Supplementary material

Supplemental table 1. The distribution and minor allele frequency of *UGT1A1* variants in 346 Gilbert syndrome and 11 Crigler-Najjar syndrome type II patients.

Nucleotide	Zygosity	N(%)	Minor Allele Frequency
c.-3275T>G	het	145(40.62%)	54.76%
	hom	123(34.45%)	
A(TA)6TAA>A(TA)7TAA	het	151(42.30%)	39.08%
	hom	64(17.93%)	
c.211G>A	het	138(38.66%)	31.37%
	hom	43(12.04%)	
c.1456T>G	het	19(5.32%)	6.58%
	hom	14(3.92%)	
c.1091C>T	het	28(7.84%)	4.76%
	hom	3(0.84%)	
c.686C>A	het	33(9.24%)	4.62%

	hom	0(0.00%)	
c.652T>G	het	0(0.00%)	0.84%
	hom	3(0.84%)	
c.1447T>G	het	0(0.00%)	0.28%
	hom	1(0.28%)	
c.264_267del	het	1(0.28%)	0.14%
	hom	0(0.00%)	
c.554A>C	het	1(0.28%)	0.14%
	hom	0(0.00%)	
c.625C>T	het	1(0.28%)	0.14%
	hom	0(0.00%)	
c.1024C>T	het	1(0.28%)	0.14%
	hom	0(0.00%)	
c.1047delG	het	1(0.28%)	0.14%
	hom	0(0.00%)	
c.1082C>T	het	1(0.28%)	0.14%
	hom	0(0.00%)	

c.1156G>T	het	1(0.28%)	0.14%
	hom	0(0.00%)	
c.1305-1G>A	het	1(0.28%)	0.14%
	hom	0(0.00%)	
c.1463C>T	het	1(0.28%)	0.14%
	hom	0(0.00%)	

Supplementary Table 2. Analysis of hemolysis-related gene variations in 39 patients with hemolytic disease-associated unconjugated hyperbilirubinemia.

Patient	GENE	Chromosomal location	Transcript	Exon / Intron	cDNA change	Protein_change	Zygosity	gnomAD_EAS	ACMG	Tbil (μmol/L)
P1	SPTB	chr14:65260058	NM_001355436.2	exon13	c.2323G>A	p.Glu775Lys	het	0.0006	Uncertain	177
	SPTB	chr14:65246494	NM_001355436.2	exon2	c.208C>T	p.Arg70Ter	het	0	Pathogenic	

P2	SPTB	chr14:65246494	NM_001355436.2	exon20	c.4421delT	p.Leu1474fs	het	0	Likely pathogenic	154.2
	PIEZO1	chr1:158582607	NM_001142864.4	exon25	c.3602C>T	p.Thr1201Met	het	0	Uncertain	
P3	SPTB	chr17:42330654-42330658	NM_001355436.2	exon20	c.4421delT	p.Leu1474fs	het	0	Likely pathogenic	127.8
	PIEZO1	chr15:43489546	NM_001142864.4	exon25	c.3602C>T	p.Thr1201Met	het	0	Uncertain	
P4	SPTA1	chr1:29323824	NM_003126.4	exon51	c.7134G>A	p.Gln2378Gln	het	0	Uncertain	103.8
P5	SLC4A1	chr14:65266584	NM_000342.4	exon17	c.2139_2142del	p.Gly713fs	het	0	Likely pathogenic	178.5
P6	EPB42	chr1:158581098	NM_000119.3	exon13	c.2120C>T	p.Thr707Ile	het	0.0002	Uncertain	72.2
P7	EPB41	chr17:42330519	NM_001166005.2	exon4	c.779C>T	p.Thr260Ile	het	0	Uncertain	117.5
P8	SPTB	chr1:15858409	NM_001355436	exon9	c.944dupT	p.Thr316Hisf	het	0	Likely	80

		1	.2			s*76			pathogenic	
P9	SPTA1	chr1:158622405	NM_003126.4	exon52	c.7214_7215delCT	p.Ser2405Trpfs*3	het	0	Likely pathogenic	122.1
P10	SLC4A1	chr1:29342254-29342256	NM_000342.4	exon17	c.2278C>T	p.Arg 760Trp	het	0	Likely pathogenic	83.1
P11	SPTA1	chr1:29319893	NM_003126.4	exon49	c.6794T>C:	p.Ile2265Thr	hom	0.6088	Benign	175
	SPTA1	chr8:41575723	NM_003126.4	exon40	c.5572C>G:	p.Leu1858Val	hom	0.1694	Benign	
	SLC4A1	chr1:158596754	NM_000342.4	exon4	c.113A>C	p.Asp38Ala	het	0.0937	Benign	
P12	SPTA1	chr14:65289713	NM_003126.4	exon23	c.3227G>A	p.Arg1076His	het	0	Uncertain	56.1
	ANK1	chr1:158590142	NM_000037.4	exon25	c.2698_2699del	p.Asp900fs	het	0	Likely pathogenic	
P13	EPB41	chr1:158597458	NM_001166005.2	exon6	c.881_882del	p.Ala294fs	het	0	Likely pathogenic	87

P14	EPB41	chr8:41655029	NM_001166005.2	exon3	c.520G>A	p.Glu174Lys	het	0	Uncertain	119.6
P15	ANK1	chr1:158624457-158624458	NM_000037.4	intron10	c.1108-1G>A	\	het	0	Pathogenic	111
P16	SPTA1	chr14:65246489	NM_003126.4	exon41	c.5708T>C	p.Ile1903Thr	het	0	Uncertain	100.1
	SPTA1	chr8:41580637-41580642	NM_003126.4	exon2	c.203G>A	p.Trp68Ter	het	0	Pathogenic	
P17	SPTB	chr8:41550246	NM_001355436.2	exon2	c.99delC	p.Asp33Glufs*13	het	0	Likely pathogenic	63.5
P18	SPTA1	chr8:41526071	NM_003126.4	exon44	c.6235C>T	p.Arg2079Trp	het	0.0008	Uncertain	84.9
P19	SPTA1	chr17:42330647	NM_003126.4	exon40	c.5621G>A	p.Arg1874Gln	het	0	Uncertain	140.2
	KCNN4	chr1:158653199	NM_002250.3	exon6	c.988C>G:	p.Arg330Gly	het	0	Uncertain	
P20	ANK1	chr1:29342222	NM_000037.4	intron	c.27+1G>	\	het	0	Likely	130.5

				1	A				pathogenic	
P21	PIEZO1	chr17:42335127	NM_001142864.4	exon40	c.5743C>T	p.Arg1915Cys	het	0.0123	Uncertain	139.8
	SPTA1	chr17:42333145	NM_003126.4	exon21	c.2980-2981insCC	p.Arg994Profs*6	het	0	Likely pathogenic	
	SPTB	chr1:158592847	NM_001355436.2	exon21	c.4427C>A	p.Ser1476Ter	het	0	Pathogenic	
P22	PIEZO1	chr14:65253869	NM_001142864.4	exon34	c.4499G>C	p.Arg1500Pro	het	0	Uncertain	80
	KCNN4	chr16:88800372-88800375	NM_002250.3	exon5	c.863_864del	p.Lys288Thrfs*4	het	0	Likely pathogenic	
P23	ANK1	chr16:88798865	NM_000037.4	intron9	c.909+1_909+5del	\	het	0	Likely pathogenic	97.3
	PIEZO1	chr16:88799772	NM_001142864.4	exon29	c.4072C>T	p.Arg1358Cys	het	0	Uncertain	
P24	ANK1	chr16:88805014	NM_000037.4	exon31	c.3778T>C	p.Cys1260Arg	het	0	Uncertain	89.6

P25	ANK1	chr16:215982-231133	NM_000037.4	exon39	c.5108G>A	p.Trp1703Ter	het	0.000001	Pathogenic	112.2
P26	SLC4A1	chr7:134346528	NM_000342.4	exon17	c.2150C>T	p.Ala717Val	het	0	Uncertain	59.5
P27	SPTA1	chrX:138878562	NM_003126.4	exon3	c.352C>T	p.Arg118Ter	het	0.000004	Likely pathogenic	152.1
P28	EPB41	chr14:65260058	NM_001166005.2	Exon6	c.848C>T	p.Thr283Ile	het	0.009	Uncertain	90.9
P29	SLC4A1	chr14:65246494	NM_000342.4	exon12	c.1331C>A	p.Thr444Asn	het	0.003	Uncertain	42.3
P30	SLC4A1	chr14:65246494	NM_000342.4	exon14	c.1696C>T	p.Pro566Ser	het	0	Uncertain	55.8
	EPB41	chr1:158582607	NM_001166005.2	exon11	c.1535G>A	p.Arg512Gln	het	0	Uncertain	
P31	SPTA1	chr17:42330654-42330658	NM_003126.4	exon43	c.6046C>T	p.Arg2016Cys	het	0	Uncertain	110.7
	SPTA1	chr15:4348954	NM_003126.4	exon4	c.5992G>C	p.Ala1998Pro	het	0	Uncertain	

		6		3						
P32	SPTB	chr1:29323824	NM_001024858. 4	exon15	c.2807_2813dupGGCAGGC	p.Phe939Alafs*67	het	0	Likely pathogenic	227.4
P33	PIEZO1	chr14:65266584	NM_001142864.4	exon17	c.2268_2270del	p.Glu756del	het	0.0012	Uncertain	80.8
	PIEZO1	chr1:158581098	NM_001142864.4	exon8	c.1006C>T	p.Arg336Cys	het	0	Uncertain	
P34	PIEZO1	chr17:42330519	NM_001142864.4	exon21	c.2869C>T	p.Gln957Ter	het	0	Pathogenic	265.3
P35	PIEZO1	chr1:158584091	NM_001142864.4	exon19	c.2578G>A	p.Val860Met	het	0	Uncertain	131.2
P36	PIEZO1	chr1:158622405	NM_001142864.4	exon6	c.596C>A	p.Ala199Asp	het	0	Uncertain	44.8
P37	chr16:215982-231133		\	\	\	\	het	0	\	46.1

	delete（contain HBA1and HBA2）									
P38	BPGM	chr1:29319893	NM_001293085.2	exon3	c.269G>A	p.Arg90His	het	0	Uncertain	81
P39	ATP11C	chr8:41575723	NM_001353812.2	exon12	c.1076T>G		het			122.7

Supplementary Table 3. Variables missing rate in the training and internal validation cohort.

	Variables	Missing rate
1	COOMBS	59(15.57%)
2	Peripheral blood smear	60(15.83%)

3	Cholelithiasis	60(15.83%)
4	Splenomegaly	60(15.83%)
5	sex	0(0%)
6	age	0(0%)
7	WBC	3(0.79%)
8	PLT	4(1.06%)
9	RBC	6(1.58%)
10	HGB	0(0%)
11	HCT	5(1.32%)
12	MCV	5(1.32%)
13	MCH	5(1.32%)
14	MCHC	5(1.32%)
15	RDW	5(1.32%)
16	RDWSD	10(2.64%)
17	N	4(1.06%)

18	L	4(1.06%)
19	M	5(1.32%)
20	NLR	6(1.58%)
21	RET	0(0%)
22	RET %	0(0%)
23	LDH	175(46.17%)
24	HCY	187(49.34%)
25	HP	172(45.38%)
26	ALT	25(6.60%)
27	AST	26(6.86%)
28	TBiL	0(0%)
29	DBiL	0(0%)
30	IBiL	0(0%)
31	IBiL/TBiL ratio	0(0%)
32	TP	28(7.39%)
33	ALB	28(7.39%)

34	GLB	28(7.39%)
35	ALP	34(8.97%)
36	GGT	31(8.18%)
37	TBA	35(9.23%)
38	BUN	42(11.08%)
39	CR	42(11.08%)
40	GFR	45(11.87%)
41	UA	50(13.19%)
42	GLU	80(21.11%)
43	TG	78(20.58%)
44	TC	76(20.05%)
45	HDL	78(20.58%)
46	LDL	78(20.58%)
47	APOB1A1	93(24.54%)
48	APOB	93(24.54%)
49	SPLEEN WIDE	76(20.05%)

50	CAP	88(23.22%)
51	KPA	88(23.22%)

ALB: Albumin, ALP: Alkaline Phosphatase, ALT: Alanine Aminotransferase, APOB:Apolipoprotein B, APOB1A1:Apolipoprotein B1A1, AST: Aspartate Aminotransferase, BUN: Blood Urea Nitrogen, CAP: Controlled Attenuation Parameter, COOMBS: Coombs Test (or Direct Antiglobulin Test), CR: Creatinine, DBiL: Direct Bilirubin, GFR: Glomerular Filtration Rate, GGT: Gamma-Glutamyl Transferase, GLB: Globulin, GLU: Glucose, HCT: Hematocrit, HCY:Homocysteine, HDL: High-Density Lipoprotein, HGB: Hemoglobin, HP:Haptoglobin, IBiL: Indirect Bilirubin, KPA: Kilopascals,L:Lymphocyte percentageLDH:Lactate Dehydrogenase, LDL: Low-Density Lipoprotein, M Monocyte percentage, MCH: Mean Corpuscular Hemoglobin, MCHC: Mean Corpuscular Hemoglobin Concentration, MCV: Mean Corpuscular Volume,N:Neutrophil percentage, PLT: Platelet, RBC: Red Blood Cell, RDW: Red Cell Distribution Width, RDWSD: Red Cell Distribution Width - Standard Deviation, RET: Reticulocyte, RET%: Reticulocyte Percentage, TBA: Total Bile Acid, TBiL: Total Bilirubin, TC: Total Cholesterol, TG: Triglycerides, UA: Uric Acid

Supplementary Table 4. Comparison of clinical characteristics between the training and internal validation cohort.

Variables	Total (n = 379)	Internal Validation Cohort(n = 114)	Training Cohort (n = 265)	Statistic	*P*

Variables	Total (n = 379)	Internal Validation Cohort(n = 114)	Training Cohort (n = 265)	Statistic	*P*
Group, n(%)				χ^2=1.22	0.269
UGT1A1 mutation group	317 (83.64)	99 (86.84)	218 (82.26)		
Hemolysis group	62 (16.36)	15 (13.16)	47 (17.74)		
Age, M (Q_1, Q_3)	39.00 (31.00, 51.00)	40.00 (30.00, 52.75)	38.00 (31.00, 51.00)	Z=-0.03	0.979
Sex,Male n(%)	260 (68.60)	78 (68.42)	182 (68.68)	χ^2=0.00	0.960
Cholelithiasis, n(%)	58 (15.30)	16 (14.04)	42 (15.85)	χ^2=0.20	0.653
Splenomegaly, n(%)	76 (20.05)	20 (17.54)	56 (21.13)	χ^2=0.64	0.424
WBC, M (Q_1, Q_3)	5.57 (4.71, 6.53)	5.60 (4.84, 6.65)	5.57 (4.66, 6.49)	Z=-0.26	0.794
PLT, M (Q_1, Q_3)	209.00 (172.00, 243.00)	206.00 (172.50, 238.00)	211.00 (172.00, 245.00)	Z=-0.52	0.604
RBC, M (Q_1, Q_3)	4.81 (4.38, 5.21)	4.83 (4.43, 5.25)	4.77 (4.35, 5.20)	Z=-1.16	0.246
HGB, M (Q_1, Q_3)	147.00 (134.00, 160.00)	150.50 (138.00, 162.00)	146.00 (132.00, 159.00)	Z=-1.73	0.083
HCT, M (Q_1, Q_3)	43.30 (39.30, 46.50)	44.30 (40.68, 46.77)	43.10 (39.00, 46.30)	Z=-1.59	0.112

Variables	Total (n = 379)	Internal Validation Cohort(n = 114)	Training Cohort (n = 265)	Statistic	*P*
MCV, M (Q_1, Q_3)	90.60 (88.10, 93.20)	90.50 (88.25, 93.27)	90.70 (88.00, 93.00)	Z=-0.03	0.978
MCH, M (Q_1, Q_3)	30.90 (30.00, 31.85)	30.90 (30.10, 31.70)	30.90 (30.00, 31.90)	Z=-0.19	0.851
MCHC, M (Q_1, Q_3)	340.00 (334.00, 348.00)	341.00 (334.00, 347.00)	340.00 (334.00, 348.00)	Z=-0.53	0.599
RDW, M (Q_1, Q_3)	12.40 (12.00, 13.00)	12.40 (11.90, 12.70)	12.40 (12.00, 13.10)	Z=-1.25	0.212
RDWSD, M (Q_1, Q_3)	40.90 (39.40, 43.40)	40.55 (39.30, 42.98)	40.90 (39.40, 43.70)	Z=-1.09	0.275
N, M (Q_1, Q_3)	58.50 (51.60, 63.82)	58.55 (53.30, 63.22)	58.10 (50.60, 64.30)	Z=-0.56	0.578
L, M (Q_1, Q_3)	32.80 (27.30, 39.00)	33.15 (28.13, 38.65)	32.70 (26.90, 39.10)	Z=-0.21	0.832
M, M (Q_1, Q_3)	6.10 (5.20, 7.30)	6.20 (5.20, 7.18)	6.10 (5.30, 7.40)	Z=-0.74	0.459
NLR, M (Q_1, Q_3)	1.77 (1.33, 2.35)	1.77 (1.36, 2.25)	1.75 (1.30, 2.38)	Z=-0.04	0.969
RET, M (Q_1, Q_3)	68.80 (57.60, 94.45)	69.20 (57.65, 93.00)	68.50 (57.60, 100.00)	Z=-0.03	0.976
RET %, M (Q_1, Q_3)	1.43 (1.19, 1.96)	1.42 (1.19, 1.83)	1.43 (1.19, 1.98)	Z=-0.46	0.648
ALT, M (Q_1, Q_3)	20.00 (15.00, 28.00)	20.50 (15.00, 28.00)	19.00 (14.00, 28.00)	Z=-0.51	0.611
AST, M (Q_1, Q_3)	23.00 (18.50, 28.85)	22.00 (19.00, 28.00)	23.00 (18.00, 29.00)	Z=-0.89	0.372

Variables	Total (n = 379)	Internal Validation Cohort(n = 114)	Training Cohort (n = 265)	Statistic	*P*
Tbil, M (Q_1, Q_3)	40.00 (30.10, 59.10)	40.35 (28.52, 59.20)	39.70 (30.50, 57.80)	Z=-0.63	0.528
Dbil, M (Q_1, Q_3)	11.30 (8.75, 14.75)	11.15 (8.30, 14.38)	11.30 (8.80, 14.80)	Z=-0.76	0.446
Ibil, M (Q_1, Q_3)	27.10 (20.30, 43.15)	27.55 (19.72, 43.27)	26.90 (20.50, 42.90)	Z=-0.42	0.678
IBiL/TBiL ratio, M (Q_1, Q_3)	0.70 (0.65, 0.76)	0.71 (0.66, 0.77)	0.70 (0.65, 0.76)	Z=-0.44	0.660
TP, M (Q_1, Q_3)	73.70 (70.90, 76.65)	73.90 (71.10, 76.75)	73.50 (70.80, 76.60)	Z=-0.53	0.600
ALB, M (Q_1, Q_3)	46.50 (44.60, 48.30)	46.70 (44.80, 48.60)	46.50 (44.50, 48.10)	Z=-0.88	0.381
GLB, M (Q_1, Q_3)	27.20 (25.10, 29.20)	27.25 (25.22, 29.17)	27.10 (25.10, 29.20)	Z=-0.00	0.997
ALP, M (Q_1, Q_3)	70.83 (58.00, 86.65)	71.00 (60.00, 86.22)	70.00 (58.00, 87.00)	Z=-0.04	0.970
GGT, M (Q_1, Q_3)	17.00 (13.00, 24.72)	16.00 (13.00, 24.00)	17.00 (12.00, 25.00)	Z=-0.17	0.864
TBA, M (Q_1, Q_3)	3.60 (2.30, 6.00)	3.60 (2.50, 5.46)	3.60 (2.20, 6.30)	Z=-0.13	0.899
BUN, M (Q_1, Q_3)	5.00 (4.25, 5.76)	4.88 (4.07, 5.79)	5.02 (4.38, 5.74)	Z=-1.18	0.237
CR, M (Q_1, Q_3)	64.00 (55.00, 72.21)	64.00 (55.25, 72.00)	64.00 (55.00, 72.74)	Z=-0.21	0.835
GFR, M (Q_1, Q_3)	112.61 (103.25, 121.85)	111.40 (104.17, 124.55)	113.30 (103.20, 121.40)	Z=-0.02	0.987

Variables	Total (n = 379)	Internal Validation Cohort(n = 114)	Training Cohort (n = 265)	Statistic	*P*
UA, M (Q_1, Q_3)	348.00 (288.57, 402.50)	348.50 (283.25, 398.93)	346.00 (291.00, 404.00)	Z=-0.43	0.664
GLU, M (Q_1, Q_3)	5.20 (4.96, 5.46)	5.21 (4.89, 5.45)	5.19 (4.99, 5.48)	Z=-0.81	0.415
TG, M (Q_1, Q_3)	1.00 (0.78, 1.40)	1.05 (0.78, 1.32)	0.99 (0.77, 1.42)	Z=-0.02	0.981
TC, M (Q_1, Q_3)	4.17 (3.63, 4.88)	4.27 (3.72, 5.01)	4.14 (3.56, 4.85)	Z=-1.28	0.200
HDL, M (Q_1, Q_3)	1.20 (1.00, 1.49)	1.21 (1.10, 1.53)	1.20 (0.97, 1.45)	Z=-1.77	0.077
LDL, M (Q_1, Q_3)	2.44 (1.85, 3.09)	2.50 (1.91, 3.10)	2.41 (1.82, 3.09)	Z=-0.82	0.410
APOB1A1, M (Q_1, Q_3)	1.82 (1.61, 2.07)	1.83 (1.69, 2.09)	1.82 (1.55, 2.06)	Z=-1.09	0.277
APOB, M (Q_1, Q_3)	0.76 (0.63, 0.95)	0.76 (0.62, 0.94)	0.76 (0.64, 0.96)	Z=-0.25	0.804
SPLEEN WIDE, M (Q_1, Q_3)	33.55 (30.87, 39.00)	33.15 (30.54, 37.00)	33.60 (31.00, 39.00)	Z=-0.52	0.604
CAP, M (Q_1, Q_3)	236.02 (206.39, 265.35)	239.84 (213.23, 267.50)	235.43 (205.00, 263.00)	Z=-0.81	0.419
KPA, M (Q_1, Q_3)	5.10 (4.40, 6.10)	5.10 (4.12, 6.10)	5.10 (4.50, 6.16)	Z=-0.53	0.596
COOMBS, n(%)				χ^2=0.24	0.625
Negative	352 (92.88)	107 (93.86)	245 (92.45)		

Variables	Total (n = 379)	Internal Validation Cohort(n = 114)	Training Cohort (n = 265)	Statistic	*P*
Positive	27 (7.12)	7 (6.14)	20 (7.55)		
Abnormal peripheral blood smear, n(%)				χ^2=0.71	0.400
Normal	338 (89.18)	104 (91.23)	234 (88.30)		
Abnormal	41 (10.82)	10 (8.77)	31 (11.70)		

Z: Mann-Whitney test, χ^2: Chi-square test

M: Median, Q_1: 1st Quartile, Q_3: 3st Quartile

ALB: Albumin, ALP: Alkaline Phosphatase, ALT: Alanine Aminotransferase, APOB:Apolipoprotein B, APOB1A1:Apolipoprotein B1A1, AST: Aspartate Aminotransferase, BUN: Blood Urea Nitrogen, CAP: Controlled Attenuation Parameter, COOMBS: Coombs Test (or Direct Antiglobulin Test), CR: Creatinine, DBiL: Direct Bilirubin, GFR: Glomerular Filtration Rate, GGT: Gamma-Glutamyl Transferase, GLB: Globulin, GLU: Glucose, HCT: Hematocrit, HCY:Homocysteine, HDL: High-Density Lipoprotein, HGB: Hemoglobin, HP:Haptoglobin, IBiL: Indirect Bilirubin, KPA: Kilopascals,L:Lymphocyte percentageLDH:Lactate Dehydrogenase, LDL: Low-Density Lipoprotein, M Monocyte percentage, MCH: Mean Corpuscular

Hemoglobin, MCHC: Mean Corpuscular Hemoglobin Concentration, MCV: Mean Corpuscular Volume,N:Neutrophil percentage, PLT: Platelet, RBC: Red Blood Cell, RDW: Red Cell Distribution Width, RDWSD: Red Cell Distribution Width - Standard Deviation, RET: Reticulocyte, RET%: Reticulocyte Percentage, TBA: Total Bile Acid, TBiL: Total Bilirubin, TC: Total Cholesterol, TG: Triglycerides, UA: Uric Acid

Supplementary Table 5. Comparison of clinical characteristics between patients with *UGT1A1*-associated and hemolytic disease-associated unconjugated hyperbilirubinemia in the external validation cohorts.

Variables	Total (n = 50)	UGT1A1 mutation group (n = 40)	Hemolysis group (n = 10)	Statistic	*P*
Age, M (Q_1, Q_3)	45.00 (35.50, 55.75)	42.50 (34.25, 55.00)	54.50 (43.50, 67.00)	Z=-1.60	0.109
Sex,Male n(%)	34 (68.00)	27 (67.50)	7 (70.00)	χ^2=0.00	1.000
Cholelithiasis, n(%)	14 (28.00)	6 (15.00)	8 (80.00)	χ^2=13.70	**<0.001**
Splenomegaly, n(%)	13 (26.00)	4 (10.00)	9 (90.00)	χ^2=22.62	**<0.001**
WBC, M (Q_1, Q_3)	5.29 (4.58, 6.11)	5.25 (4.54, 6.11)	5.44 (5.21, 6.37)	Z=-1.03	0.303

Variables	Total (n = 50)	UGT1A1 mutation group (n = 40)	Hemolysis group (n = 10)	Statistic	*P*
PLT, M (Q_1, Q_3)	204.50 (158.75, 243.75)	208.50 (164.00, 245.00)	195.50 (153.25, 221.50)	Z=-0.74	0.459
RBC, M (Q_1, Q_3)	4.58 (4.20, 5.14)	4.74 (4.35, 5.22)	3.44 (3.19, 4.08)	Z=-3.92	**<0.001**
HGB, M (Q_1, Q_3)	144.50 (132.50, 155.75)	149.00 (138.25, 159.75)	121.50 (102.25, 133.00)	Z=-3.74	**<0.001**
HCT1, M (Q_1, Q_3)	42.45 (38.55, 45.55)	43.20 (40.67, 47.10)	34.95 (29.12, 38.57)	Z=-4.10	**<0.001**
MCV, M (Q_1, Q_3)	90.90 (87.57, 94.38)	89.95 (86.90, 93.85)	94.50 (92.15, 100.75)	Z=-2.07	**0.038**
MCH, M (Q_1, Q_3)	31.15 (29.38, 32.65)	31.00 (29.20, 31.75)	33.80 (31.93, 35.27)	Z=-2.74	**0.006**
MCHC, M (Q_1, Q_3)	343.00 (334.00, 351.00)	341.00 (333.00, 348.50)	350.00 (342.25, 356.25)	Z=-2.02	**0.043**

Variables	Total (n = 50)	UGT1A1 mutation group (n = 40)	Hemolysis group (n = 10)	Statistic	*P*
RDWSD, M (Q_1, Q_3)	41.35 (39.45, 45.30)	40.50 (39.18, 43.00)	55.90 (53.60, 57.38)	Z=-3.89	**<0.001**
N, M (Q_1, Q_3)	57.30 (54.52, 62.12)	55.50 (54.08, 58.83)	65.95 (61.85, 76.38)	Z=-3.94	**<0.001**
L, M (Q_1, Q_3)	33.70 (29.05, 37.18)	35.55 (30.55, 37.83)	24.55 (15.00, 30.18)	Z=-3.74	**<0.001**
M, M (Q_1, Q_3)	5.95 (5.43, 7.00)	5.90 (5.47, 7.00)	6.15 (5.12, 7.20)	Z=-0.38	0.707
NLR, M (Q_1, Q_3)	1.70 (1.46, 2.21)	1.56 (1.42, 1.92)	2.70 (2.02, 5.47)	Z=-3.78	**<0.001**
RET%, M (Q_1, Q_3)	1.60 (1.38, 2.33)	1.52 (1.28, 1.82)	7.90 (4.36, 9.40)	Z=-4.84	**<0.001**
ALT, M (Q_1, Q_3)	20.00 (15.00, 33.75)	21.50 (15.00, 35.75)	19.50 (18.00, 29.25)	Z=-0.12	0.903
AST, M (Q_1, Q_3)	24.00 (19.25, 28.00)	24.00 (19.00, 27.25)	25.50 (21.50, 38.50)	Z=-1.03	0.302

Variables	Total (n = 50)	UGT1A1 mutation group (n = 40)	Hemolysis group (n = 10)	Statistic	*P*
TBiL, M (Q_1, Q_3)	43.55 (30.15, 62.83)	39.15 (27.55, 51.40)	69.60 (57.25, 86.70)	Z=-3.24	**<0.001**
DBiL, M (Q_1, Q_3)	10.75 (8.43, 15.07)	9.45 (8.07, 13.95)	17.55 (12.12, 31.43)	Z=-3.06	**0.002**
IBiL, M (Q_1, Q_3)	28.20 (21.30, 45.95)	27.30 (20.00, 37.60)	49.85 (33.20, 75.90)	Z=-2.84	**0.005**
IBiL/TBiL, M (Q_1, Q_3)	0.71 (0.65, 0.75)	0.70 (0.66, 0.75)	0.72 (0.64, 0.81)	Z=-0.11	0.914
TP, M (Q_1, Q_3)	73.10 (71.08, 76.27)	73.40 (71.22, 76.23)	72.70 (70.35, 76.17)	Z=-0.06	0.952
ALB, M (Q_1, Q_3)	45.55 (44.70, 47.38)	45.55 (44.70, 47.32)	45.25 (43.80, 47.55)	Z=-0.28	0.780
GLB, M (Q_1, Q_3)	27.50 (26.00, 30.17)	27.50 (26.30, 30.13)	28.20 (24.80, 30.45)	Z=-0.10	0.923
ALP, M (Q_1, Q_3)	65.00 (53.25,	65.00 (52.75, 78.25)	73.00 (59.50,	Z=-0.91	0.363

Variables	Total (n = 50)	UGT1A1 mutation group (n = 40)	Hemolysis group (n = 10)	Statistic	*P*
	82.00)		105.25)		
GGT, M (Q_1, Q_3)	20.00 (15.25, 29.00)	19.00 (14.75, 26.25)	31.50 (21.75, 33.75)	Z=-2.31	**0.021**
TBA, M (Q_1, Q_3)	4.30 (2.60, 5.55)	4.15 (2.40, 5.38)	5.35 (4.25, 10.35)	Z=-1.70	0.089
BUN, M (Q_1, Q_3)	5.00 (4.24, 6.21)	5.00 (4.33, 6.25)	5.13 (4.15, 5.71)	Z=-0.68	0.497
CR, M (Q_1, Q_3)	62.00 (52.50, 67.75)	65.00 (54.00, 70.25)	54.50 (49.00, 61.75)	Z=-2.09	**0.037**
GFR, M (Q_1, Q_3)	109.15 (99.00, 117.97)	109.55 (99.45, 117.93)	103.95 (98.93, 122.23)	Z=-0.10	0.923
UA, M (Q_1, Q_3)	301.00 (260.25, 385.75)	306.50 (259.25, 391.25)	278.00 (261.50, 380.00)	Z=-0.13	0.894
GLU, M (Q_1, Q_3)	5.21 (4.94, 5.47)	5.20 (4.92, 5.38)	5.25 (5.21, 6.15)	Z=-1.21	0.225
TG, M (Q_1, Q_3)	0.96 (0.78, 1.21)	0.94 (0.73, 1.20)	1.26 (0.98, 1.65)	Z=-1.98	**0.048**

Variables	Total (n = 50)	UGT1A1 mutation group (n = 40)	Hemolysis group (n = 10)	Statistic	*P*
TC, M (Q_1, Q_3)	4.44 (3.56, 4.74)	4.48 (3.83, 4.83)	3.76 (3.27, 4.50)	Z=-1.89	0.058
HDL, M (Q_1, Q_3)	1.18 (1.00, 1.44)	1.25 (1.12, 1.54)	0.98 (0.62, 1.18)	Z=-2.65	**0.008**
LDL, M (Q_1, Q_3)	2.31 (1.95, 2.85)	2.37 (2.00, 2.77)	2.31 (1.46, 2.85)	Z=-0.74	0.459
APOB1A1, M (Q_1, Q_3)	1.70 (1.48, 1.92)	1.72 (1.63, 1.95)	1.45 (1.07, 1.70)	Z=-2.11	**0.034**
APOB, M (Q_1, Q_3)	0.80 (0.65, 0.90)	0.80 (0.65, 0.86)	0.80 (0.63, 0.93)	Z=-0.10	0.923
CAP, M (Q_1, Q_3)	230.00 (214.25, 255.25)	230.00 (204.75, 241.00)	230.00 (230.00, 275.25)	Z=-1.50	0.134
KPA, M (Q_1, Q_3)	4.65 (4.50, 5.57)	4.65 (4.38, 5.50)	4.95 (4.50, 6.00)	Z=-1.18	0.237
COOMBS, n(%)				-	1.000
0	46 (92.00)	37 (92.50)	9 (90.00)		
1	4 (8.00)	3 (7.50)	1 (10.00)		
Abnormal peripheral blood				-	**0.013**

Variables	Total (n = 50)	UGT1A1 mutation group (n = 40)	Hemolysis group (n = 10)	Statistic	P
smear, n(%)					
0	43 (91.49)	38 (97.44)	5 (62.50)		
1	4 (8.51)	1 (2.56)	3 (37.50)		

Z: Mann-Whitney test, χ^2: Chi-square test, -: Fisher exact

M: Median, Q_1: 1st Quartile, Q_3: 3st Quartile

ALB: Albumin, ALP: Alkaline Phosphatase, ALT: Alanine Aminotransferase, APOB:Apolipoprotein B, APOB1A1:Apolipoprotein B1A1, AST: Aspartate Aminotransferase, BUN: Blood Urea Nitrogen, CAP: Controlled Attenuation Parameter, COOMBS: Coombs Test (or Direct Antiglobulin Test), CR: Creatinine, DBiL: Direct Bilirubin, GFR: Glomerular Filtration Rate, GGT: Gamma-Glutamyl Transferase, GLB: Globulin, GLU: Glucose, HCT: Hematocrit, HCY:Homocysteine, HDL: High-Density Lipoprotein, HGB: Hemoglobin, HP:Haptoglobin, IBiL: Indirect Bilirubin, KPA: Kilopascals,L:Lymphocyte percentageLDH:Lactate Dehydrogenase, LDL: Low-Density Lipoprotein, M Monocyte percentage, MCH: Mean Corpuscular Hemoglobin, MCHC: Mean Corpuscular Hemoglobin Concentration, MCV: Mean Corpuscular Volume,N:Neutrophil percentage, PLT: Platelet, RBC: Red Blood Cell, RDW: Red Cell Distribution Width, RDWSD: Red Cell Distribution Width - Standard

Deviation, RET: Reticulocyte, RET%: Reticulocyte Percentage, TBA: Total Bile Acid, TBiL: Total Bilirubin, TC: Total Cholesterol, TG: Triglycerides, UA: Uric Acid

Supplementary Table 6. Assessment of multicollinearity using variance inflation factors for variables selected by LASSO regression.

	Abnormal peripheral blood smear	Splenomegaly	HCT	RDWSD	RET %	IBiL/TBiL ratio
VIP	2.293954	1.856931	2.183073	4.214209	3.3355720	1.189163

HCT: Hematocrit, RDWSD: Red cell distribution width-standard deviation, RET: Reticulocyte, IBiL: Indirect bilirubin, TBiL: Total bilirubin.

Supplementary Table 7. The diagnostic performance indicators of the nomogram model across the training, internal validation, and external validation cohorts.

	AUC (95%CI)	Accuracy (95%CI)	Sensitivity (95%CI)	Specificity (95%CI)	PPV (95%CI)	NPV (95%CI)	F1 (95%CI)	AUPRC (95%CI)
Training cohort	0.994	0.981	0.957	0.986	0.937	0.991	0.947	0.978

	(0.985-1.000)	(0.961-0.999)	(0.895-1.000)	(0.934-1.000)	(0.759-1.000)	(0.977-1.000)	(0.852-0.989)	(0.95-1.000)
Internal validation cohort	0.958 (0.877-1.000)	0.974 (0.901-0.999)	0.933 (0.727-1.000)	0.980 (0.95-1.000)	0.875 (0.722-1.000)	0.990 (0.968-1.000)	0.903 (0.8-1.000)	0.943 (0.829-1.000)
External validation cohort	0.986 (0.963-1.000)	0.920 (0.808-0.978)	1.000 (0.867-1.000)	0.900 (0.794-1.000)	0.714 (0.441-1.000)	1.000 (0.971-1.000)	0.833 (0.667-1.000)	0.938 (0.738-1.000)

AUC:Area under the Receiver Operating Characteristic curve, AUPRC:Area under the Precision-Recall curve, NPV:Negative predictive value, PPV:Positive predictive value.