance estimator of $\hat{\beta}_{MI}$ is given by

$$\hat{V}_{MI} = \widehat{W} + \left(1 + \frac{1}{K}\right)\hat{B}$$

with the within- and between-imputation variance

$$\widehat{W} = \frac{1}{K}\sum_{k=1}^{K}\hat{\sigma}_k^2$$

and

$$\hat{B} = \frac{1}{K=1}\sum_{k=1}^{K}\left(\hat{\beta}_k - \hat{\beta}_{MI}\right)^2$$

The standard error of the multiple imputation estimator $\hat{\beta}_{MI}$ was obtained by $\sqrt{\hat{V}_{MI}}$ and this was used to obtain 95% confidence intervals of $\hat{\beta}_{MI}$. Given that *K* was large, the variance reduces to $\hat{V}_{MI} = \widehat{W} + \hat{B}$ and the Bayesian posterior of $\hat{\beta}_{MI}$can be interpreted with $\hat{\beta}_{MI} \sim N\left(\beta, \widehat{W} + \hat{B}\right)$. Subsequently a 100(1-$\alpha$)% confidence interval is given by

$$\left(\hat{\beta}_{MI} - z_{1-\frac{\alpha}{2}}\sqrt{\hat{V}_{MI}}\ , \hat{\beta}_{MI} - z_{\frac{\alpha}{2}}\sqrt{\hat{V}_{MI}}\ \right)$$

Coefficients that were not included in one of the *B* x *K* models due to the feature selection properties of the Elastic Net were treated as zero, i.e. $\hat{\beta}_k$ = 0.

After determining the coefficients, the linear predictor βX was calculated, and the model intercept was determined by a generalized linear model for *response* ~ βX (in R notation).

Calculation of model scores `by hand`

For all continuous predictors the age-corrected values *x'* should be obtained first by calculating the residual on the expectation value of *x* at a certain age, given by the linear predictor: $x' = x - (\beta_0 + \beta_{age} * age)$, where age is given in years. A full list of age coefficients for all continuous predictors is given in Supplementary Table 3.

Variable	**Intercept** (β_0)	**Age coefficient** (β_{age})
Platelet count, x10^9/L	7.141	-0.080
Hemoglobin, mmol/L	7.141	0.100
Neutrophils, x10^9/L	4.539	-0.194
TPO, U/L	62.323	-1.972

Non-standardized coefficients

In the main text of the manuscript standardized coefficients are provided to allow comparison, i.e. the odds ratio for a change in one standard deviation of the predictor. For calculation of scores directly from the age-corrected predictors, the coefficients below should be used that are directly obtained from the elastic net models and pooled by Rubin's Rules.

		Coefficient Molecular risk score		**Coefficient Model #2**
Intercept	β_0	0.590	β_0	0.877
Hemoglobin, mmol/L	β_1	-1.180	β_1	-1.031
Platelet count, x10^9/L	β_2	-0.071	β_2	-0.069
FCGR2C ORF absent		Ref		Ref
FCGR2C ORF present	β_3	-0.832	β_3	-0.677
MAIPA IgG positive	β_4	-0.757	β_4	-0.606
Preceding vaccination	β_5	1.133	β_5	1.385
TPO, U/L			β_6	0.008
Neutrophils, x10^9/L			β_7	-0.111
FCGR2B p.232 I/I				Ref
FCGR2B p.232 I/T			β_8	0.048
FCGR2B p.232 T/T			β_9	1.265

The calculation of the score (linear predictor) is performed to obtain the probability of absence of a sustained response (ASR). For this, considering all *i* variables *x,* the estimated coefficients $\beta_1, \ldots, \beta_i$ are multiplied with the respective (age-corrected) value of x_i and summed up with the model intercept β_0:

$$score = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i$$

The score finally gives the probability for ASR:

$$p(ASR) = \frac{1}{1 + e^{-score}}$$

Note that 1 - p(ASR) is the corresponding probability for CSR.

References

1. Breunis WB, van Mirre E, Bruin M, Geissler J, de Boer M, Peters M, et al. Copy number variation of the activating FCGR2C gene predisposes to idiopathic thrombocytopenic purpura. 2008 Feb 1;111(3):1029–38.

2. Tsang-A-Sjoe MWP, Nagelkerke SQ, Bultink IEM, Geissler J, Tanck MWT, Tacke CE, et al. Fc-gamma receptor polymorphisms differentially influence susceptibility to systemic lupus erythematosus and lupus nephritis. Rheumatology. 2016 Apr 14;55(5):939–48.

3. Borne von dem AEGK, Verheugt FWA, Oosterhof F, Riesz E, Rivière AB, Engelfriet CP. A Simple Immunofluorescence Test for the Detection of Platelet Antibodies. British Journal of Haematology. Blackwell Publishing Ltd; 1978 Jun;39(2):195–207.

4. Porcelijn L, Huiskes E, Oldert G, Schipperus M, Zwaginga JJ, de Haas M. Detection of platelet autoantibodies to identify immune thrombocytopenia: state of the art. British Journal of Haematology. 2018 May 29;39(Suppl. 1):195.

5. Schmidt DE, Heitink-Pollé KMJ, Porcelijn L, van der Schoot CE, Vidarsson G, Bruin MCA, et al. Anti-Platelet Antibodies in Childhood Immune Thrombocytopenia: Prevalence and Prognostic Implications. J Thromb Haemost. 2020 Feb 13;:jth.14762.

6. Schmidt DE, Lakerveld AJ, Heitink-Pollé KMJ, Bruin MCA, Vidarsson G, Porcelijn L, et al. Anti-platelet antibody immunoassays in childhood immune thrombocytopenia: a systematic review. Vox Sang. 2020 Feb 20;133:257.

7. Porcelijn L, Folman CC, Bossers B, Huiskes E, Overbeeke MA, d Schoot von CE, et al. The diagnostic value of thrombopoietin level measurements in thrombocytopenia. Thromb Haemost. 1998 Jun;79(6):1101–5.

8. Fritsch RD, Shen X, Sims GP, Hathcock KS, Hodes RJ, Lipsky PE. Stepwise differentiation of CD4 memory T cells defined by expression of CCR7 and CD27. J Immunol. 2005 Nov 15;175(10):6489–97.

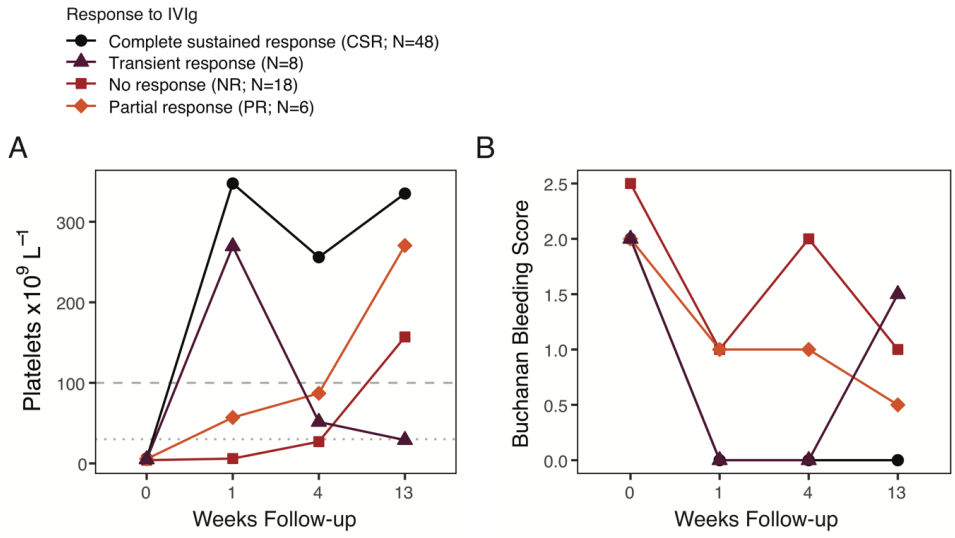
9. Mahnke YD, Brodie TM, Sallusto F, Roederer M, Lugli E. The who's who of T-cell differentiation: human memory T-cell subsets. Eur J Immunol. 2013 Nov;43(11):2797–809.

10. van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software. 2011 Dec 1;45(3).

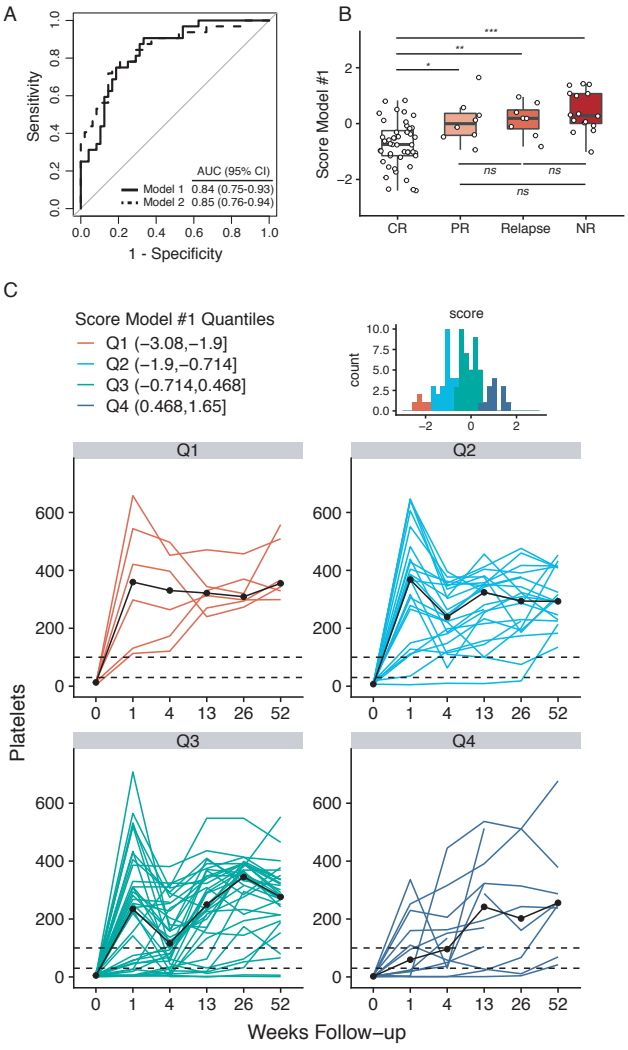
11. Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software. 2010;33(1):1–22.

Supplementary Figures

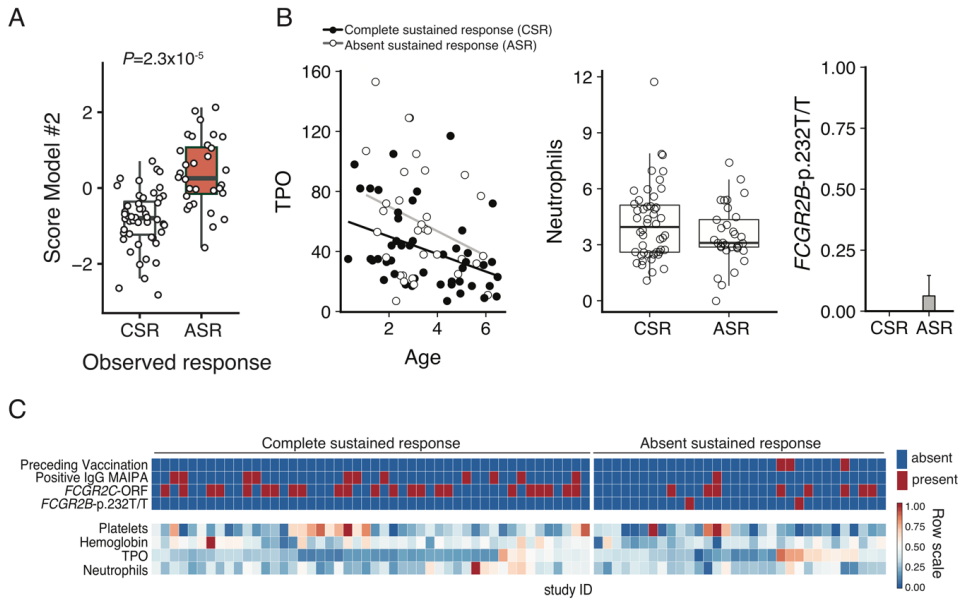
Supplementary Figure 1. Response to IVIg among childhood ITP patients. (A) The response to IVIg correlated with platelet counts during three months follow-up. Based on the platelet counts, we categorized patients as either IVIg responders and non-responders (Rodeghiero et al. Blood 2009; 113(11)). The complete sustained responders (CSR) showed a platelet count > 100 x 10^9 L^{-1} one week after treatment and the response was sustained one month later. If an initial complete response was not maintained at the one-month follow-up, patients were classified as having a transient response. Partial responders (PR) showed an increase to >30 x10^9 L^{-1} with a two-fold increase of baseline platelet count one week after IVIg. Non-responders (NR) showed neither a complete nor a partial response. PR, NR, and transient responders were grouped together as absence of a sustained response (ASR). (B) The response to IVIg is correlated with bleeding symptoms during the follow-up as expected.



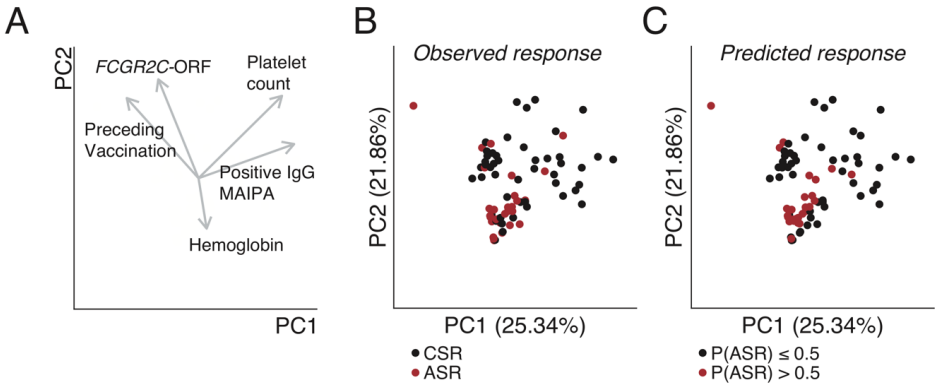
Supplementary Figure 2. The molecular risk score provided similar predictions across all non-response subgroups and discriminated complete sustained responders from non-responders by enrichment around the score extremes. (A) Patients with a complete sustained response (CR) were discriminated from partial responders (PR), patients with a transient response to IVIg (relapse) and non-responders (NR). There were no significant differences in the scores between non-responders. (B) A 5-parameter model (Model #1) with an alpha of 1.0 (elastic-net penalty of lasso) showed similar discrimination by the area under the curve of the receiver operating characteristic as an 8-parameter model (Model #2) with an alpha of 0.2 (elastic-net penalty close to ridge). Derivation of these parameters is given in the main manuscript, Figure 2A. (C) When the score was split in four quantiles from low to high (Q1-Q4), the ratio of complete sustained responders to non-responders decreased from Q1 to Q4. Patients with chronic and prolonged ITP were enriched in Q3 and Q4. Line graphs indicate individual patient's platelet trajectories; black line indicates median platelet counts.



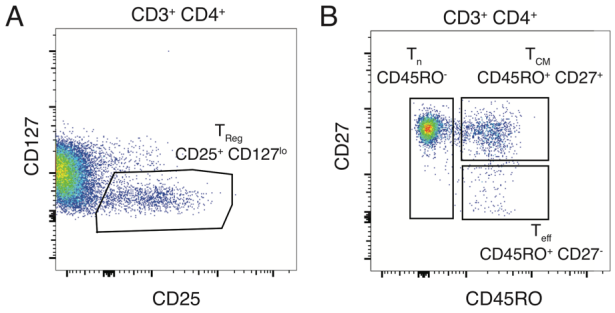
Supplementary Figure 3. Prognosis after IVIg-treatment as predicted by the alternative model (Model #2; α=0.2). (A) Elastic net model (score on log odds scale) displayed discrimination of IVIg responders (CSR; N = 48) and non-responders (ASR; N = 32). The 8 included parameters of the model included all 5 parameters of Model #1 (α=1.0). (B) Variables additionally included in Model #2 were TPO (Wilcoxon rank-sum test; P = 0.033) and neutrophils (Welch t-test; P = 0.117). Moreover, the *FCGR2B*-p.232T/T genotype was solely present in non-responders (Fisher's exact test, P = 0.287). Some of these values did not reach univariate significance and may not have been selected by traditional modelling strategies. (C) Aggregated view on all predictor variables included in the eight-parameter model.



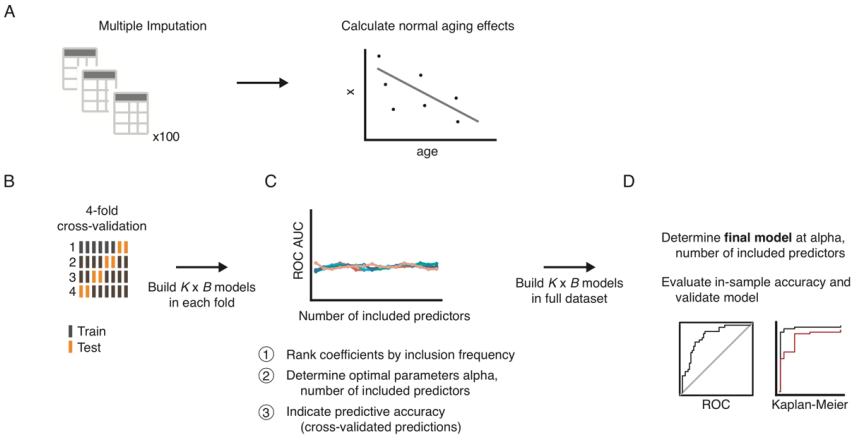
Supplementary Figure 4. A principal component analysis (PCA) of the variables that were included in the molecular risk score. (A) The first two principal components (PC) are displayed along the axes, and the arrows indicate a projection of the loadings. (B) Plotting of the 80 individuals of the IVIg-treated cohort in the PCA with the observed response to IVIg. (C) Similar to B, with the predicted response by the molecular risk score posterior probability of ASR>0.5 (N = 29) *vs*. ≤ 0.5 (N = 51). While some patients were misclassified (switch of the color labels between B and C), the overall classification was favorable.



Supplementary Figure 5. Determination of T cell subsets frequencies by flow cytometry on fresh whole blood. Separate tubes were stained for surface markers of regulatory T cell (A) and naive, central memory and effector T cells (B). A preselection was made on CD3+ CD4+ cells and lymphocytes in FSC/SSC. Gates were set manually and inspected visually for each sample.



Supplementary Figure 6. Modeling strategy for determination of the molecular risk score. (A) Multiple imputation was performed for missing data using *mice*. Aging effects present in the data were removed by constructing a linear model for each continuous variable and using the residuals for further modeling. (B) The 80 IVIg-treated patients in the dataset were distributed among 4 folds for external 4-fold cross-validation, using stratified sampling from response subgroups (CR/NR/PR/relapse). At each candidate elastic net *alpha* value, a model was built in three folds (training data) and evaluated in the remaining fold (test data). This was repeated until all folds were tested once. For building the model in three folds, (1) within each imputed dataset *K*, *B* bootstrap samples were taken from the training data and *glmnet* was run in each, with the predictor coefficients saved. (2) Coefficients were pooled across the imputed datasets using Rubin's Rules. (3) Using the model's results, the predictors were ranked according to their inclusion frequency and the top *s* {1, … , 20} predictors were chosen to predict the response in the test data. The whole procedure was repeated until all folds were predicted exactly once at each *alpha*. (B) Predictions for a number of *s* parameters at each *alpha* were compared to select the best performing parameters during cross-validation. The discriminative performance (predictive accuracy on cross-validated predictions) was evaluated. (C) A final model was built with the determined *alpha,* and number of *s* parameters by running *glmnet* in each imputed dataset *K* for a number of *B* times with bootstrapping. Again, the *K* x *B* model results were combined by Rubin's Rules.



Supplementary Tables

Supplementary Table 1. Predictors for molecular risk score.

Variable	Type
gender	dichotomous
precedingInfection	dichotomous
precedingVaccination	dichotomous
CNR2	dichotomous
F2A_H131R	categorical
F2A_Q27W	categorical
F3B_HNA1c	dichotomous
F2B_I232T	categorical
F2C_nc_ORF	dichotomous
F2B_4	dichotomous
MAIPA_IgM_GPIIbIIIa_dich	dichotomous
MAIPA_IgM_GPIbIX_dich	dichotomous
MAIPA_IgM_GPV_dich	dichotomous
MAIPA_IgG_GPIb_dich	dichotomous
MAIPA_IgG_GPV_dich	dichotomous
MAIPA_IgG_GPIIbIIIa_dich	dichotomous
MAIPA_IgM_dich	dichotomous
MAIPA_IgG_dich	dichotomous
PIFT_dich	dichotomous
TPObinary	dichotomous
MLPA_ORF	dichotomous
MLPA_CNR1	categorical
MLPA_HNA1_Nr_a	categorical
MLPA_158_Nr_V	categorical
bleeding_diagnosis	dichotomous
disease_onset	dichotomous
age_years	continous
hb	continous
leukocytes	continous
neutrophils	continous
monocytes	continous
TPO	continous
Total_IgG	continous
IP_lympho_total	continous
CD3_total	continous
CD4_total	continous
CD8_total	continous
CD4CD8ratio_total	continous

Variable	Type
CD19_total	continous
NK_total	continous
CD3_relative	continous
CD4_relative	continous
CD8_relative	continous
CD19_relative	continous
NK_relative	continous
plt	continuous

Supplementary Table 2. Missing predictor data before multiple imputation

Variable	N missing (IVIG)	N missing (Observation)
PIFT	0	7
precedingInfection	0	2
precedingVaccination	0	2
plt	0	1
BuchananScore	0	1
leukocytes	0	1
durationDiagnosis	0	1
pid	0	0
gender	0	0
group	0	0
age_years	0	0
plt_month1	1	3
hb	1	2
bodyweight	1	2
plt_week1	1	1
CNR1	2	16
CNR2	2	16
CNR3	2	16
F2A.H131R	2	16
F2A.Q27W	2	16
F3A.V158F	2	16
F2C.Stop	2	16
F2C.ORF	2	16
F3B.HNA1c	2	16
F2B.I232T	2	16
F2C.nc.ORF	2	16
F2B.4	2	16
F2BC.2B.2	2	16
lymphocytes	2	3
MAIPA_IgG_GPIb	3	16
MAIPA_IgG_GPV	3	16
MAIPA_IgG_GPIIbIIIa	3	16
neutrophils	3	5
monocytes	4	5
MAIPA_IgM_GPIIbIIIa	6	19
MAIPA_IgM_GPV	6	19
MAIPA_IgM_GPIbIX	7	20
TPO	8	17
CD4CD8ratio_total	10	22
CD8_total	11	23

Variable	N missing (IVIG)	N missing (Observation)
lympho_total	11	22
CD3_total	11	22
CD4_total	11	22
CD8_relative	11	22
CD3_relative	11	21
CD4_relative	11	21
CD19_relative	11	21
freqTreg	12	26
freqTnCD45-	12	25
freqTcmCD45RO+CD27+	12	25
freqTeffCD45RO+CD27-	12	25
CD19_total	12	22
NK_total	12	22
NK_relative	12	22
Total_IgG	18	33

Supplementary Table 3. Coefficients for age correction of numerical variables. A linear model was used to correct numerical variables for age (in R notation: variable ~ age) to calculate residuals used for downstream modeling. Age was entered in years.

	Intercept	Coefficient for age
hb	7.141	0.100
leukocytes	11.382	-0.584
lymphocytes	6.620	-0.565
neutrophils	4.539	-0.194
monocytes	0.814	-0.044
eosinophils	0.280	0.001
basophils	0.096	-0.013
mpv	12.728	-0.027
TPO	62.323	-1.972
Total_IgG	7.383	0.645
bodyweight	7.170	2.493
durationDiagnosis	4.804	0.226
CRP	1.331	0.196
IP_lympho_total	6.346	-0.554
CD3_total	4.214	-0.332
CD4_total	2.836	-0.294
CD8_total	1.249	-0.053
CD4CD8ratio_total	2.596	-0.140
CD19_total	1.621	-0.196
NK_total	0.525	-0.038
CD3_relative	65.605	1.187
CD4_relative	43.731	-0.673
CD8_relative	18.746	1.449
CD19_relative	26.705	-1.644
NK_relative	7.599	0.367
plt	7.141	-0.080

Supplementary Table 4. Elastic net-derived model characteristics and coefficients for prediction of the IVIg treatment response.

	Model #1	**Model #2**
Elastic net alpha	1.0	0.2
Parameters	5	8
Odds ratio (95% CI)*		
Intercept		
Hemoglobin, mmol/L	0.54 (0.41, 0.70)	0.58 (0.48, 0.70)
Platelet count, x10^9/L	0.70 (0.59, 0.82)	0.70 (0.62, 0.77)
FCGR2C ORF present	0.44 (0.32, 0.59)	0.51 (0.43, 0.61)
MAIPA IgG positive	0.47 (0.30, 0.74)	0.55 (0.42, 0.72)
Preceding vaccination	3.11 (1.87, 5.15)	4.00 (2.71, 5.88)
TPO, U/L		1.36 (1.00, 1.36)
Neutrophils, x10^9/L		0.82 (0.73, 0.91)
FCGR2B p.232 I/I		Reference
FCGR2B p.232 I/T		1.05 (0.89, 1.24)
FCGR2B p.232 T/T		3.54 (2.08, 6.03)
ROC AUC (95% CI)	0.84 (0.75, 0.93)	0.85 (0.76, 0.94)
Nagelkerke R^2	0.428	0.471
Brier Score	0.163	0.151

Model #1 and #2 are models with the optimal ROC AUC as determined by cross-validation. Odds ratios (OR) are provided for the absence of sustained response (ASR; logistic regression with outcome ASR *vs.* complete sustained response). OR for continuous predictors are standardized coefficients (i.e., the OR for the effect of a change in one standard deviation on the scale of the predictor). All continuous variables were adjusted for age as outlined in the Supplementary Methods. The calculation of an individual's score is shown in the Supplementary Methods. *The 95% confidence intervals for the coefficients are reported to illustrate the model variability using bootstrapping, but they are biased due to the use of penalized regression.

Supplementary Table 5. Contingency table to discriminate patients with different IVIg responses.

		Observed response	
		Absent sustained response (ASR)	Complete sustained response (CSR)
Predicted response	ASR	29	16
	CSR	3	32

Molecular risk score with predicted response dichotomized at the optimal Youden's J value (cut-off at -0.54). Sensitivity: 0.91; Specificity: 0.67; Positive predictive value (predict ASR): 0.64; Negative predictive value (predict CSR): 0.91

Supplementary Table 6. Model accuracy at various score cut-offs.

	Sensitivity maximized	*Best Youden's J*	*Specificity maximized*
Model 1			
Threshold	- 1.00	- 0.54	0.76
Specificity	0.46	0.67	0.98
Sensitivity	0.97	0.91	0.25
PPV	0.54	0.64	0.89
NPV	0.96	0.91	0.66
Model 2			
Threshold	- 1.05	- 0.15	0.56
Specificity	0.33	0.83	0.98
Sensitivity	0.97	0.75	0.41
PPV	0.49	0.75	0.93
NPV	0.94	0.83	0.71

Score thresholds as determined by the optimal threshold values using *pROC.* ROC curves are presented in Supplementary Figure 3.

PPV, positive predictive value; NPV, negative predictive value.

Supplementary Table 7. Proportion of patients with bleeding events during one-year follow-up, stratified by the molecular risk score.

	IVIg		Observation	
	P(ASR) ≤ 0.5	**P(ASR) > 0.5**	**P(ASR) ≤ 0.5**	**P(ASR) > 0.5**
At least extensive petechiae and/or large bruises (Buchanan Score >1)				
Week 1	0.17	0.25	0.32	0.60
Month 1	0.12	0.32	0.19	0.27
Month 3	0.10	0.21	0.03	0.23
Month 6	0.06	0.18	0.08	0.13
Month 12	0.04	0.07	0.00	0.07
At least mucosal bleeding (Buchanan Score >2)				
Week 1	0.02	0.11	0.08	0.30
Month 1	0.04	0.14	0.16	0.13
Month 3	0.06	0.07	0.03	0.13
Month 6	0.06	0.11	0.05	0.07
Month 12	0.02	0.04	0.00	0.03

Model #1 score. Data are the proportion of patients experiencing the indicated modified Buchanan Score (Buchanan J Pediatr 2002; Bennett Blood 2006).

Supplementary Table 8. Proportion of patients in observation group exhibiting complete recovery during one-year follow-up, stratified by both a clinical prediction score and the and molecular risk score.

	High Probability of Recovery		Intermediate Probability		Low Probability	
	P(ASR) ≤ 0.5	**P(ASR) > 0.5**	**P(ASR) ≤ 0.5**	**P(ASR) > 0.5**	**P(ASR) ≤ 0.5**	**P(ASR) > 0.5**
Month 3	0.87 (20/23)	0.68 (13/19)	0.80 (8/10)	0.38 (3/8)	0/2	0/2
Month 6	0.96 (22/23)	0.84 (16/19)	0.91 (10/11)	0.62 (5/8)	1/2	0/2
Month 12	1.00 (23/23)	0.95 (18/19)	1.00 (11/11)	0.75 (6/8)	2/2	0/2

Data are the proportion (n/N) of patients exhibiting a complete recovery at the indicated timepoint. Shown is the probability of recovery by the clinical prediction score (NOPHO high/intermediate/low) and molecular score (P[ASR]). The probability of recovery is indicated by NOPHO groups, high (10-14), intermediate (5-9), low (0-4). For two patients the NOPHO score could not be calculated due to missing data.