Annals of Hematology (2019) 98:2283–2291
https://doi.org/10.1007/s00277-019-03770-3

ORIGINAL ARTICLE

Prediction of thrombosis risk in patients with paroxysmal nocturnal hemoglobinuria

Yuzhou Huang[1] · Xinjian Liu[2] · Fangfei Chen[1] · Wenzhe Zhou[1] · Hongmin Li[1] · Zhangbiao Long[1] · Chen Yang[1] · Miao Chen[1] · Bing Han[1]

Received: 29 November 2018 / Accepted: 29 July 2019 / Published online: 8 August 2019

Abstract

Paroxysmal nocturnal hemoglobinuria (PNH) is an acquired hemolytic disease with thrombosis as a major complication. The mechanism of thrombosis and related risk factors in PNH patients are still not well characterized. We retrospectively enrolled 99 patients with newly diagnosed PNH at our institute from 2011 to 2016. According to binary logistic regression model analysis, we first identified four baseline clinical risk factors which may be associated with incidence of thrombosis in the PNH cohort, including PNH clone sizes (fluorescent aerolysin of neutrophil) ≤ 80 (OR 1.056, 95%CI 1.016–1.097, $P = 0.005$), hemoglobin ≤ 75 g/L (OR 4.202, 95%CI 0.984–17.954, $P = 0.053$), platelet > 100×10^9/L (OR 6.547, 95%CI 1.490–28.767, $P = 0.013$) and *rs495828* = G (OR 5.243, 95%CI 1.314–20.916, $P = 0.019$). These independent risk factors were combined together to develop a risk model to evaluate thrombosis risk (AUC = 0.756, 95%CI 0.607–0.905, $P < 0.001$). Our risk model revealed a higher cumulative incidence of thrombosis and an earlier thrombosis events in PNH patients with high risk (risk score ≥ 23) compared with those with low risk (risk score < 23, $P < 0.001$ and $P = 0.043$, respectively). Although with some limitations, we set up a prediction model for thrombosis risk in patients with PNH for the first time, but it needed to be verified in a prospective study with larger patients and longer follow-up time in the future.

Keywords Paroxysmal nocturnal hemoglobinuria (PNH) · Thrombosis · Risk score · Prediction model

Introduction

Paroxysmal nocturnal hemoglobinuria (PNH) is a hemolytic disease caused by a loss of the glycosylphosphatidylinositol (GPI) anchor protein due to mutation of the *PIGA* gene in hematopoietic stem cells, which is characterized with thrombosis as one of the main complications [1–3]. A global PNH patient registration survey found that the proportion of thrombosis in 1610 PNH patients was 16% [4], besides, the incidence of thrombosis in PNH patients is 10.61/100 per year, which is much higher than that of the general population (0.75~2.69/1000 per year) [4–7]. Forty to 67% of PNH death was related with thrombosis, indicating the magnificent influence of this complication on the quality of life for PNH patients [6, 8–11].

The mechanism of high incidence of PNH thrombosis is not fully understood. It is believed that excessive consumption of nitric oxide (NO) and abnormal activation of platelets caused by repeated hemolysis are the main causes of thrombosis in PNH patients [3, 6, 12]. Studies have shown that the size of PNH clones is associated with the risk of thrombosis: A 10% increase in PNH clones can bring a 1.64-fold increase in thrombosis risk [13]. Even though, some PNH patients with small PNH clone size can also occur thrombosis, suggesting that clone size is not the only factor affecting thrombosis [4]. Our previous study also found that patients with ABO gene *rs495828* and *rs2519093* mutations are prone to develop thrombosis, which is a risk factor independent of PNH clone size [14].

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00277-019-03770-3) contains supplementary material, which is available to authorized users.

✉ Bing Han
hanbing_li@sina.com.cn

[1] Department of Hematology, Peking Union Hospital, CAMS & PUMC, Beijing 100730, China

[2] Department of Hematology, Henan Cancer Hospital, Institute of Hematology of Henan Province, Affiliated Cancer Hospital of Zhengzhou University, 127 Dongming Road, Zhengzhou, China

This study mainly aimed to set up a thrombotic prediction system by retrospectively analysis of all-known risk factors possibly associated with thrombosis and separate patients prone to thrombosis events by different risk scores, which may help the decision of prophylactic use of anticoagulation in high-risk patients in the future.

Patients and methods

Patients

We did a retrospective search of medical records to identify patients with diagnosis of PNH in Peking Union Medical College Hospital (PUMCH) from October 2011 to October 2016. All patients planned to be included in study had a clear diagnosis of PNH, stored with blood specimen at the time of diagnosis and had been followed up regularly in PUMCH. The patients with unclear diagnosis, uncomplete clinical record or unqualified blood specimen, or losing contact information were excluded from the study. The diagnosis was confirmed according to the international diagnostic criteria for PNH [15, 16].

The detailed clinical features at diagnosis including sex, age, medical history, and physical examination were obtained from clinical record. Laboratory data including routine blood test, liver and kidney function test, coagulation test, and flow cytometry for PNH clone were tested and collected at diagnosis and follow-up time. PNH clone was assayed by CD55/CD59 levels of RBC/neutrophil and FLAER (fluorescent aerolysin) negative percentage of neutrophil/monocytes.

According to our previous work for prothrombotic mutations, we selected SNP *rs495828*/*rs2519093* in ABO gene [14], which were identified as the independent risk factors for thrombosis to represent for prothrombotic gene in the following analysis. Blood samples at the time of diagnosis were thawing. DNA was extracted through DNA purification kits (Qiangen, Hilden, Germany), and DNA concentration was measured by spectrometry. SNP *rs495828*/*rs2519093* was genotyped by Sequenom MassARRAY system (Sequenom, Inc., San Diego, CA, USA) according to relevant protocol. The primers used to sequence *rs495828* and *rs2519093* were *TTGCAAGGGAGGTAAA* and *cccccTAAGCCACCGACTGA*, respectively.

Patients who had been followed up regularly were recorded the clinical course or complications. Clinical symptoms and laboratory results were found from medical record or obtained by telephone or mail. The thrombosis-related information, including starting time, location of thrombosis events, number of thrombosis, treatment, and clinical outcome, were documented. All thrombosis events were confirmed by ultrasound and/or angiography. The terminal of follow-up was death, diagnosis with thrombosis, or reach the endpoint time of October 2018. Each patient enrolled was ensured to have at least 2-year follow-up time. The medium follow-up time was 48 months (24–91 months). No patients enrolled lost to follow-up before terminal time. Finally, 99 patients with relatively complete data were included in the model establishment.

Informed consents were obtained from all patients before the study start, and the study was in accordance with the Declaration of Helsinki and was approved by the Peking Union Medical College Hospital Ethics Committee.

Modeling methods

Our thrombosis risk scoring system mainly referred to the Framingham coronary heart disease risk model [17]. The possible related risk factors were analyzed by multivariate binary logistic regression method, and the risk score was evaluated as follows:

a) We first selected the risk factors either statistically significant related to thrombosis formation in the univariate analysis in this study, or recognized by previous evidence-based medicine as possible meaningful clinical predictors.
b) Construct a multivariate binary logistic regression equation to find out the relevant risk factors for thrombosis and determine the regression parameters β and OR values.
c) Classify continuous variables in the regression equation. Take the PNH clone size as an example: Since there was a significant difference in the incidence of thrombosis between patients with PNH clones ≤ 80% and those with > 80%, but no difference in other cut-off point, we classified the PNH clones as 5 categories: "PNH clones ≤ 80%," "81–85%," "86–90%," "91–95%," and "96–100%." Those with PNH clone "≤ 80%" was taken the mean value of 41%, and patients with clone size "81–85%" was calculated as the mean value of 83%.
d) Set up the basic value for the risk parameters, for example, patients with PNH clone "≤ 80%" as the basic risk factor value. For a two-category variable, the basic risk value was considered as 0.
e) Determine the distance value from the basic value in each category by using $D_{ij} = \beta_i(W_{ij} - W_{iREF})$ in regression units (β_i is the coefficient of the variable no. i in the multivariate binary logistic regression model, W_{ij} is the value of category j for the risk factor i; W_{iREF} is the baseline value of risk factor i).
f) Set the constant B for the integrating system. The constant had to give the consideration of both the risk assessment time period and the refer risk factor itself. Take PNH clone as an example, constant B had to consider both the thrombosis in 5 years and the size of PNH clone, so we set the constant B as $5 \times \beta_{\text{PNH clone}}$.

g) Determine the score of each category for each risk factor by Points $ij = D_{ij}/B$. Risk scores are generally rounded up for clinical facilities.
h) Determine the correspondence between the total score of risk factors and the probability of future risk of thrombosis by $p = \frac{1}{1+\exp\left(-\sum_{i=0}^{n}\beta_i X_i\right)}$, and $\sum_{i=0}^{n}\beta_i X_i \approx \sum_{i=0}^{n}\beta_i X_{iREF} + B \times P_{total}$.
i) Stratify the risk level according to combination between groups with similar trends of thrombosis incidence.

Data analysis

All data analysis involved in this study was performed by SPSS 22.0 software. Continuous variables were described using standard summary statistics, and categorical variable was shown as frequency and percentage. Group differences were determined by Student's *t* test for continuous variables with normal distribution, and two-sided Wilcoxcon rank sum test for continuous variables with non-normal distribution. Pearson chi-square or Fisher's exact test was used for categorical variables. The group differences of two ordered categorical variables were determined by Kruskal-Wallis test. Independent risk factors were analyzed by LR (logistic regression) model of multivariate binary logistic regression. The discrimination ability of the regression model was determined by ROC curve, and the calibration ability was determined by Hosmer-Lemeshow test. The scoring system was established by multivariate binary logistic regression equation [17]. The effect of scoring on the cumulative incidence of thrombotic events was achieved by Kaplan-Meier function and Log-rank test, and the hazard ratio was calculated by Mantel-Haenszel test. A $P < 0.05$ was considered as statistically significant.

Results

Patients characteristics

The clinical characteristics of 99 patients with PNH are shown in Table 1. For all patients, the male to female ratio was 1.25:1, and the median age at diagnosis was 40 (18–78) years old. In total, 40.4% of patients were diagnosed with PNH/AA (aplastic anemia). During the follow-up period, 17 (17.2%) patients with PNH developed thrombosis events. Among them, 3 patients had arterial thrombosis (1 case with myocardial infarction combining with cerebral infarction, 1 case with pulmonary embolism combining with cerebral infarction, and 1 case with coronary thrombosis); 16 patients had venous thrombosis (10 cases of abdominal venous thrombosis, 3 cases of intracranial thrombosis, 2 cases of lower extremity venous thrombosis, and 1 case of thrombophlebitis; 2 cases had both venous and arterial thrombosis). Detailed information of 17 PNH patients with thrombosis was described in Online Resource 1. Four patients had multiple thrombi in multiple sites, and one patient had multiple thrombi at the same site. The median time to thrombotic events was 12 m (1–67 m) after diagnosis of PNH, and 10 of them (58.8%) developed thrombosis within 12 months after diagnosis. Four patients died during the follow-up period: two died of thrombosis, one died of infection, and one of colon cancer, respectively.

The establishment of thrombosis risk scoring system

Table 2 shows the risk factors that may be associated with thrombosis in our study. We selected variables with statistical significance from previous studies [4, 14, 18] or possibly clinical relevance according to previous evidence-based medicine [11, 19–21]. The PNH clone size (FLAER-% of PMN) and the presence of the thrombogenic gene (GENE, −*rs495828 GT* or *T* genotype) were significantly related to the thrombosis formation ($P = 0.019$ and 0.003, respectively). Because FLAER-% is the most important marker for clone size [15], to make it simple, in our following steps, we took FLAER-% of PMN as the risk factor for PNH clone size. Due to the genetic linkage relationship of *rs495828* and *rs2519093*, we took the *rs*495828 (*GT* or *T*) as the representation for thrombogenic gene in this study. Although not significant, platelet count (PLT), age (AGE), hemoglobin (HGB), and PC and PS level were also considered as potential risk factors in the following analysis because they could be clinically relevant due to small *P* value ($P < 0.3$) in this study.

Then, we used those possible risk factors to do multivariate binary logistic regression analysis. Finally, PLT ($Plt > 100 \times 10^9$/L, OR 6.547, 95%CI 1.490–28.767, $P = 0.013$), HGB ($Hgb \leq 75$ g/L, OR 4.202, 95%CI 0.984–17.954, $P = 0.053$), PNH clone (FLAER-% of PMN, OR 1.056, 95%CI 1.016–1.097, $P = 0.005$), and GENE (*rs495828* = G, OR 5.243, 95%CI 1.314–20.916, $P = 0.019$; Table 3) were shown to be the independent risk factors for thrombosis. We selected "PNH clone ≤ 80%, Hg ≤ 75 g/L, $Plt > 100 \times 10^9$/L, *rs2519093 G*" as the referent baseline risk variables for assignment analysis (Online Resource 2). The final risk score model and estimated risk levels are shown in Table 4. The study cohort could be divided into two categories (low risk, score < 23; high risk, score ≥ 23; Online Resources 3, 4). At the cutoff value for high risk (risk score ≥ 23), the risk score system had a positive predictive value of 62.5%, a negative predictive value of 91.25%, and a specificity of 92.4% in our patient cohort. Meanwhile, the ROC curve showed AUC = 0.756 (95%CI 0.607–0.905, $P < 0.001$) and the Hosmer-Lemeshow test showed $X^2 = 1.166$ ($P =$

Table 1 Clinical characters of the 99 patients enrolled

Characteristic	Whole cohort	With thrombosis	Without thrombosis	*P* value
Cases, *n* (%)	99	17 (17.2%)	82 (82.8%)	
Age (years)	40 (18–78)	31 (24–55)	39 (13–73)	0.250
Male gender, *n* (%)	55 (55.6%)	10 (58.8%)	45 (54.9%)	0.766
PNH/AA, *n* (%)	40 (40.4%)	10 (58.5%)	30 (36.6%)	
RET%	4.84 ± 3.32	4.16 ± 1.97	5.01 ± 3.56	0.622
RBC ($\times 10^{12}$/L)	2.55 ± 0.80	2.67 ± 0.70	2.52 ± 0.82	0.478
WBC ($\times 10^{9}$/L)	3.66 (1.02–9.48)	3.40 (1.65–9.90)	3.71 (1.02–9.96)	0.868
Hgb (g/L)	79 (18–149)	80 (44–112)	79 (18–149)	0.878
Plt ($\times 10^{9}$/L)	81.0 (5.0–340.0)	54.0 (22.0–340.0)	95.5 (5.0–331.0)	0.179
LDH (U/L)	1135.5 (105.0–3950.0)	1163.0 (206.0–2930.0)	1123.00 (105.0–3950.0)	0.592
D-Dimer (mg/L)	0.33 (0.11–4.03)	0.32 (0.12–3.40)	0.34 (0.11–4.03)	0.908
PC (%)	111.52 ± 21.30	110.13 ± 27.21	111.80 ± 20.10	0.776
PS (%)	106.484 ± 20.01	101.44 ± 25.71	107.51 ± 18.69	0.223
FLAER-monocytes (%)	84.6 (2.0–99.0)	91.0 (61.0–99.0)	76.0 (2.0–99.0)	0.003
FLAER-PMN (%)	84.5 (2.0–99.0)	92.0 (61.0–99.0)	78.0 (2.0–99.0)	0.004
CD59-PMN (%)	85.0 (1.0–99.0)	91.00 (61.0–99.0)	74.00 (1.0–99.0)	0.003
rs495828---GT, *n* (%)	36 (36.4%)	11 (64.7%)	25 (30.5%)	0.003
rs2519093---TC, *n* (%)	36 (36.4%)	12 (70.6%)	24 (29.3%)	0.002

Data normally distributed were shown as the mean ± standard deviation; data non-normally distributed were shown as the median (range); *P* value represented the difference between patients with and without thrombosis

PNH paroxysmal nocturnal hemoglobinuria, *AA* aplastic anemia, *FLAER* fluorescent aerolysin, *PMN* polymorphonuclear, *PC* protein C, *PS* protein S, *PMN* polymorphonuclear

0.468 > 0.05), indicating the good discrimination and calibration ability of the regression model (Fig. 1).

Thrombosis incidence analysis in study cohort

In our study cohort, 1-year and 3-year overall cumulative incidence of thrombosis were 13.1% and 16. 8%, respectively (Fig. 2a). According to our risk stratification, high-risk group had much higher risk score than low-risk group (25.5 vs. 12.3, $P < 0.001$), and the true incidence of thrombosis during follow-up period was 62.5% and 8.8% for those two risk groups, respectively ($P = 0.001$). We did Kaplan-Meier analysis to evaluate cumulative incidence of thrombosis for different risk levels. Patients with high risk had higher cumulative incidence of thrombotic events compared with those with low risk ($P < 0.001$, Fig. 2b). Detail analysis showed that the high-risk group had significantly higher cumulative incidence of thrombosis at 12 months (50.0 vs 2.5%, $P < 0.01$) and 36 months (65.0% vs 7.9%, $P < 0.01$) after diagnosis than low-risk group. Meanwhile, the median time for thrombosis onset after diagnosis for patients with high risk and low risk was 8.5 m (1–36 m) and 18 m (5–67 m), respectively ($P = 0.043$, Fig. 2c). For patients who developed thrombosis, 80.0% and 28.6% of patients had thrombotic events within the first year after diagnosis for high-risk and low-risk patients, respectively ($P = 0.040$).

Discussion

PNH is one of the common thrombophilia diseases [1, 4, 8], and thrombosis is also a common cause of death [10, 11]. The incidence of thrombosis in Chinese patients is lower, but the predilection sites are similar to the West, as the venous thrombosis, especially abdominal veins thrombosis the most common ones [9, 22–24]. In addition, 29 to 44% of patients with PNH have had at least one thrombus during the progression of the disease, and more than 20% of patients have multiple sites of thrombosis [4]. Even though, anticoagulation therapy is not commonly used in China due to the high incidence of bleeding and thrombocytopenia in PNH. More accurate and effective evaluation for thrombotic risk is necessary.

In addition to PNH clone size [4, 8], which is widely accepted as the most important risk factor for thrombosis, our previous study had shown that patients with *rs495828* locus of ABO gene exon $G > T$ mutation and $C > T$ mutation of the *rs2519093* locus for ABO intron gene have 3.03 and 4.24 times higher rate of thrombosis than those of the non-

Table 2 The categories of clinical risk factors associated with thrombosis in PNH patients

Risk factors	Categories	Thrombi events	*P* value
HGB (Hgb, g/L)	≤ 75	5 (12.50%)	0.278
	> 75	12 (21.50%)	
PLT (Plt, × 10^9/L)	> 100	14 (7.10%)	0.019
	≤ 100	3 (25.50%)	
GENE (*rs495828*)	*G*	5 (8.20%)	0.003
	GT or *T*	13 (31.60%)	
AGE	≤ 40	12 (22.22%)	0.146
	> 40	5 (11.11%)	
LDH	> 1.5ULN	3 (16.70%)	0.789
	≤ 1.5ULN	14 (19.40%)	
D-DIMER (mg/L)	≤ 0.4	12 (22.20%)	0.485
	> 0.4	6 (16.20%)	
WBC (× 10^9/L)	≤ 4.0	11 (19.6%)	0.597
	> 4.0	6 (14.6%)	
RBC (× 10^{12}/L)	≤ 3.5	15 (17.9%)	0.828
	> 3.5	2 (15.4%)	
RET%	≤ 5	12 (21.8%)	0.527
	> 5	5 (16.1%)	
PC%	≤ 80	2 (40%)	0.157
	> 80	14 (15.6%)	
PS%	≤ 85	4 (30.8%)	0.151
	> 85	12 (14.6%)	

P value represented the difference between patients with different risk factor categories

HGB hemoglobin, *PLT* platelet, *GENE rs495828*, *AGE* age, *LDH* lactate dehydrogenase, *D-DIMER* d-dimer, *WBC* white blood cell, *RBC* red blood cell, *RET%* percent of reticulocyte, *PC* protein C, *PS* protein S, *ULN* upper limit of normal

mutated patients, respectively [14]. Other studies showed that LDH > 1.5ULN (upper limit of normal) is an independent risk factor [4, 18]. While factors like ethnicity, D-dimer, age over 55 years, and PNH type were described as risk factors in some study [11, 19, 20], but were not verified by others. Some factors like PC and PS, which were closely associated with thrombosis [21, 25, 26], were also considered. Finally, all the variables possibly associated with thrombosis taken either from the literatures or clinical experience were selected as the primary factors for analysis in our study.

Table 3 Multivariate binary logistic regression analysis of risk factors associated with thrombosis in PNH patients

Risk factor	β_{ij}	*P*	OR	95%CI for OR
PNH clone (%)	0.054	0.005	1.056	1.016–1.097
Hgb ≤ 75 g/L	1.436	0.053	4.202	0.984–17.954
Plt > 100 × 10^9/L	1.879	0.013	6.547	1.490–28.767
rs495828 = G	1.657	0.019	5.243	1.314–20.916
Intercept	− 8.721	> 0.001		

OR odds ratio, *CI* confidence interval, *PNH clone* FLAER of PMN, β coefficient

In the univariate analysis, we found that the size of PNH clones and the presence of both *G* > *T* mutation of rs495828 locus and *C* > *T* mutation of rs2519093 locus of ABO gene exon were significantly different between patients with thrombosis and those without (*P* = 0.019, 0.003, and 0.002, respectively). Due to the genetic linkage between *rs2519093* and *rs495828* [14], we only selected *rs495828* as the gene indicator in the following analysis. Other factors like platelet count, age, hemoglobin, and PC/PS may also be considered for following logistic analysis due to small *P* value (*P* < 0.3, Table 2). By multivariate binary logistic analysis, PNH clone sizes ≤ 80, hemoglobin ≤ 75 g/L, platelet > 100 × 10^9/L, and gene mutation (*rs495828* = G) were considered as independent risk factors for thrombosis in PNH patients. Though lysis of platelets occurs and contributes to thrombocytopenia in some PNH patients, aberrant activation of platelets resulting from hemolysis and MAC (membrane attack complex) deposition contributes to a prothrombotic state and initiates clot formation in PNH patients [27–29]. Hemoglobin was another risk factor included in our risk model. Though its *P* value was over 0.05 in logistic regression, SPSS software automatically recognized it as independent risk factor due to its *P* value closed enough to significance. Though patients with higher hemoglobin appeared to have a lower degree of hemolysis, a high blood viscosity at baseline time was easier to cause platelet aggregation and adhesion later on. Combined together, high platelet count and high blood viscosity were prone to thrombosis for PNH patients in our study, which is in accordance with clinical experience. Although age is a common risk factor for thrombosis in other diseases, it was excluded from our risk model, probably because that most of PNH patients suffered from thrombosis at a younger age in our study. PC/PS deficiency were very common in many thrombosis diseases [26], but not a dependent risk factor in our risk model. Our previous study had also found no difference in the expression of PC/PS between thrombotic and non-thrombotic PNH patients [14, 30]. Parameters like LDH and D-dimer did not show significant difference between patients with or without thrombosis, probably because the data we collected was from the time of initial diagnosis, not the time when the thrombosis happened.

We set up the risk model based on the methods from Framingham coronary heart disease risk model [17], which was well accepted and widely used for many non-heart diseases [31]. The model also works for our patients: high-risk patients (score ≥ 23) and low-risk patients (score < 23) was well discriminated (ROC AUC = 0.756, 95%CI 0.607–0.905, *P* < 0.001). Patients scored as high risk were much easier to have thrombosis with a higher cumulative incidence of thrombosis at 1 year and at

Table 4 Risk score system to predict thrombosis risk in PNH patients

Risk factor	Categories	Points	Point total	Estimated risk	Risk level
PNH clone (%)	≤ 80	0	0–22	< 40%	Low risk
	81–85	8			
	86–90	9			
	91–95	10			
	96–100	11			
HGB (g/L)	≤ 75	0			
	> 75	5	23–29	> 40%	High risk
PLT (× 10^9/L)	> 100	0			
	≤ 100	7			
GENE	*G*	0			
	GT/T	6			

HGB hemoglobin, *PLT* platelet, *GENE rs495828*, *PNH clone* FLAER of PMN

3 years after diagnosis. Meanwhile, they had an earlier onset of thrombosis. Previous study reported a score system to evaluate probability of death in PNH patients according to sites of occlusion [9], suggesting sites of thrombosis influencing prognosis. However, there was no clear relationship between the location of the thrombi and the risk score in our risk model. Although the model we set up was based on a retrospective analysis and needed to be validated in a prospective trial, it was tested to have the ability to classify patients for their thrombotic risk at the time of diagnosis. It may be suggestive and practical in clinical use. For future patients, all the relevant prediction factors found in our study should be examined at the time of diagnosis and patients should be scored after diagnosis workup. For those with high risk (score ≥ 23), a more close observation of thrombotic events in the follow-up, or even suggestion for prophylactic anticoagulation or eculizumab [8, 32] if needed, may be recommended. Of course, all patients with different scores should be documented for their frequency and time

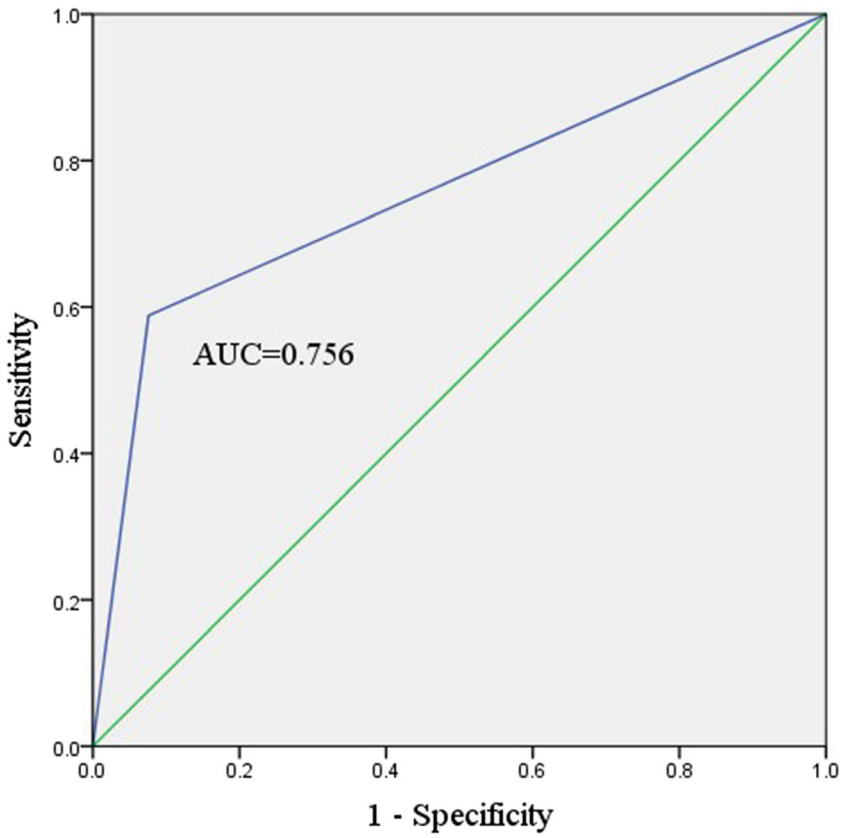

Fig. 1 ROC curve for predicting incidence of thrombosis after diagnosis. Discrimination with the use of predicting incidence of thrombosis was shown by ROC curve and AUC = 0.756 (95%CI 0.607–0.905, $P < 0.001$). ROC curve was completed by SPSS 22.0

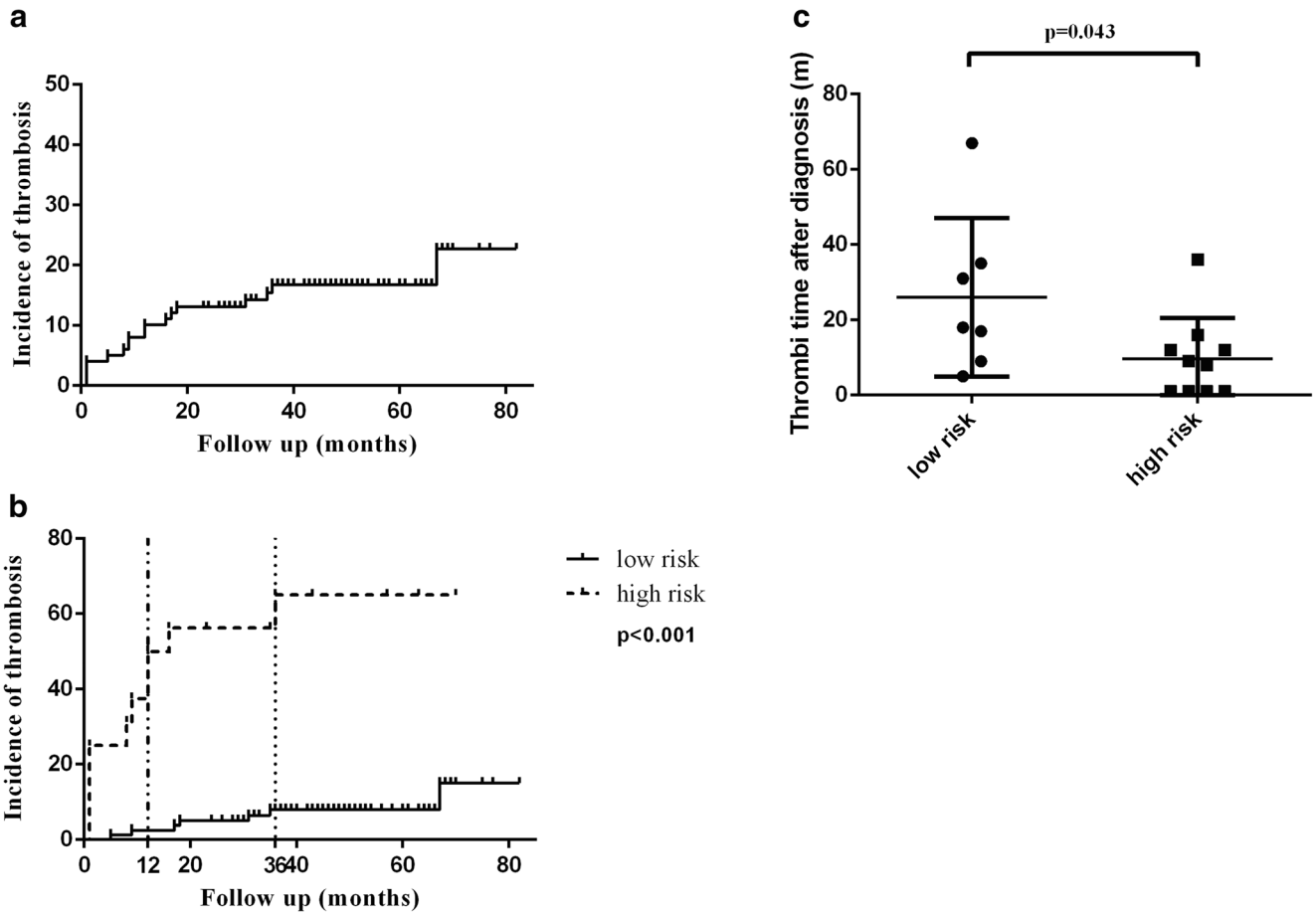

Fig. 2 Thrombosis incidence analysis in study cohort. **a**, **b** The cumulative incidence of whole cohort and different risk groups, respectively. The vertical dot line showed the time points of the 12th month and 36th month after the diagnosis of PNH. **c** The onset time of thrombosis for PNH patients after diagnosis (high risk 8.5 m vs. low risk 18 m, $P = 0.043$)

of thrombosis onset after diagnosis and follow-up for longer time to further test the model.

To the best of our knowledge, this is the first PNH model for thrombosis prediction so far. Most of the parameters we used were easy to access, and the scoring can be calculated automatically after value input. However, our study had some limitations. First, this was a retrospective study. Some patients did not have sufficient laboratory results, and evaluation of thrombosis were prone to variation. Second, due to the limited knowledge of thrombotic risk factors, the prothrombotic factors enrolled in this study were probably insufficient, i.e., there is not much convincing data on the thrombotic genes for PNH so far; we only took the genes which had been verified to be thrombotic related in our previous study [14, 33]. Third, the number of PNH patients, especially those with thrombosis, was relatively small, which may cause the bias and uncertainty of the model. However, PNH is a rare disease [22], and it is extremely difficult to collect patients in such a number with clinical data integrated enough for the analysis. Finally, the follow-up time was limited. It is known that patients may develop thrombosis later in their follow-up and may be considered as non-thrombotic patients at the time of analysis. Although we tried to overcome the defect by following up each patient for at least 2 years, longer follow-up time is still needed.

With the limitations mentioned above, future work will focus on finding new related markers and verifying the current model in a well-designed, prospective study with larger PNH cohort and longer follow-up time.

Authors' contribution Y.H. and B.H. designed the study and wrote the manuscript. Y.H. analyzed clinical and experimental data. X.L. helped analyzed the risk model. C.Y. and M.C. helped to design the study. Z.L. and W.Z. collected the samples, performed genotyping, and helped to analyze the data. H.L. and F.C. helped to perform genotyping. All authors reviewed the manuscript finally. Y.H., M.C., and B.H. approved the final submission of the manuscript.

Funding information This study was supported by grants from the Chinese Academy of Medical Sciences (CAMS) innovation fund for medical sciences (2016-I2M-3-004), The National Key Research and Development Program of China (2016YFC0901500), and Beijing Natural Science Foundation (7192168).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Research involving human participants All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study was approved by the PUMCH Ethics Committee.

Informed consent Informed consent was obtained before study start from all individuals included in the study.

References

1. Hill A, DeZern AE, Kinoshita T, Brodsky RA (2017) Paroxysmal nocturnal haemoglobinuria. Nat Rev Dis Primers 3:17028. https://doi.org/10.1038/nrdp.2017.28
2. Mohammed AA, El-Tanni H, Atiah TA, Atiah AA, Atiah MA, Rasmy AA (2016) Paroxysmal nocturnal hemoglobinuria: from bench to bed. Indian J Hematol Blood Transfus 32(4):383–391. https://doi.org/10.1007/s12288-016-0654-2
3. Boyce S, Eren E, Lwaleed BA, Kazmi RS (2015) The activation of complement and its role in the pathogenesis of thromboembolism. Semin Thromb Hemost 41(6):665–672. https://doi.org/10.1055/s-0035-1556732
4. Schrezenmeier H, Muus P, Socie G, Szer J, Urbano-Ispizua A, Maciejewski JP, Brodsky RA, Bessler M, Kanakura Y, Rosse W, Khursigara G, Bedrosian C, Hillmen P (2014) Baseline characteristics and disease burden in patients in the international paroxysmal nocturnal hemoglobinuria registry. Haematologica 99(5):922–929. https://doi.org/10.3324/haematol.2013.093161
5. Day ISCWT (2014) Thrombosis: a major contributor to global disease burden. Thromb Res 134(5):931–938. https://doi.org/10.1016/j.thromres.2014.08.014
6. Hill A, Kelly RJ, Hillmen P (2013) Thrombosis in paroxysmal nocturnal hemoglobinuria. Blood 121(25):4985–4996. https://doi.org/10.1182/blood-2012-09-311381
7. DeZern AE, Brodsky RA (2015) Paroxysmal nocturnal hemoglobinuria: a complement-mediated hemolytic anemia. Hematol Oncol Clin North Am 29(3):479–494. https://doi.org/10.1016/j.hoc.2015.01.005
8. Griffin M, Munir T (2017) Management of thrombosis in paroxysmal nocturnal hemoglobinuria: a clinician's guide. Ther Adv Hematol 8(3):119–126. https://doi.org/10.1177/2040620716681748
9. Ziakas PD, Poulou LS, Rokas GI, Bartzoudis D, Voulgarelis M (2007) Thrombosis in paroxysmal nocturnal hemoglobinuria: sites, risks, outcome. An overview. J Thromb Haemost 5(3):642–645. https://doi.org/10.1111/j.1538-7836.2007.02379.x
10. Hillmen P, Muus P, Duhrsen U, Risitano AM, Schubert J, Luzzatto L, Schrezenmeier H, Szer J, Brodsky RA, Hill A, Socie G, Bessler M, Rollins SA, Bell L, Rother RP, Young NS (2007) Effect of the complement inhibitor eculizumab on thromboembolism in patients with paroxysmal nocturnal hemoglobinuria. Blood 110(12):4123–4128. https://doi.org/10.1182/blood-2007-06-095646
11. de Latour RP, Mary JY, Salanoubat C, Terriou L, Etienne G, Mohty M, Roth S, De GS, Maury S, Cahn JY (2008) Paroxysmal nocturnal hemoglobinuria: natural history of disease subcategories. Blood 112(8):3099–3106
12. Brodsky RA (2014) Paroxysmal nocturnal hemoglobinuria. Blood 124(18):2804–2811. https://doi.org/10.1182/blood-2014-02-522128
13. Hoekstra J, Leebeek FW, Plessier A, Raffa S, Darwish Murad S, Heller J, Hadengue A, Chagneau C, Elias E, Primignani M, Garcia-Pagan JC, Valla DC, Janssen HL (2009) European network for vascular disorders of the L (2009) paroxysmal nocturnal hemoglobinuria in Budd-Chiari syndrome: findings from a cohort study. J Hepatol 51(4):696–706. https://doi.org/10.1016/j.jhep.2009.06.019
14. Long Z, Du Y, Li H, Han B (2017) Polymorphism of the ABO gene associate with thrombosis risk in patients with paroxysmal nocturnal hemoglobinuria. Oncotarget 8(54):92411–92419. https://doi.org/10.18632/oncotarget.21361
15. Dezern AE, Borowitz MJ (2018) ICCS/ESCCA consensus guidelines to detect GPI-deficient cells in paroxysmal nocturnal hemoglobinuria (PNH) and related disorders part 1 - clinical utility. Cytometry B Clin Cytom 94(1):16–22. https://doi.org/10.1002/cyto.b.21608
16. Borowitz MJ, Craig FE, Digiuseppe JA, Illingworth AJ, Rosse W, Sutherland DR, Wittwer CT, Richards SJ, Clinical Cytometry S (2010) Guidelines for the diagnosis and monitoring of paroxysmal nocturnal hemoglobinuria and related disorders by flow cytometry. Cytometry B Clin Cytom 78(4):211–230. https://doi.org/10.1002/cyto.b.20525
17. Sullivan LM, Massaro JM, Sr D'Agostino RB (2004) Presentation of multivariate data for clinical use: the Framingham study risk score functions. Stat Med 23(10):1631–1660. https://doi.org/10.1002/sim.1742
18. Lee JW, Jang JH, Kim JS, Yoon SS, Lee JH, Kim YK, Jo DY, Chung J, Sohn SK (2013) Clinical signs and symptoms associated with increased risk for thrombosis in patients with paroxysmal nocturnal hemoglobinuria from a Korean registry. Int J Hematol 97(6):749–757. https://doi.org/10.1007/s12185-013-1346-4
19. Nishimura J, Kanakura Y, Ware RE, Shichishima T, Nakakuma H, Ninomiya H, Decastro CM, Hall S, Kanamaru A, Sullivan KM, Mizoguchi H, Omine M, Kinoshita T, Rosse WF (2004) Clinical course and flow cytometric analysis of paroxysmal nocturnal hemoglobinuria in the United States and Japan. Medicine 83(3):193–207
20. Helley D, de Latour RP, Porcher R, Rodrigues CA, Galy-Fauroux I, Matheron J, Duval A, Schved JF, Fischer AM, Socie G, French Society of H (2010) Evaluation of hemostasis and endothelial function in patients with paroxysmal nocturnal hemoglobinuria receiving eculizumab. Haematologica 95(4):574–581. https://doi.org/10.3324/haematol.2009.016121
21. Kinoshita S, Iida H, Inoue S, Watanabe K, Kurihara M, Wada Y, Tsuda H, Kang D, Hamasaki N (2005) Protein S and protein C gene mutations in Japanese deep vein thrombosis patients. Clin Biochem 38(10):908–915. https://doi.org/10.1016/j.clinbiochem.2005.05.006
22. Yu F, Du Y, Han B (2016) A comparative analysis of clinical characteristics of patients with paroxysmal nocturnal hemoglobinuria between Asia and Europe/America. Int J Hematol 103(6):649–654. https://doi.org/10.1007/s12185-016-1995-1
23. Zou N, Han B, Cai H, Xu Y, Wang X, Li RS, Shen T (2012) Clinical features of 76 Chinese patients with paroxysmal nocturnal haemoglobinuria. Zhonghua xue ye xue za zhi = Zhonghua xueyexue zazhi 33(6):471–474
24. Ge M, Li X, Shi J, Shao Y, Zheng Y (2012) Clinical features and prognostic factors of Asian patients with paroxysmal nocturnal hemoglobinuria: results from a single center in China. Ann Hematol 91(7):1121–1128. https://doi.org/10.1007/s00277-012-1413-6
25. Green M, Styles T, Russell T, Sada C, Jallow E, Stewart J, Lazariashvili O, Lubomirova I, Cotlarciuc I, Sharma S, Han TS, Sharma P (2018) Non-genetic and genetic risk factors for adult

cerebral venous thrombosis. Thromb Res 169:15–22. https://doi.org/10.1016/j.thromres.2018.07.005
26. Taran LD (2003) Protein C system and thrombosis diseases. Ukr Biokhim Zh (1999) 75(1):18–24
27. Peacock-Young B, Macrae FL, Newton DJ, Hill A, Ariens RAS (2018) The prothrombotic state in paroxysmal nocturnal hemoglobinuria: a multifaceted source. Haematologica 103(1):9–17. https://doi.org/10.3324/haematol.2017.177618
28. Blair P, Flaumenhaft R (2009) Platelet alpha-granules: basic biology and clinical correlates. Blood Rev 23(4):177–189. https://doi.org/10.1016/j.blre.2009.04.001
29. Wiedmer T, Esmon CT, Sims PJ (1986) Complement proteins C5b-9 stimulate procoagulant activity through platelet prothrombinase. Blood 68(4):875–880
30. Du Y, Long Z, Xie H, Zhuang J, Han B (2016) The preliminary research in paroxysmal nocturnal hemoglobinuria with thrombosis. Zhonghua xue ye xue za zhi = Zhonghua xueyexue zazhi 37(4): 318–323. https://doi.org/10.3760/cma.j.issn.0253-2727.2016.04.014
31. Hosmer DW, Hosmer T, Le Cessie S, Lemeshow S (1997) A comparison of goodness-of-fit tests for the logistic regression model. Stat Med 16(9):965–980
32. Van Bijnen ST, Van Heerde WL, Muus P (2012) Mechanisms and clinical implications of thrombosis in paroxysmal nocturnal hemoglobinuria. J Thromb Haemost 10(1):1–10. https://doi.org/10.1111/j.1538-7836.2011.04562.x
33. Morange PE, Suchon P, Tregouet DA (2015) Genetics of venous thrombosis: update in 2015. Thromb Haemost 114(5):910–919. https://doi.org/10.1160/TH15-05-0410