NEWBORN SCREENING FOR G6PD DEFICIENCY IN HEFEI, FUYANG AND ANQING, CHINA: PREVALENCE, CUT-OFF VALUE, VARIANT SPECTRUM

SKRINING NOVOROĐENČADI NA NEDOSTATAK G6PD U GRADOVIMA HEFEI, FUYANG I ANQING, KINA: PREVALENCA, GRANIČNA VREDNOST, VARIANTNI SPEKTAR

Hui Li[1], Yah Ch'ih[2], Meiling Li[1], Yulei Luo[3], Hao Liu[4], Junyang Xu[1],
Wangsheng Song[1], Qingqing Ma[1], Ziyu Shao[1*]

[1]HeFei Women and Children Medical Care Center, HeFei City, Anhui Province, 230601, China
[2]Zhejiang Biosan Biochemical Technologies Co., Ltd, Hangzhou City, Zhejiang Province, 310012, China
[3]FuYang Maternal and Child Health Family Planning Service Center,
FuYang City, Anhui Province, 236000, China
[4]AnQing Maternal and Child Health Family Planning Service Center, AnQing City,
Anhui Province, 246000, China

Summary

Background: Glucose-6-phosphate dehydrogenase (G6PD) deficiency is an X-linked recessive Mendelian genetic disorder characterized by neonatal jaundice and hemolytic anemia, affecting more than 400 million people worldwide. The purpose of this research was to investigate prevalence rates of G6PD deficiency and to evaluate and establish specific cut-off values in early prediction of G6PD deficiency by regions (HeFei, FuYang, AnQing) on different seasons, as well as to investigate the frequencies of G6PD gene mutations among three regions mentioned above.

Methods: A total of 31,482 neonates (21,402, 7680, and 2340 for HeFei, FuYang, and AnQing cities, respectively) were recruited. Positive subjects were recalled to attend genetic tests for diagnosis. G6PD activity on the Genetic screening processor (GSP analyzer, 2021-0010) was measured following the manufacturer's protocol. The cut-off value was first set to 35 U/dL. The receiver operating characteristics (ROC) curve was employed to assess and compare the efficiency in predicting G6PD deficiency among HeFei, FuYang, and AnQing cities in different seasons.

Results: 31482 participants were screened for G6PD deficiency, and 29 neonates were finally identified with G6PD deficiency, with a prevalence rate of 0.09%. The incidence

Kratak sadržaj

Uvod: Nedostatak glukoza-6-fosfat dehidrogenaze (G6PD) je recesivni Mendelov genetski poremećaj povezan sa X hromozomom koji karakteriše neonatalna žutica i hemolitička anemija, koji pogađa više od 400 miliona ljudi širom sveta. Svrha ovog istraživanja je bila da se istraže stope prevalencije nedostatka G6PD i da se procene i utvrde specifične granične vrednosti u ranom predviđanju nedostatka G6PD po regionima (HeFei, FuLang, AnKing) u različitim godišnjim dobima, kao i da se ispita učestalost G6PD genske mutacije u okviru tri gore pomenuta regiona.

Metode: Ukupno je uključeno 31.482 novorođenčadi (21.402, 7.680 i 2.340 za gradova HeFei, FuYang i AnQing). Pozitivni subjekti su pozvani da prisustvuju genetskim testovima radi postavljanja dijagnoze. Aktivnost G6PD na procesoru genetskog skrininga (GSP analizator, 2021-0010) je merena prema protokolu proizvođača. Granična vrednost je prvo postavljena na 35 U/dL. Kriva radnih karakteristika prijemnika (ROC) je korišćena da se proceni i uporedi efikasnost u predviđanju nedostatka G6PD u gradovima HeFei, FuYang i AnQing u različitim godišnjim dobima.

Rezultati: Na nedostatak G6PD je testirano 31.482 novorođenčadi, a 29 novorođenčadi je identifikovano sa nedo-

Address for correspondence:

Ziyu Shao
No. 38, HeFei Women and Children Medical Care Center,
Yonghong Road, Luyang District HeFei City,
Anhui Province, 230601, China
e-mail: shaoziyu@xmmc-edu.cn

rate of the G6PD deficiency was 0.10% (22/21,380) for HeFei, 0.04%(3/7,667) for FuYang, and 0.17% (4/2,396) for AnQing. The optimal cut-off values of the ROC curve in the prediction of G6PD deficiency was 26.55 U/dL for HeFei (spring: 21.80 U/dL, summer: 26.55 U/dL, winter: 23.16 U/dL), 6.35 U/dL for FuYang, 7.15 U/dL for AnQing. R463H (17/57, 29.82%) was the most frequent variation, followed by R459L (15/57, 26.31%), H32R (8/57, 14.04%), L342F (6/57, 10.53%), V291M (6/57, 10.53%).

Conclusion: We have established the optimal cut-off values of G6PD activity in predicting G6PD deficiency among HeFei, FuYang and AnQing in different seasons. This will aid the early detection of G6PD deficiency among neonates of the regions mentioned above. We have also investigated the prevalence of the regions and discovered frequency mutations. Those results will help G6PD screening promotion and precision diagnosis for Anhui province.

Keywords: G6PD deficiency, newborn screening, ROC curve, variant spectrum, prevalence

statkom G6PD, sa stopom prevalencije od 0,09%. Stopa incidencije nedostatka G6PD je bila 0,10% (22/21.380) za HeFei, 0,04% (3/7.667) za FuYang i 0,17% (4/2.396) za AnQing. Optimalne granične vrednosti ROC krive u predviđanju nedostatka G6PD su bile 26,55 U/dL za HeFei (proleće: 21,80 U/dL, leto: 26,55 U/dL, zima: 23,16 U/dL), 6,35 U/dL za grad FuYang, 7,15 U/dL za AnQing. R463H (17/57, 29,82%) je bila najčešća varijacija, zatim R459L (15/57, 26,31%), H32R (8/57, 14,04%), L342F (6/57, 10,53%), V291M (6/57, 10,53%).

Zaključak: Ustanovili smo optimalne granične vrednosti G6PD aktivnosti u predviđanju nedostatka G6PD u gradovima HeFei, FuYang i AnQing u različitim godišnjim dobima. Ovo će pomoći u ranom otkrivanju nedostatka G6PD kod novorođenčadi u gore pomenutim regionima. Takođe smo istražili prevalence po regionima i otkrili mutacije frekvencije. Ovi rezultati će pomoći u promociji G6PD skrininga i postavljanju precizne dijagnoze za provinciju Anhui.

Ključne reči: nedostatak G6PD, skrining novorođenčadi, ROC kriva, varijantni spektar, prevalencija

Introduction

Glucose-6-phosphate dehydrogenase (G6PD) deficiency, known as favism, is an X-linked recessive Mendelian genetic disorder (1), which is characterized by neonatal jaundice and hemolytic anemia (2) when triggered by oxidative drugs, infection and the intake of fava beans (3). Mutations in the G6PD gene can result in reduced activity and stability of the G6PD enzyme, which is the crucial enzyme of the pentose phosphate pathway that produces reduced nicotinamide adenine dinucleotide phosphate (NADPH) to maintain sufficient levels of reduced glutathione (4, 5). Red blood cells would lyse when exposed to the factors mentioned above if there was an insufficient supply of NADPH and glutathione, which is the reason for the aforementioned clinical symptoms (6). G6PD deficiency affects more than 400 million people worldwide (5, 7, 8), especially prevalent in malaria-endemic areas (9), such as Africa, Oceania, Asia, and Mediterranean Europe (9, 10). G6PD deficiency is also common in China (11, 12). Nevertheless, the prevalence of G6PD deficiency, related to specific geographic regions and associated with ethnicities, is highly variable around China (13). The highest incidence of G6PD deficiency has been reported in southern China (14–16), for instance, in Guangdong, Yunnan, and Guangxi provinces. However, we have not yet known the precise incidence of G6PD deficiency in Anhui Province.

There is no cure for G6PD deficiency, as is true for many genetic disorders, and most G6PD deficiency individuals are asymptomatic. Hence, performing a G6PD deficiency screening program effectively facilitates early diagnosis and timely intervention. G6PD enzyme activity tests were performed within Newborn Screening programs in most provinces of China (17). According to the largest Youden Index, a specific optimal cut-off value generated by receiver operating characteristic (ROC) curve analysis was often to be used to predict G6PD deficiency and evaluate for diagnostic accuracy based on sensitivity, specificity, and the area under the ROC curve (AUC), having to discriminate all patients with G6PD levels accurately below a preset cut-off value, which is often followed with manufacturer's protocol. It is widely known that enzymatic activity may be affected by seasons and temperatures in vitro and, consequently, possess a different level of G6PD activity in different parts and seasons. Furthermore, through random X-chromosome inactivation, females heterozygous for G6PD could manifest various degrees of G6PD deficiency (18), which was called intermediate G6PD activities between typical normal and deficient G6PD activities (19). Therefore, the cut-off values for newborn screening of G6PD deficiency should optimize based on geographical regions, seasons, and gender.

Cut-off values of G6PD activity to determine G6PD deficiency differ for genetic background, for the G6PD gene is highly polymorphic with many variants, which could lead to decreased G6PD activity in erythrocytes or G6PD deficiency (20).To date, 270 mutations of the G6PD gene have been identified as pathogenic or likely pathogenic (https://www.ncbi.nlm.nih.gov/clinvar/?term=g6pd%5Bgene%5D&redir=gene), and over 200 pathogenic variants have been reported in China (21). As a previous study reported (21), G6PD gene variants K459M, L463H, A32G, P342S, and Q291K account for approximately 95% of the causative reasons for G6PD deficiency individuals in the Chinese population. However, the distribution of G6PD gene variants is related to geo-

graphical regions and ethnic groups, as reported. Hence, it is believed to be important to discover characteristics of the native G6PD variant spectrum.

In this study, we first established G6PD activity cut-off values of HeFei, FuYang, and AnQing because of geographical regions and seasons, based on a population cohort of 31482 neonates participating in Newborn Screening between December 2020 and October 2021. We also profiled the characteristics of the G6PD variant spectrum of the areas mentioned above and revealed the prevalence of the regions.

Materials and Methods

Subjects

A total of 31482 neonates (21402, 7680, and 2340 for HeFei, FuYang, and AnQing, respectively) were recruited through the Newborn Screening program at the newborn screening centre of HeFei Women and Children Medical Care Centre, FuYang Maternal and Child Health Family Planning Service Center, and AnQing Maternal and Child Health Family Planning Service Center between December 2020 and October 2021. Details are available in *Table I*. Dried blood spots (DBS) were collected from a heel stick within 48 hours after birth. The guardians of all subjects in the study signed informed consent, which was approved by the Medical Ethics Committee of the hospital mentioned above.

Determination of G6PD enzyme activity

G6PD enzyme activity was determined using a Genetic screening processor (GSP analyzer, 2021-0010), Panthera-PuncherTM 9 blood spot punching system, and Nenatal G6PD Kit, which were all purchased from Perkin Elmer (Perkin Elmer, Waltham, Massachusetts, United States), according to the manufacturer's instructions. In brief, G6PD calibrators, G6PD controls, and DBS samples were added separately to 96-well microplates, and the detection was performed using a GSP analyzer. According to the manufacturer's recommendation, the cut-off values of G6PD activity for G6PD deficiency were set to 35 U/dL. Subjects with G6P-D activity less than 35 U/dL were assayed using another DBS again.

Diagnosis test for G6PD deficiency

Subjects whose G6PD activity was less than 35 U/dL after the second test were defined as having positive results and recalled to attend a genetic test. Venous blood was gathered from each positive subject and then sent to the Clinical Laboratory of Zhejiang Biosan Biochemical Technologies Co., Ltd for laboratory tests. A client was diagnosed with G6PD deficiency positive if the genetic test results were positive.

A genetic test examined the 12 common hotspot mutation sites of the G6PD gene in the Chinese population, accounting for more than 95% of G6PD deficiency (21). Twelve common G6PD variants were: c.1360C>T(R454C), c.1376G>T(R459L), c.1388G>A(R463H), c.871G>A(V291M), c.1004C>A(A335D), c.1024C>T(L342F), c.95A>G(H32R), c.383T>C(L128P), c.392G>T(G131V), c.487G>A(G163S), c.592C>T(R198C), c.517T>C(F173L). The reaction was performed on a SLAN-96S real-time PCR system (Hongshi, Shanghai, China). *Supplementary Table S1* outlines PCR amplification and melting curve conditions in detail. By plotting the negative derivative of fluorescence versus temperature (dF/dT) as a function of temperature, Tm values and melting curves were generated automatically, being carried out with the software SLAN 8.0 (Hongshi).

Statical Analysis

Statical Analysis was performed using the SPSS analysis software(SPSS version 25.0 for Windows, SPSS Inc. Chicago, IL, United States). Non-normally distributed data was conducted on the Mann-Whitney U test and expressed as quartiles P50 (P25~P75). $P < 0.05$ was considered statistically significant. Receiver operative characteristics (ROC) curves were performed in MedCalc (MedCalc Soft-ware v.19.05, Mariakerke, Belgium). Mapping was made using GraphPad Prism 9.00 (GraphPad Software, La Jolla, CA, USA).

Results

A review of Newborn Screening and diagnosis of glucose-6 phosphate dehydrogenase deficiency

As depicted in *Table I*, 31482 participants were screened for G6PD deficiency, including 29 (0.09%) neonates finally identified with G6PD deficiency. Among the 29 neonates, 22 were males, and seven were not documented, of whom 22 (0.10%) were from HeFei, 3 (0.04%) were from FuYang, and 4 (0.17%) were from AnQing. Several characteristics differed significantly across age at initial testing, gender, and gestational age between the two groups, but had no significant differences in Birth Weight and Region. The normal group (Newborns without G6PD deficiency) showed smaller values in age at initial testing and Gestational age. Neonates confirmed G6PD deficiency were predominantly males (p = 0.002 < 0.05).

Effects of geographic regions and season on G6PD activity

We compared G6PD activity between geographical regions and seasons. In general, the distribution medians were significantly different (p<0.001) (59.58 (range 39.76–69.10)), (63.90

Table I Characteristics of newborns screened by G6PD newborn screening.

	Normal N = 31453	Patients N = 29	H	P
Age at initial testing[a]	4.00 (3.00~5.00)	4.00 (4.00~7.00)	8.809*	0.033
Gender				
Male	11405	22	9.995*	0.002
Female	9974	0		
No record	10074	7		
Gestational age[a] (weeks)	39.29 (38.57~40.14)	39.71 (37.00~39.93)	4.984*	0.026
Birth Weight (g)[a]	3350 (3050~3650)	3200 (2850~3530)	2.805	0.094
Region				
HF	21380	22	0.365	0.546
FY	7677	3		
AnQ	2396	4		

Normal: Neonates who were not diagnosed with G6PD deficiency by Newborn Screening.
Normal: Nenoates, diagnosed with G6PD deficiency by Newborn Screening.
H: Kruskal-Wallis Test; a: indicating that it did not fit with a Gaussian distribution after being assessed by the KS test.
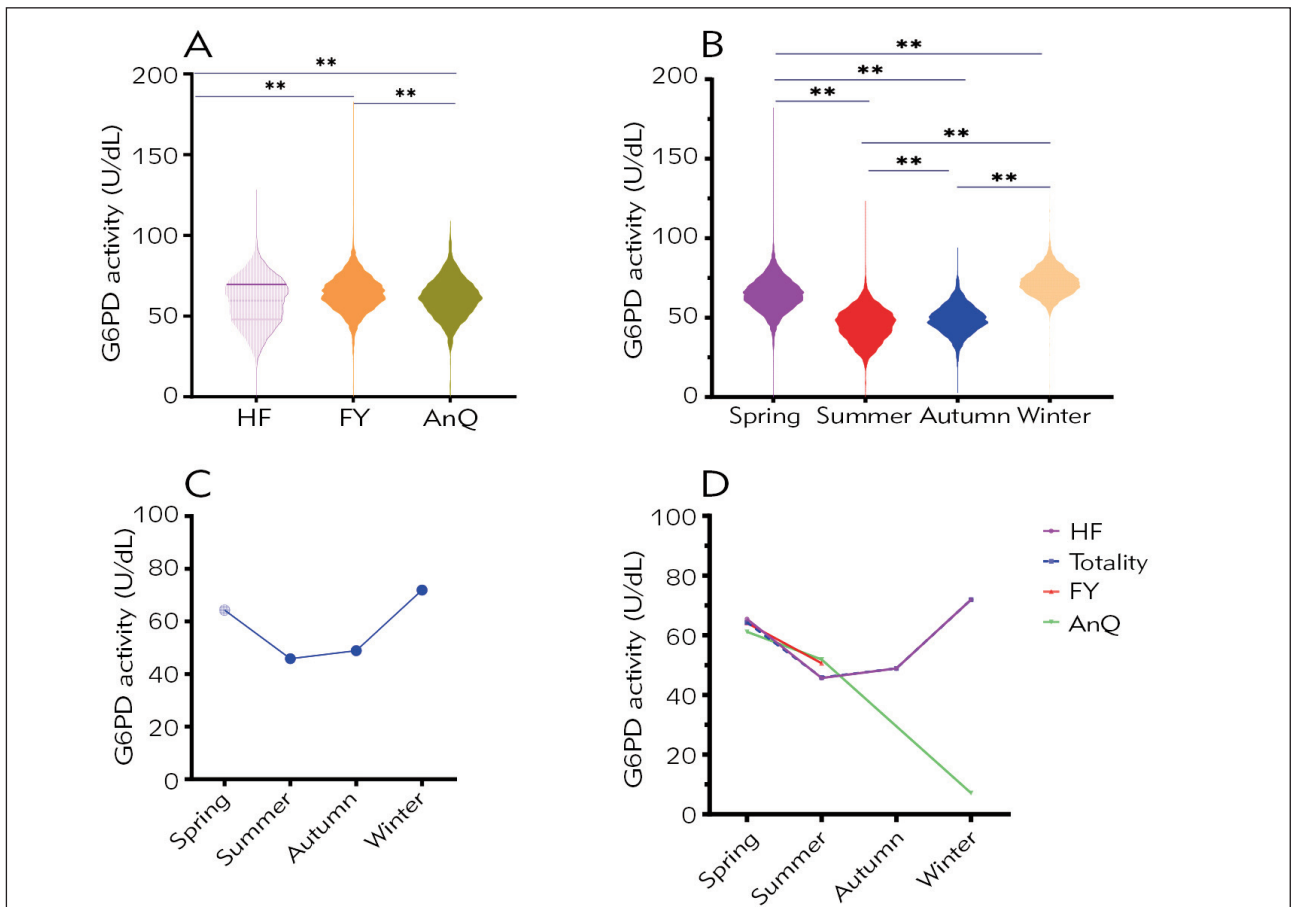
Figure 1 Effects of geographic regions and season on G6PD activity. ** represents p < 0.01 Abbreviations: HF, HeFei city, FY, FuYang city, AnQ for AnQing.

1A Comparison of enzymatic activity of HF, FY, and AnQ.
1B Comparison of spring, summer, autumn, and winter enzymatic activity.
1C Median of G6PD activity changes from Seasons.
1D G6PD activity distributions by regions (HeFei, FuYang, AnQing) in different seasons.

(range 57.00–71.00)) and (63.14 (range 53.29–69.05)) between HeFei, FuYang, and AnQing, respectively, as presented in *figure 1A*. As shown in *Figures 1B* and *1C*, significant seasonal differences existed in G6PD activity. G6PD activity presented declining trends during spring and summer and increasing trends from summer to winter (1C). In contrast, G6PD activity was significantly different between 4 categories: median values for spring, summer, autumn, and winter were 64.3000 (range 57.80–70.91), 45.83 (range 38.14–52.74), 48.89 (range 43.37–54.74), 71.8900 (range 66.43–77.65), respectively (*Table S1*). Due to no participants engaged in screening in certain seasons (FuYang in autumn, AnQing in autumn and winter), we could only display the whole season G6PD activity of HeFei, which was consistent with the tendency of Totality (*Figure 1D*).

Comparison of ROC curves between geographic regions or seasons

We performed ROC analysis to compare the diagnosis efficiency of geographical regions (*Figure S1, Table II*). There were no significant differences between HeFei, FuYang, and AnQing in the area under the ROC curves (AUC), as seen in *Table II*. Nevertheless, cut-off values determined by the ROC curves are differentiated. The ROC curve analysis revealed a high-predictive value of geographical regions (*Table II*) G6PD activity for G6PD deficiency, which in HeFei had AUC (CI 95%) of 0.999, with the optimal cut-off value of 26.55 U/dl (sensitivity, SS: 10.00%; and specificity, SP: 98.92%), in FuYang, having the AUC (CI 95%) of 1.000 with the optimal cut-off value of 6.35 U/dL (SS:100.00%; SP: 99.99%), and in AnQing, had the AUC (CI 95%) of 0.999 with the optimal cut-off value of 7.15 U/dL (SS:100.00%; SP: 99.87%).
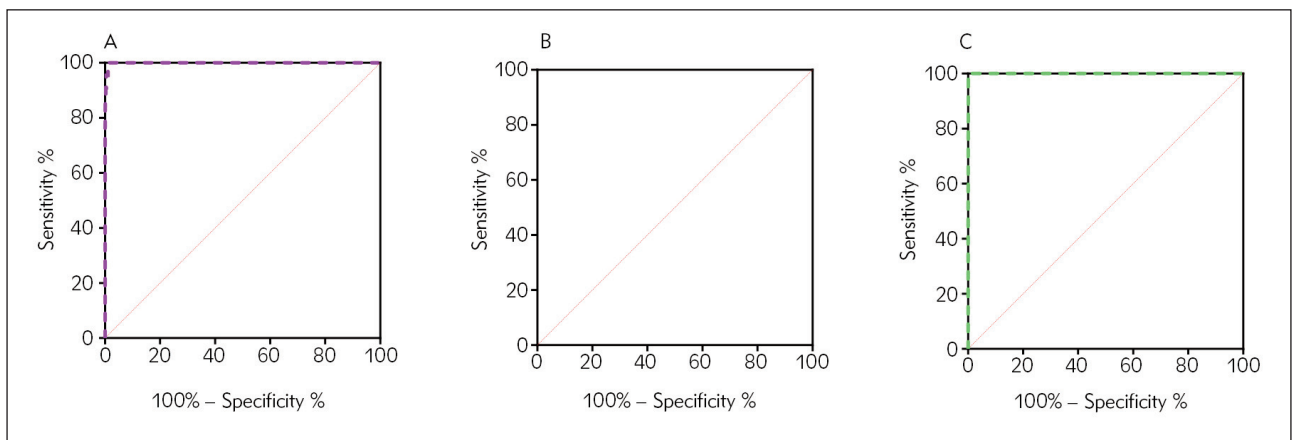
Figure S1 ROC curves of G6PD activity in the prediction of G6PD deficiency of Hefei, Fuyang, and Anqing city.
1A Hefei city AUC(95% CI) = 0.999, sensitivity = 100.00%, specificity = 98.92%
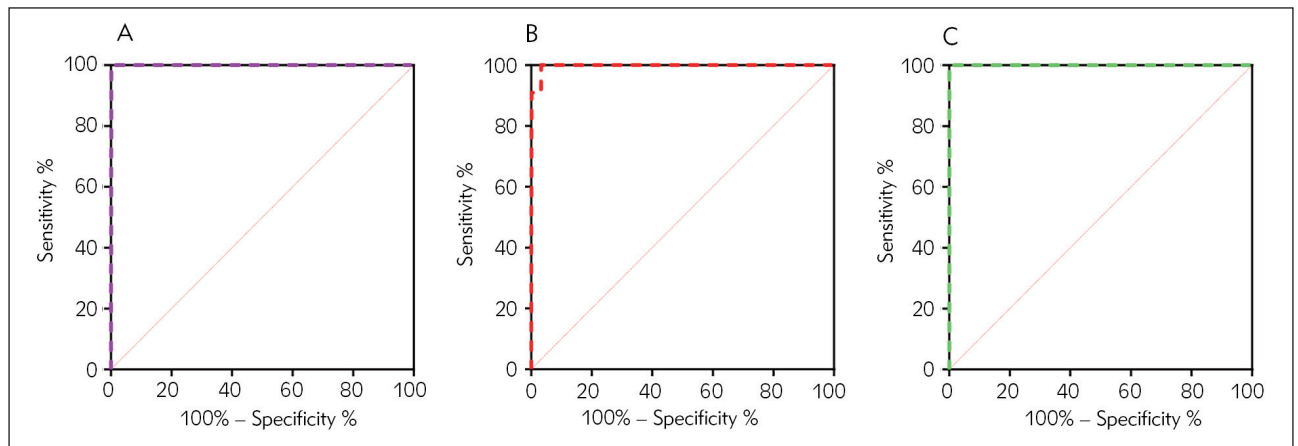1B Fuyang city AUC(95% CI) = 1.000, sensitivity = 100.00%, specificity = 99.99%
1C Anqing city AUC(95% CI) = 0.999, sensitivity = 100.00%, specificity = 99.87%

Table II Comparison of ROC curves of G6PD activity in predicting G6PD deficiency of HeFei, FuYang, and AnQing.

Variables	Cut-off values[a]	AUC (95% CI)	P values	SS (%)	SP (%)	Pairwise comparison of ROC curves		
						Variables	AUC difference	P value
HF	26.55	0.999 (0.998~0.999)	<0.0001	100.00	98.92	HF~FY	0.000974	0.0651
FY	6.35	1.000 (0.999~1.000)	<0.0001	100.00	99.99	FY~AnQ	0.000530	0.2309
AnQ	7.15	0.999 (0.997~1.000)	<0.0001	100.00	99.87	AnQ~HF	0.000443	0.5166

Abbreviations: HF, HeFei; FY, FuYang; AnQ, AnQing; SP, sensitivity; SP, specificity.
a:U/dL

Figure 2 ROC curves of G6PD activity in predicting G6PD deficiency between different seasons. 2A spring AUC (95% CI) = 1.000, sensitivity = 100.00%; specificity = 99.92%, 1B summer AUC(95% CI) = 0.997, sensitivity = 100.00%; specificity = 96.72% and 1C winter AUC(95% CI) = 1.000, sensitivity 100.00%; specificity = 99.98%.

Table III Comparison of ROC curves of G6PD activity in predicting G6PD deficiency between seasons.

Variables	Cut-off values	AUC (95% CI)	P values	SS (%)	SP (%)	Pairwise comparison of ROC curves		
						Variables	AUC difference	Pvalue
Spring	21.80	1.000 (0.999~1.000)	<0.0001	100.00	99.92	Spring~Summer Spring~Autumn	0.00279NA	0.3467NA
Summer	26.55	0.997 (0.995~0.998)	<0.0001	100.00	96.72	Summer~Autumn Summer~Winter	NA 0.00320	NA 0.2799
Autumn	NA	NA	NA	NA	NA	Autumn~Winter	NA	NA
Winter	23.16	1.000 (0.999~1.000)	<0.0001	100.00	99.98	Winter~Spring	0.000409	0.0209b

Abbreviations: SP, sensitivity; SP, specificity; NA indicates no data available.
b:P < 0.05

Table IV Comparison of ROC curves of G6PD activity in predicting G6PD deficiency between different seasons of HeFei.

Variables	Cut-off values	AUC (95% CI)	P values	SS (%)	SP (%)	Pairwise comparison of ROC curves		
						Variables	AUC difference	P value
Spring	21.80	1.000 (0.999~1.000)	<0.0001	100.00	99.98	Spring~Summer Spring~Autumn	0.00383 NA	0.2944 NA
Summer	26.55	0.997 (0.994~0.998)	<0.0001	100.00	96.69	Summer~Autumn Summer~Winter	NA 0.00395	NA 0.2789
Autumn	NA	NA	NA	NA	NA	Autumn~Winter	NA	NA
Winter	23.16	1.000 (0.999~1.000)	<0.0001	100.00	99.98	Winter~Spring	0.000121	0.5053

Abbreviations: SP, sensitivity; SP, specificity; NA indicates no data available.
b:P < 0.05

Table V Comparison of ROC curves of G6PD activity in predicting G6PD deficiency between different seasons of FuYang.

Variables	Cut-off values	AUC (95% CI)	P values	SS (%)	SP (%)	Pairwise comparison of ROC curves		
						Variables	AUC difference	Pvalue
Summer	6.35	1.000 (0.976~1.000)	<0.0001	100.00	100.00	Summer~Autumn Summer~Winter	NANA	NANA

Abbreviations: SP, sensitivity; SP, specificity; NA indicates no data available.

Table VI Comparison of ROC curves of G6PD activity in predicting G6PD deficiency between different seasons of AnQing.

Variables	Cut-off values	AUC (95% CI)	P values	SS (%)	SP (%)	Pairwise comparison of ROC curves		
						Variables	AUC difference	P value
Winter	7.15	1.000 (1.000~1.000)	0.0833	100.00	100.00	Winter~Spring	0.000418	1.0000

Abbreviations: SP, specificity; SP, specificity; NA indicates no data available.
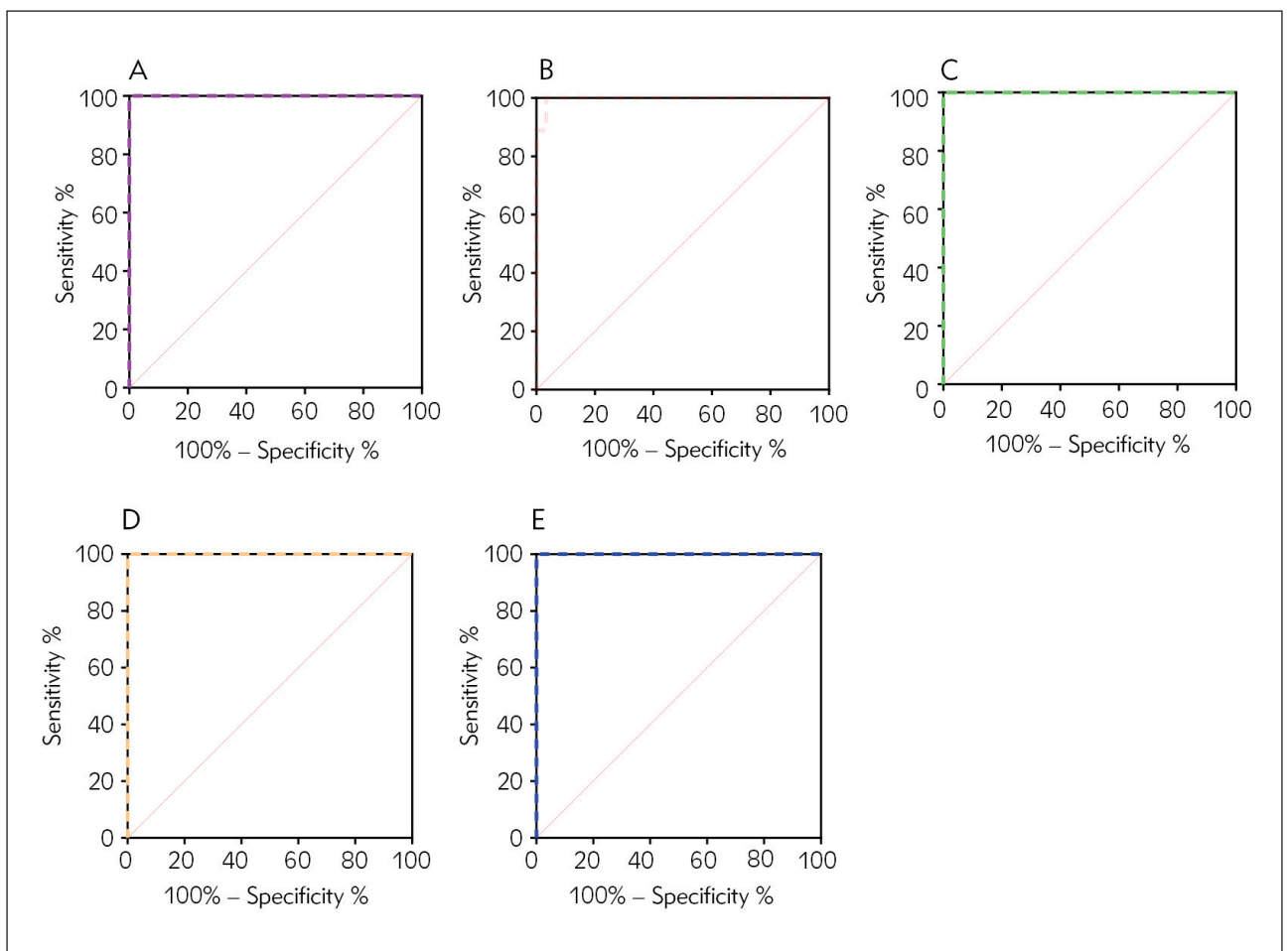
Figure S2 Comparison of ROC curves of G6PD activity in predicting G6PD deficiency between different seasons of different geographic regions

4A–4C: for the region of Hefei city; 4A Spring AUC(95% CI) = 1.000, sensitivity = 100.00%, specificity = 99.98% , 4B Summer AUC(95% CI) = 0.997, sensitivity = 100.00%, specificity = 96.69%, and 4C Winter AUC(95% CI) = 1.000, sensitivity = 100.00%, specificity = 99.98%
4D for the region of Fuyang city; Summer AUC(95% CI) = 1.000, sensitivity = 100.00%, specificity = 100.00%
4E for the region of Anqing city; AUC(95% CI) = 1.000, sensitivity = 100.00%, specificity = 100.00%

ROC analysis was also conducted on seasons (*Figure 2, Table III*). AUC values were not significantly different between each group (as there was no G6PD deficiency confirmed in autumn, values could not be calculated by ROC analysis), except for the significant difference between winter and spring. Accordingly, optimal cut-off values were 21.80 U/dL (SS: 100.00%; SP: 99.92%, AUC (CI 95%) = 1.000), 26.55 U/dL (SS: 100.00%; SP: 96.72%, AUC (CI 95%) = 0.997), 23.16 U/dL (SS:100.00%; SP: 99.98%, AUC (CI 95%) =1.000), of spring, summer, and winter, respectively.

Comparison of ROC curves between geographic regions in different seasons

To analyze the seasonal characteristics of G6PD activity in different geographical parts, the ROC analysis was performed during different seasons in HeFei, FuYang and AnQing (*Table IV, V, VI* and *Figure 2S*).

Concerning HeFei (*Table IV*), significant differences between each subgroup were not observed. And based on ROC curve analysis (*Table IV*), there was a high predictive value for seasons (G6PD activity) when predicting G6PD deficiency, with the AUC (CI 95%) of 1.000, 0.997, 1.000 (*Figure S2A, S2B, S2C*), and the optimal cut-off values of 21.80 U/dL (sensitivity, SS: 10.00%; and specificity, SP: 99.98%), 26.55 U/dL (SS: 100.00%; SP: 96.69%), 23.16 U/dL (SS: 100.00%; SP: 99.98%), in spring, summer and winter, respectively. As for FuYang and AnQing, sufficient data being not available for between subgroup analysis, we could only obtain the ROC curve (*Figure 2D, 2E*) characteristics of summer (for FuYang) (*Table V*) and winter (for AnQing) (*Table VI*), and all the AUC (CI 95%) equalled 1.000, with the optimal cut-off values of 6.35 U/dl (sensitivity, SS: 10.00%; and specificity, SP: 100.00%), 7.15 U/dL (SS:100.00%; SP: 100.00%), respectively. However, the p-value of AnQing (subgroup: winter) was 0.0833 > 0.05, and the result of ROC analysis for this category was unreliable.
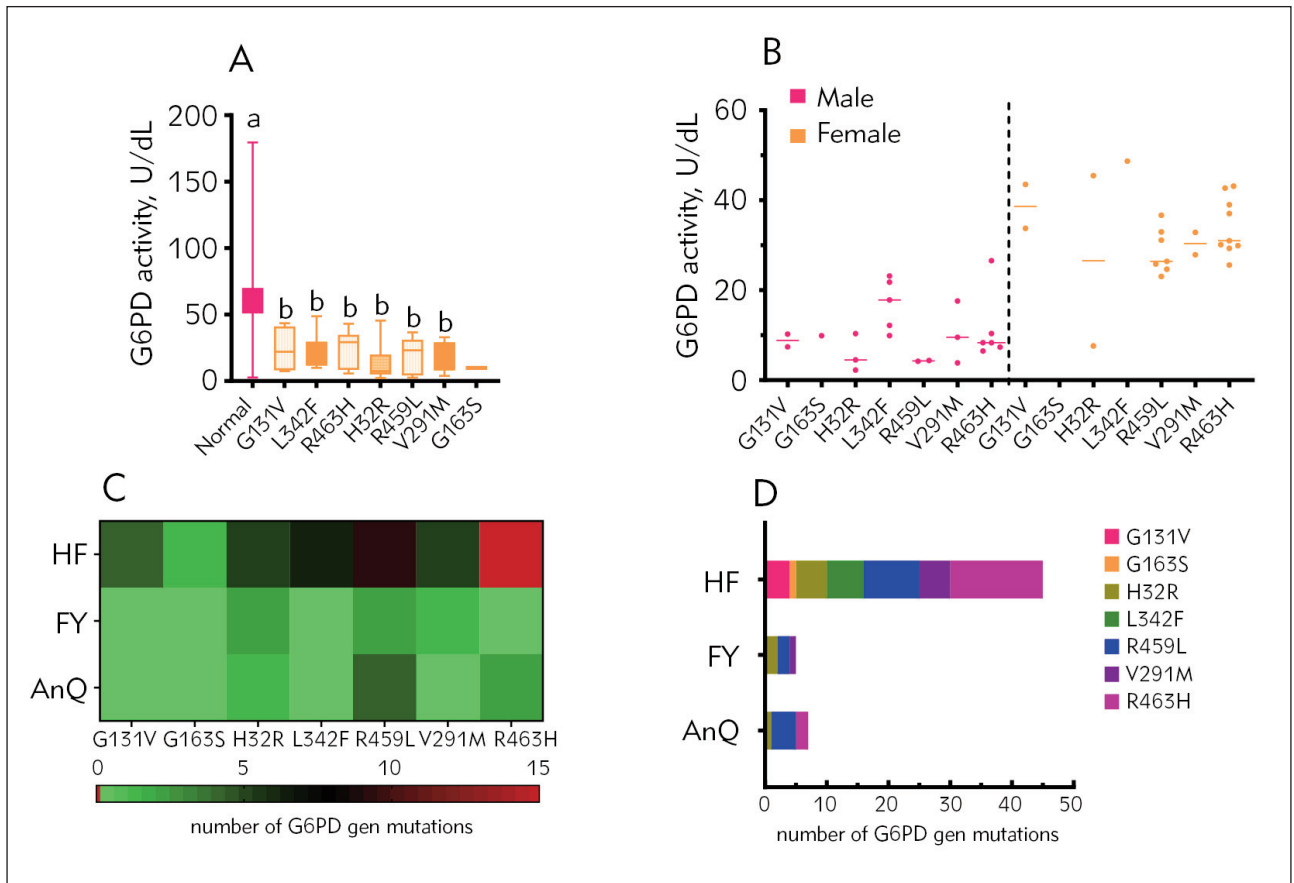
Figure 3 G6PD activity of G6PD gene mutations and its distribution of mutation frequencies among the geographic regions. 5A Comparison of G6PD enzymatic activity of Normal and mutated G6PD enzyme. 5B Comparison of G6PD enzymatic activity between different mutations in Male neonates and Female nenates.5C Heatmap of mutations in G6PD gene in individuals from different geographic regions. 5D The plot shows the number of G6PD gene mutations in each region. Different G6PD genotypes show significant differences in G6PD activity when boxplots are marked with different letters and nonsignificant differences in G6PD activity when boxplots are marked with the same letters.

Abbreviations: HF, HeFei; FY, FuYang; AnQ, AnQing.

G6PD activity of G6PD gene mutations and its distribution of mutation frequencies among the geographic regions

The phrase »Normal« represented subjects who did not harbour G6PD gene variants. The G6PD activity of the Normal group was significantly higher than all groups of newborns with G6PD gene variation (*Figure 5A*). Nevertheless, no significant difference was observed for any grouping of G6PD gene variants (*Figure 3A*). Significant differences existed between males and females regarding G6PD activity, with females having a high G6PD activity (*Figure 3B*). We also found that G6PD variant R463H was the most frequent (17/57, 29.82%) variation (*Figure 3C*), followed with R459L (15/57, 26.31%), H32R (8/57, 14.04%), L342F (6/57, 10.53%), V291M (6/57, 10.53%) (*Table S3, Figure 3D*). Concerning HeFei, R463H was the most commonly mutated allele (15/45, 33.33%), followed by R459L (9/45, 20.00%), L342F (6/45, 13.33%), H32R (5/45, 11.11%), V291M (5/45, 11.11%) and G131V (4/45, 8.89%) (*Table S4*). Except for H32R, R459L, and R463H (*Table S2*), no other variants were found in FuYang or AnQing.

Discussion

In this study, we first reported the prevalence of G6PD deficiency, cut-off values of G6PD activity optimized based on regions and seasons, and the G6PD gene variant spectrum for HeFei, FuYang and AnQing city. It would provide reference data values for preventing and treating G6PD deficiency in Anhui province.

Anhui province has a large population of about 60,000,000; every year, the population is estimated at approximately 400,000 births. For G6PD deficiency, no epidemiological data was previously reported. Our study revealed a prevalence of 0.09%, which is roughly consistent with most provinces in North China and lower than most provinces in South China except Hubei province, neighbouring Anhui province (21). HeFei, as the capital of Anhui province, located in the middle of Anhui province, had a higher incidence rate (0.10%) than FuYang, located in the north of Anhui province and population numbers being reasonably consistent with HeFei, possessed the incidence rate of 0.04%. But the incidence rate of HeFei was lower than that of AnQing, located along the Yangtze River, which was thought to be associated with the high incidence of G6PD deficiency (22), and had an incidence rate of 0.17%. These results were approximately consistent with the trend of south-high-north-low. We also found significant differences in age at initial testing and gestational age between the normal and G6PD deficiency groups, which likely happened by chance. A major drawback of this research is that no deficient sample was identified in

the female population by the method used to determine G6PD activity. This is indicative because gene analysis indicated the presence of multiple variants in the female population.

We found significant variation in the median results by comparing the G6PD activity obtained for HeFei, FuYang, and AnQing groups, which decreased in the following order: FuYang > AnQing > HeFei. Whereas, uncomprehensively, the data of FuYang and AnQing which obscured the fact was responsible for the conclusion. As presented in Figure 1D, comparing the differences should depend on the spring data rather than the overall data. In fact, the median of G6PD activity was observed in the following order: HF (65.42 (range 59.79–71.25)) > FY (64.10 (range 57.30–71.10)) > AnQ (61.1700 (range 53.37–69.076)). The cause of these discrepancies is poorly defined and remains to be further investigated. Moreover, we revealed that the changing trend of the levels of G6PD activity was consistent with that first decreasing and then increasing with the season, followed the order of winter (71.9 range (66.44–77.65)) > spring (65.42 range (59.74–71.2375)) > autumn (48.89 range (43.48–54.79)) > summer (45.65 range (37.91–52.47)). One plausible mechanism is that temperature may affect the activity of G6PD. High temperatures could decrease the activity of G6PD in summer, while low temperatures increase it in winter.

We have established cut-off values predicated on geographic regions and seasons and demonstrated that cut-off values varied with different geographic regions and seasons. The optimal derived cut-offs for spring, summer and winter to detect G6PD deficiency in the region of HeFei, were 21.80 U/dl (SS:100.00%; SP: 99.92%, AUC(CI 95%) = 1.000), 26.55 U/dl (SS: 100.00%; SP: 96.72%, AUC (CI 95%) = 0.997) and 23.16 U/dl (SS: 100.00%; SP: 99.98%, AUC (CI 95%) = 1.000), respectively, which agreed with the overall results (*Table III*). Additionally, the best cut-off values of FuYang and AnQing differed from those of HeFei, which was also supported by the result of the ROC curve, and the cut-off values of FuYang and AnQing were 6.35 (sensitivity 100%, specificity 99.99%) and 7.15 (sensitivity 100%, specificity 99.87%), respectively.

In our research, the results were considered the most suitable cut-off values for FuYang and AnQing because there was insufficient data to establish different cut-off values constructed on seasons. To our knowledge, no previous study reported accurate cut-off values of Anhui province relying on seasonal factors. It was filled in this research, and before this study, the cut-off values were set to 35 U/dl, followed by the manufacturer's instructions, which was significantly higher than the result of our research. Consequently, our results provided a precise approach to facilitate G6PD deficiency screening,

achieving higher screening accuracy, reducing unnecessary anxiety resulting from false positives, and saving medical resources.

We found that G6PD gene variant R463H was a significant source of hotspot mutations in HeFei and identified the gene mutated frequency arranged in the order of R459L (24.56%) > H32R (14.04%) > L342F (10.53%) > V291M (10.53%), of which overlapped with the gene frequency of Chongqing and Guangxi (21). R463H emerged in the past 5000 to 6000 years (23) and was considered to have a significantly higher bilirubin level than that of R459L among hyperbilirubinemia neonates with G6PD deficiency (24). Moreover, R459L is thought to have commenced in the past 3125 to 3750 years (24), located at 12 bases from R463H. Those two sites are located in the conserved regions, which contain more than 10 highly amino acids, and more importantly, these two variants stay closer to the first binding site of NADP (386–387) in space. Consequently, the potential mechanisms underlying the effect of the binding process between the G6PD enzyme and NADP (25) may influence the G6PD activity. Our research results likely indicated that individuals, especially males carrying R463H and R459L, had to attach greater importance and higher priority to manage.

There remain several shortcomings in this study. Firstly, due to the limited investigation and the lack of diagnosis data from autumn, the G6PD activity cut-off values have not been established for autumn. Likewise, relying on different seasons, we could not establish the best suitable cut-off values of G6PD activity for FuYang and AnQing city. Thirdly, a small proportion of G6PD deficiency cases may be missed because we evaluated the positive subjects relying exclusively on hotspot mutations of the G6PD gene, which accounts for approximately 95% of causative reasons for G6PD deficiency individuals in the Chinese population, as the previous study reported (21). Actually, there are continuous reports describing additional findings in novel mutations of the G6PD gene (26, 27), which is essential for identifying G6PD deficient patients. With the development of gene sequencing, high throughput, low-cost assay for detecting the mutations of the G6PD gene may soon become a reality. Next Generation Sequencing(NGS) is rapidly becoming an essential diagnostic tool for genetic disease patients (28, 29). The advantage of NGS, in contrast, is that we can direct sequence the 13 exons of the flanking sequences of the G6PD gene other than G6PD hotspot mutations.

In addition, we have first investigated the prevalence, cut-off values and variant spectrum of G6PD deficiency patients, depending on a population cohort of 31482 neonates from Newborn Screening among HeFei, FuYang, and AnQing. Our study will provide data to promote G6PD deficiency screening being integrated into the Newborn Screening program in Anhui province for earlier identification of neonates at risk of G6PD deficiency.

Authors contributions

The study was conceived and designed by all authors. Hui Li is the study leader. Yah Ch'ih conducted the genetic test for G6PD deficiency. Meiling Li, Yulei Luo and Hao Liu performed the G6PD newborn screening test. Junyang Xu conducted the quality control of the newborn screening test. Qingqing Ma and Wangsheng Song were responsible for diagnosing and treating positive subjects. Hui Li gathered the newborn screening data and performed the statistical analysis of experimental data. Ziyu Shao undertook the project coordination. Yah Ch'ih wrote and revised the manuscript. All authors have read and supported the final draft.

Conflict of interest statement

All the authors declare that they have no conflict of interest in this work.

References

1. Anderle A, Bancone G, Domingo G J, et al. Point-of-Care Testing for G6PD Deficiency: Opportunities for Screening. Int J Neonatal Screen 2018; 4(4): 34.

2. Hamali H A. Glucose-6-Phosphate Dehydrogenase Deficiency: An Overview of the Prevalence and Genetic Variants in Saudi Arabia. Hemoglobin 2021; 45(5): 287–95.

3. Beauverd Y, Adam Y, Assouline B, et al. COVID-19 infection and treatment with hydroxychloroquine cause severe haemolysis crisis in a patient with glucose-6-phosphate dehydrogenase deficiency. Eur J Haematol 2020; 105(3): 357–9.

4. Tu D, Gao Y, Yang R, et al. The pentose phosphate pathway regulates chronic neuroinflammation and dopaminergic neurodegeneration. J Neuroinflammation 2019; 16(1): 255.

5. Ryan K, Tekwani B L. Current investigations on clinical pharmacology and therapeutics of Glucose-6-phosphate

dehydrogenase deficiency. Pharmacol Ther 2021; 222: 107788.

6. Lee H Y, Ithnin A, Azma R Z, et al. Glucose-6-Phosphate Dehydrogenase Deficiency and Neonatal Hyperbilirubinemia: Insights on Pathophysiology, Diagnosis, and Gene Variants in Disease Heterogeneity. Front Pediatr 2022; 10: 875877.

7. Garcia A A, Koperniku A, Ferreira J C B, et al. Treatment strategies for glucose-6-phosphate dehydrogenase deficiency: past and future perspectives. Trends Pharmacol Sci 2021; 42(10): 829–44.

8. Belfield K D, Tichy E M. Review and drug therapy implications of glucose-6-phosphate dehydrogenase deficiency. Am J Health Syst Pharm 2018; 75(3): 97–104.

9. Dechyotin S, Sakunthai K, Khemtonglang N, et al. Prevalence and Molecular Characterization of Glucose-6-Phosphate Dehydrogenase (G6PD) Deficiency in Females from Previously Malaria Endemic Regions in Northeastern Thailand and Identification of a Novel G6PD Variant. Mediterr J Hematol Infect Dis 2021; 13(1): e2021029.

10. Ashley E A, Poespoprodjo J R. Treatment and prevention of malaria in children. Lancet Child Adolesc Health 2020; 4(10): 775–89.

11. Wang X, Xia Z, He Y, et al. Newborn Screening for G6PD Deficiency in Xiamen, China: Prevalence, Variant Spectrum, and Genotype-Phenotype Correlations. Front Genet 2021; 12: 718503.

12. Sathupak S, Leecharoenkiat K, Kampuansai J. Prevalence and molecular characterization of glucose-6-phosphate dehydrogenase deficiency in the Lue ethnic group of northern Thailand. Sci Rep 2021; 11(1): 2956.

13. Ntaios G, Chatzinikolaou A, Tomos C, et al. Prevalence of glucose-6-phosphate dehydrogenase deficiency in Northern Greece. Intern Med J 2008; 38(3): 204–6.

14. Peng Q, Li S, Ma K, et al. Large cohort screening of G6PD deficiency and the mutational spectrum in the Dongguan District in Southern China. PLoS One 2015; 10(3): e0120683.

15. He M, Lin K, Huang Y, et al. Prevalence and Molecular Study of G6PD Deficiency in the Dai and Jingpo Ethnic Groups in the Dehong Prefecture of the Yunnan Province. Hum Hered 2018; 83(2): 55–64.

16. Fu C, Luo S, Li Q, et al. Newborn screening of glucose-6-phosphate dehydrogenase deficiency in Guangxi, China: determination of optimal cut-off value to identify heterozygous female neonates. Sci Rep 2018; 8(1): 833.

17. Delfavero J J, Jnah A J, Newberry D. Glucose-6-Phosphate Dehydrogenase Deficiency and the Benefits of Early Screening. Neonatal Netw 2020; 39(5): 270–82.

18. Wang J, Xiao Q Z, Chen Y M, et al. DNA hypermethylation and X chromosome inactivation are major determinants of phenotypic variation in women heterozygous for G6PD mutations. Blood Cells Mol Dis 2014; 53(4): 241–5.

19. Chu C S, Freedman D O. Tafenoquine and G6PD: a primer for clinicians. J Travel Med 2019; 26(4).

20. Luzzatto L, Ally M, Notaro R. Glucose-6-phosphate dehydrogenase deficiency. Blood 2020; 136(11): 1225–40.

21. Liu Z, Yu C, Li Q, et al. Chinese newborn screening for the incidence of G6PD deficiency and variant of G6PD gene from 2013 to 2017. Hum Mutat 2020; 41(1): 212–21.

22. Yang H, Wang Q, Zheng L, et al. Incidence and molecular characterization of Glucose-6-Phosphate Dehydrogenase deficiency among neonates for newborn screening in Chaozhou, China. Int J Lab Hematol 2015; 37(3): 410–9.

23. Zheng Y, Wang J, Liang X, et al. Epidemiology, evolutionary origin, and malaria-induced positive selection effects of G6PD-deficient alleles in Chinese populations. Mol Genet Genomic Med 2020; 8(12): e1540.

24. Wisnumurti D A, Sribudiani Y, Porsch R M, et al. G6PD genetic variations in neonatal Hyperbilirubinemia in Indonesian Deutromalay population. BMC Pediatr 2019; 19(1): 506.

25. Chen Y, Xiu W, Dong Y, et al. Mutation of glucose-6-phosphate dehydrogenase deficiency in Chinese Han children in eastern Fujian. Medicine (Baltimore) 2018; 97(30): e11553.

26. Malik S, Zaied R, Syed N, et al. Seven novel glucose-6-phosphate dehydrogenase (G6PD) deficiency variants identified in the Qatari population. Hum Genomics 2021; 15(1): 61.

27. Arunachalam A K, Sumithra S, Maddali M, et al. Molecular Characterization of G6PD Deficiency: Report of Three Novel G6PD Variants. Indian J Hematol Blood Transfus 2020; 36(2): 349–55.

28. Yang Y, Wang Y, Zhou L, et al. Molecular Genetic Screening of Neonatal Intensive Care Units: Hyperbilirubinemia as an Example. Appl Clin Genet 2022; 15: 39–48.

29. Klau J, Abou Jamra R, Radtke M, et al. Exome first approach to reduce diagnostic costs and time – retrospective analysis of 111 individuals with rare neurodevelopmental disorders. Eur J Hum Genet 2022; 30(1): 117–25.