

CLINICAL RESEARCH ARTICLE



# A new scoring system for coronary artery abnormalities in Kawasaki disease

Hui Hui Liu<sup>1</sup>, Wei Xia Chen<sup>1</sup>, Man Man Niu<sup>1</sup>, Qi Jiang<sup>1</sup>, Zhen Qiu<sup>1</sup>, Guo Zhen Fan<sup>1</sup>, Rui Xue Li<sup>1</sup>, Goshgar Mammadov<sup>1</sup>, Yang Fang Wu<sup>1</sup>, Huang Huang Luo<sup>1</sup>, Dong Dong Zhang<sup>1</sup> and Peng Hu<sup>1</sup>✉

© The Author(s), under exclusive licence to the International Pediatric Research Foundation, Inc 2021

**BACKGROUND:** In China, coronary artery abnormalities (CAAs) secondary to Kawasaki disease (KD) tend to have an increased occurrence. We hypothesize that Chinese children with KD may possess several unique CAA risks, and the predictive efficacy of multiple scoring systems in Chinese patients are still to be further studied.

**METHODS:** Two hundred and three KD patients were recruited. Using multivariable analysis, independent predictors of CAAs were combined into a scoring system. Subsequently, CAA risks of our patients were evaluated by the newly established scoring system and eight other published scoring systems.

**RESULTS:** Seventeen (8.37%) KD patients were identified as CAAs. The newly established scoring system contained the following 5 independent predictors: days of illness at initial treatment  $\geq 7$ , redness and swelling of extremities, hematocrit  $\leq 33\%$ , percentage of monocytes  $\geq 8.89\%$ , and procalcitonin  $\geq 0.5$  ng/mL. The AUC value of newly established scoring system was 0.685 with a sensitivity of 41.18% and a specificity of 84.41%, higher than Harada score, Egami score, Kobayashi score, Sato score, San Diego score, Formosa score, and Tang score, whereas lower than Hua score.

**CONCLUSIONS:** Days of illness at initial treatment  $\geq 7$  and procalcitonin are unique predictors of CAAs in newly established scoring system. Taking into account different identification criteria and analytical methodologies, there is still some heterogeneity among different scoring systems.

*Pediatric Research* (2022) 92:275–283; <https://doi.org/10.1038/s41390-021-01752-8>

**IMPACT:**

- The newly established scoring system contains the five independent predictors.
- Days of illness at initial treatment  $\geq 7$  and PCT are unique predictors of CAAs in our study, compared with 8 other systems.
- The AUC value of newly established scoring system is 0.685, similar to Hua score.
- There is some heterogeneity among different scoring systems.

**INTRODUCTION**

Kawasaki disease (KD) is a systemic vasculitis of childhood. According to a recent epidemiological study in Shanghai, China, the average annual incidence rate of KD was 50.5 per 100,000 children.<sup>1</sup> Coronary artery abnormalities (CAAs) secondary to KD represent the major contributors to morbidity and mortality in both the acute stage and long term. Although intravenous immunoglobulin (IVIG) plus aspirin is the mainstay of KD treatment, approximately 5% of patients still developed CAAs. In a study of long-term outcomes during an average follow-up period of 14.6 years, Holve et al.<sup>2</sup> observed that, of the 546 KD patients, 12.8% had CAAs in the acute stage, and moreover 5% developed persistent CAAs at least 1 year after the KD illness.

In the past decade, several epidemiological surveys and meta-analyses have found some risk factors possibly related to CAAs. A recent retrospective study encompassing 1016 KD patients from China identified that male, albumin, erythrocyte sedimentation rate (ESR), *Mycoplasma pneumoniae* infection, IVIG started after the

tenth day of illness, and IVIG non-responders were predictive of CAAs.<sup>3</sup> Based on the data from 92 studies encompassing 8330 KD patients, a meta-analysis by Chen et al.<sup>4</sup> indicated that platelet count (PLT), neutrophils count, platelet hematocrit, platelet distribution width, mean platelet volume, ESR, cardiac troponin I (cTnI), endothelin-1, albumin, and hemoglobin may be the potential candidates for CAA predictors. To date, about 10 scoring systems have been formulated to predict CAA risks. However, the criteria in these scoring systems are inconsistent with the findings of previous studies completely and have significant variations in predictive efficacy among different populations. Several risk factors, such as age, C-reactive protein (CRP), and fever duration are commonly included in multiple scoring systems, whereas their associations with CAAs are seldom confirmed by previous studies. Recently, Kobayashi score was applied to predict CAA risks in non-Japanese populations; the sensitivities were 50, 60, and 15% in Germany,<sup>5</sup> British,<sup>6</sup> and Polish<sup>7</sup> respectively, far lower than that in Japanese (90.6%).<sup>8</sup> On the other hand, Fabi et al.<sup>9</sup> used Kobayashi,

<sup>1</sup>Department of Pediatrics, The First Affiliated Hospital of Anhui Medical University, No. 218 Ji-Xi Road, 230022 Hefei, PR China. ✉email: hupeng28@aliyun.com

Received: 12 April 2021 Revised: 9 August 2021 Accepted: 31 August 2021  
Published online: 28 September 2021

Egami, and Formosa score to evaluate CAA risks in Italian simultaneously and found that Formosa score had the highest predictive efficacy with a sensitivity of 72.0% and a specificity of 45.5%, followed by Egami and Kobayashi scores.

In China, CAAs tend to have an increased occurrence relative to the other Asian countries. Several previous epidemiological surveys covering 2 decades indicated that the CAA morbidities were 15.9% in Shanghai,<sup>1</sup> 20.6% in Beijing,<sup>10</sup> and 24.1% in Hangzhou.<sup>11</sup> Given this background, we hypothesize that Chinese KD patients may possess several unique risk factors of CAAs, and the predictive efficacy of multiple scoring systems in Chinese patients are still to be further studied.

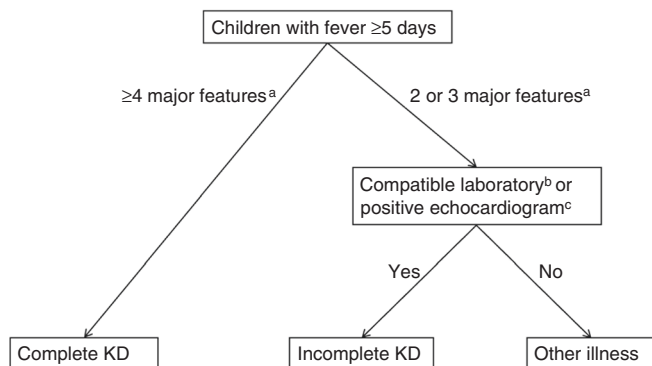
## METHODS

### Subjects

The present retrospective study included a total of 203 KD children between July 2015 and March 2019 in our center (Fig. 1). There was no fatal case during the observational period. According to the 2017 American Heart Association (AHA) guidelines,<sup>12</sup> the diagnostic criteria for complete KD comprise the presence of  $\geq 5$  days of fever and  $\geq 4$  of the following 5 major features: (1) bilateral conjunctival injection without exudates; (2) changes in the oral mucosa, such as erythema and cracking lips, erythema of the pharynx, strawberry tongue; (3) changes in extremities, such as redness and swelling in the acute phase, periungual desquamation in subacute phase; (4) polymorphous exanthema; and (5) cervical lymphadenopathy ( $\geq 1.5$  cm in diameter), usually unilateral. Patients with fever for  $\geq 5$  days and  $\geq 2$  of the major features with compatible laboratory or echocardiographic findings can be diagnosed as incomplete KD, if no other disease processes could explain the illness. The first day of the illness is defined as the first day of fever. All patients received the standard therapy for KD, including a single infusion of IVIG (2 g/kg) and aspirin (30–50 mg/kg/day). IVIG non-responsiveness is defined as persistent or recrudescence fever  $\geq 36$  h after completion of the initial IVIG infusion. In this circumstance, a second dose of IVIG would be given alone or combined with corticosteroids simultaneously. Delayed diagnosis of KD is defined as days of initial IVIG treatment  $> 10$  days. Echocardiography was performed in the acute episode of KD. Two echocardiographic evaluative criteria of CAAs are adopted in the present study. One proposed by the Japanese Ministry of Health (JMH) classifies coronary arteries as abnormality if the internal lumen diameter is  $> 3.0$  mm in children  $< 5$  years of age or  $> 4.0$  mm in children  $\geq 5$  years of age, if the internal diameter of a segment measures  $\geq 1.5$  times that of an adjacent segment, or if the coronary artery lumen is clearly irregular. The other defines CAAs on the basis of a coronary artery Z score (standard deviation units from the mean, normalized for body surface area)  $\geq 2.5$ . In the present study, CAAs could be identified if any patient fulfilled either of the above two criteria.

### Data collection

The investigation conforms to the principles outlined in the Declaration of Helsinki. Approval for this research was acquired from our medical ethic committee (No. LLSC20150009) and the informed consent was obtained from the parents of each patient. Demographic, clinical characteristics, electrocardiography, and echocardiography were collected. Laboratory



**Fig. 1 Patient selection.** CAA coronary artery abnormality, KD Kawasaki disease.

testing within 24 h pre-IVIG therapy included white blood cell counts with differential, hemoglobin, hematocrit (Hct), PLT, alanine aminotransferase, aspartate aminotransferase, total bilirubin, direct bilirubin, indirect bilirubin,  $\gamma$ -glutamyltransferase, lactate dehydrogenase, cTnI, creatine kinase MB (CK-MB), albumin, sodium, CRP, ESR, procalcitonin (PCT), interleukin-6, and tumor necrosis factor- $\alpha$ .

### Scoring systems for CAAs

Eight scoring systems were adopted in the present study. The Harada score (Japan),<sup>13</sup> Egami score (Japan),<sup>14</sup> Kobayashi score (Japan),<sup>8</sup> Sato score (Japan),<sup>15</sup> San Diego score (the United States),<sup>16</sup> Formosa score (China),<sup>17</sup> Tang score (China),<sup>18</sup> and Hua score (China)<sup>11</sup> are shown in Table 1. Besides predicting IVIG-nonresponsive KD, these scoring systems also identified CAA risks in their KD patients simultaneously. Their sensitivity and specificity for CAAs are also depicted in Table 1.

### Statistical analysis

Continuous data were expressed as mean  $\pm$  SD or median (interquartile range). Categorical variables were presented as count with percentage. Data were analyzed using two independent samples *t* test or Mann–Whitney *U* test for continuous variables and Chi-squared test or Fisher's exact test for categorical variables. A  $p < 0.05$  was considered significant. Risk factors for CAAs selected by univariate analysis were considered for multivariable analysis. Results are expressed as an odds ratio with a 95% confidence interval (CI). Goodness of fit of the regression model was tested with Hosmer–Lemeshow test, with  $p > 0.05$  considered to indicate lack of deviation between the predictive model and the sample data. To facilitate clinical application, continuous variables were turned into dichotomous variables. Cut-off values were chosen based on the receiver operating characteristic (ROC) curve of each variable, which were dependent on the maximum value of Youden index according to the sensitivity and specificity and were justified medically. The score point of each predictor was assigned based on the value of logistic coefficient. ROC curve and the area under the curve (AUC) were used to assess the predictive efficacy, sensibility, and specificity of these scoring systems. Statistical analyses were performed using the statistical package for social studies SPSS ver 16.0.

## RESULTS

### Patient characteristics

The demographic features of KD patients are shown in Table 2. In the present study, 203 patients with KD were included with a mean age of  $35.15 \pm 28.95$  months and a range from 3 months to 13 years. Of the 203 KD patients, 167 (82.27%) were younger than 5 years, 62 (30.54%) were younger than 1 year, and 16 (7.88%) were younger than 6 months. Males outnumbered females by 1.28:1.

**Clinical classification.** All 5 major features for KD were met in 59 cases (29.06%), 4 features in 80 (39.40%), 3 features in 38 (18.72%), and 2 features in 20 (9.85%). Therefore, 139 patients (68.47%) had complete KD, including 83 males and 56 females with the mean age of  $36.46 \pm 29.25$  months; 64 patients (31.53%) had incomplete KD, including 31 males and 33 females with the mean age of  $32.31 \pm 28.33$  months. The mean age and male/female ratio were almost identical in the 2 groups ( $p > 0.05$ ).

**IVIG therapy.** Among the 203 KD patients, 11 (7.75%) of them were identified as IVIG nonresponders and had persistent fever for about  $71.00 \pm 12.27$  h after IVIG therapy, including 4 males and 7 females with the mean age of  $33.73 \pm 27.46$  months. In contrast, 192 patients were identified as IVIG responders, including 110 males and 82 females with the mean age of  $35.23 \pm 29.10$  months. The mean age and male/female ratio were no significant differences between IVIG responders and IVIG nonresponders ( $p > 0.05$ ).

### Cardiovascular findings

Based on the 2017 AHA guidelines, 17 (8.37%) KD patients were identified as CAAs (left coronary artery:  $3.16 \pm 0.57$  mm; right coronary artery:  $3.34 \pm 1.35$  mm), including 11 males and 6 females

**Table 1.** Eight scoring systems to predict CAAs in the present study.

Nation	Scoring systems	Risk factors	Points	Predicted risk (score)	Sensitivity	Specificity
Japan	Harada score <sup>13</sup>	WBC >12 × 10 <sup>9</sup> /L	1	Low risk (0–3) High risk (≥4)	91%	30%
		PLT <350 × 10 <sup>9</sup> /L	1			
		CRP >30 mg/L	1			
		Hct <35%	1			
		Alb <35 g/L	1			
		Age ≤12 months	1			
		Male	1			
	Egami score <sup>14</sup>	Days of illness at initial treatment ≤4	1	Low risk (0–2) High risk (≥3)	61%	81%
		Age <6 months	1			
		PLT ≤300 × 10 <sup>9</sup> /L	1			
		CRP ≥80 mg/L	1			
		ALT ≥80 IU/L	2			
	Kobayashi score <sup>8</sup>	Days of illness at initial treatment ≤4	2	Low risk (0–3) High risk (≥4)	91%	60%
		Sodium ≤133 mmol/L	2			
		AST ≥100 IU/L	2			
N% ≥80%		2				
CRP ≥100 mg/L		1				
Age ≤12 months		1				
PLT ≤300 × 10 <sup>9</sup> /L		1				
Sato score <sup>15</sup>	Serum IL-6 ≥70 pg/mL but <140 pg/mL	1	Low risk (0–2) High risk (≥3)	69%	70%	
	Serum IL-6 ≥140 pg/mL	2				
	N% ≥75%	2				
The United States	San Diego score <sup>16</sup>	Days of illness at initial treatment ≤4	1	Low risk (0–1) High risk (≥2)	72%	58%
		Age-adjusted hemoglobin ≤−2	1			
		GGT ≥60 IU/L	1			
		Bands ≥20%	2			
China	Formosa score <sup>17</sup>	Alb <35 g/L	1	Low risk (0–2) High risk (≥3)	33%	76%
		N% ≥60%	2			
		Positive lymphadenopathy	1			
	Tang score <sup>18</sup>	Redness and swelling of extremities	1	Low risk (0–2) High risk (≥3)	59%	79%
		Rash	1			
		N% ≥80%	1			
		Age <6 months	2			
		Alb <35 g/L	2			
	Hua score <sup>11</sup>	Male	1	Low risk (0–2) High risk (≥3)	51%	68%
		Total fever duration ≥8 days	1			
		IVIG nonresponse	1	Low risk (0–2) High risk (≥3)	65%	81%
		Alb ≤35.9 g/L	1			
MO% ≥5.9%		1				
If the patient is ≤6 months old		2				
Total fever duration ≥8 days		2				
Delayed diagnosis		1				
Alb ≤35.9 g/L	1					

Age-adjusted hemoglobin = [(observed hemoglobin] – [mean hemoglobin for age])/standard deviation for age.

Alb albumin, ALT alanine aminotransferase, AST aspartate aminotransferase, CAA coronary artery abnormality, CRP C-reactive protein, GGT  $\gamma$ -glutamyltransferase, Hct hematocrit, IL-6 interleukin-6, IVIG intravenous immunoglobulin, MO% percentage of monocytes, N% percentage of neutrophils, PLT platelet count, WBC white blood cells counts.

with the mean age of 38.47 ± 33.78 months. In contrast, 186 (91.63%) patients had normal coronary artery (left coronary artery: 2.07 ± 0.37 mm, right coronary artery: 1.99 ± 0.32 mm), including 103 males and 83 females with the mean age of 34.85 ± 28.56 months.

By echocardiography, the other abnormalities included tricuspid regurgitation (112 cases, 55.17%), mitral regurgitation (77 cases, 37.93%), aortic regurgitation (1 case, 0.49%), and pericardial effusion (14 cases, 6.90%). By electrocardiography, abnormal changes were observed in 11 patients (5.41%) who dominantly suffered from ST-T changes (7 patients), prolonged PR (2 patients), ventricular premature beat (1 patient), and pre-excitation syndrome (1 patient). In addition, 33 patients had acute myocarditis, featured by elevated CK-MB (37.38 ± 13.03  $\mu$ L from 32 cases) or cTnI (0.53 ng/mL from 1 case).

### The establishment of logistic regression model for CAA predictors

The clinical and laboratory characteristics of KD patients are shown in Table 2. For clinical characteristics, significant differences were found in days of illness at initial treatment and redness and swelling of extremities between KD patients with and without CAAs; for laboratory data, percentage of monocytes (MO%), CRP, and PCT was significantly elevated, whereas Hct was lower between KD patients with and without CAAs ( $p < 0.05$ ). Therefore, the above six variables were subsequently included in the multivariable logistic regression analysis. Days of illness at initial treatment, redness and swelling of extremities, Hct, MO%, and PCT were independent predictors of CAAs ( $p < 0.05$ ). The Hosmer–Lemeshow statistic was not significant ( $p > 0.05$ ), and the AUC value of this regression model was 0.770 (95% CI 0.590–0.950).

**Table 2.** Demographic, clinical and laboratory characteristics of KD patients with and without CAAs.

Variables	Total (n=203)	CAAs (n=17)	Without CAAs (n=186)	z value	t value	$\chi^2$ value	p value
Age, mean $\pm$ SD, months	35.15 $\pm$ 28.95	38.47 $\pm$ 33.78	34.85 $\pm$ 28.56	-0.19			0.849 <sup>a</sup>
$\leq$ 5 years, n (%)	167 (82.27)	12 (70.59)	155 (83.33)			0.971	0.325 <sup>b</sup>
$\leq$ 12 months, n (%)	62 (30.54)	5 (29.41)	57 (30.65)			0.011	0.916 <sup>b</sup>
$\leq$ 6 months, n (%)	16 (7.88)	2 (11.76)	14 (7.53)			0.023	0.880 <sup>b</sup>
Male, n (%)	114 (56.16)	11 (64.71)	103 (55.38)			0.551	0.458 <sup>b</sup>
Days of illness at initial treatment, mean $\pm$ SD, days	6.78 $\pm$ 1.75	8.19 $\pm$ 2.11	6.64 $\pm$ 1.66	-2.782			0.004 <sup>a*</sup>
Delayed diagnosis, n (%)	3 (1.48)	2 (11.76)	1 (0.54)				0.200 <sup>c</sup>
Clinical characteristics							
Total fever duration, mean $\pm$ SD, days	6.84 $\pm$ 1.84	8.19 $\pm$ 2.11	6.64 $\pm$ 1.66	-1.614			0.107 <sup>a</sup>
Redness and swelling of extremities, n (%)	126 (62.07)	8 (47.06)	118 (63.44)			4.209	0.040 <sup>b*</sup>
Rash, n (%)	149 (73.40)	13 (76.47)	136 (73.12)			0.000	1.000 <sup>b</sup>
Conjunctivitis, n (%)	165 (81.28)	15 (88.24)	150 (80.65)			0.074	0.785 <sup>b</sup>
Oral changes, n (%)	168 (82.76)	13 (76.47)	155 (83.33)			0.428	0.513 <sup>b</sup>
Cervical lymphadenopathy, n (%)	142 (69.95)	11 (64.71)	131 (70.43)			0.125	0.724 <sup>b</sup>
Number of symptoms, mean $\pm$ SD	3.82 $\pm$ 1.05	3.82 $\pm$ 1.24	3.82 $\pm$ 1.04	-0.267			0.789 <sup>a</sup>
Clinical classification							
Complete, n (%)	139 (68.47)	10 (58.82)	129 (69.35)			0.800	0.371 <sup>b</sup>
Incomplete, n (%)	64 (31.53)	7 (41.18)	57 (30.65)				
IVIG therapy							
Responder, n (%)	192 (94.58)	15 (88.24)	177 (95.16)				0.232 <sup>c</sup>
Nonresponder, n (%)	11 (5.41)	2 (11.76)	9 (4.84)				
WBC, mean $\pm$ SD, $\times 10^9/L$	13.56 $\pm$ 5.05	13.14 $\pm$ 4.73	13.59 $\pm$ 5.09		0.354		0.724 <sup>d</sup>
N%, mean $\pm$ SD	65.53 $\pm$ 14.75	65.89 $\pm$ 14.62	61.66 $\pm$ 16.02		1.133		0.259 <sup>d</sup>
L%, mean $\pm$ SD	26.08 $\pm$ 12.63	27.49 $\pm$ 12.82	25.95 $\pm$ 12.64		-0.450		0.653 <sup>d</sup>
MO%, mean $\pm$ SD	5.84 $\pm$ 2.73	7.08 $\pm$ 3.41	5.73 $\pm$ 2.65		-2.287		0.023 <sup>d*</sup>
EO%, median (IQR)	1.30 (0.34, 3.43)	2.10 (0.29, 3.94)	1.27 (0.34, 3.39)	-0.346			0.729 <sup>a</sup>
Neutrophil count, mean $\pm$ SD, $\times 10^9/L$	9.14 $\pm$ 4.70	8.42 $\pm$ 4.41	9.21 $\pm$ 4.73		0.665		0.507 <sup>d</sup>
Lymphocyte count, mean $\pm$ SD, $\times 10^9/L$	3.38 $\pm$ 1.90	3.32 $\pm$ 1.78	3.38 $\pm$ 1.91	-0.055			0.956 <sup>a</sup>
PLT, mean $\pm$ SD, $\times 10^9/L$	350.31 $\pm$ 121.86	361.53 $\pm$ 143.20	349.26 $\pm$ 120.01		-0.396		0.692 <sup>d</sup>
NLR, mean $\pm$ SD	3.78 $\pm$ 3.44	3.85 $\pm$ 3.26	3.78 $\pm$ 3.37	-0.666			0.505 <sup>a</sup>
PLR, mean $\pm$ SD	134.77 $\pm$ 80.19	138.20 $\pm$ 85.25	134.46 $\pm$ 79.99	-0.143			0.886 <sup>a</sup>
HB, mean $\pm$ SD, g/L	111.62 $\pm$ 11.30	114.13 $\pm$ 12.01	111.37 $\pm$ 11.24		-0.929		0.354 <sup>d</sup>
Hct, mean $\pm$ SD, %	33.31 $\pm$ 2.99	33.20 $\pm$ 2.93	34.44 $\pm$ 3.44		-2.048		0.042 <sup>d*</sup>
CRP, mean $\pm$ SD, mg/L	61.05 $\pm$ 46.88	62.80 $\pm$ 47.68	42.47 $\pm$ 33.03		2.218		0.028 <sup>d*</sup>
ESR, mean $\pm$ SD, mm/h	61.64 $\pm$ 23.92	57.25 $\pm$ 24.84	62.05 $\pm$ 23.87		0.766		0.444 <sup>d</sup>
PCT, median (IQR), ng/mL	0.34 (0.13, 0.90)	0.35 (0.14, 0.94)	0.22 (0.10, 0.52)	-1.972			0.049 <sup>a*</sup>
IL-6, median (IQR), pg/mL	16.70 (6.34, 53.30)	17.50 (2.00, 90.00)	16.60 (6.63, 52.45)	-0.291			0.771 <sup>a</sup>
TNF- $\alpha$ , mean $\pm$ SD, pg/mL	16.30 $\pm$ 11.66	17.12 $\pm$ 10.17	16.24 $\pm$ 11.82	-0.276			0.782 <sup>a</sup>
CK-MB, mean $\pm$ SD, $\mu/L$	16.92 $\pm$ 14.07	18.80 $\pm$ 15.20	16.70 $\pm$ 13.97	-0.386			0.699 <sup>a</sup>
ALT, median (IQR), IU/L	28.50 (16.00, 51.00)	24.00 (14.00, 43.00)	29.00 (16.00, 52.00)	-0.681			0.496 <sup>a</sup>
AST, median (IQR), IU/L	27.00 (22.00, 38.00)	32.00 (24.00, 40.00)	27.00 (21.00, 38.00)	-0.846			0.398 <sup>a</sup>
TBIL, median (IQR), $\mu\text{mol/L}$	6.65 (5.07, 9.23)	6.40 (5.70, 9.09)	6.70 (4.92, 9.32)	-0.169			0.866 <sup>a</sup>
DBIL, median (IQR), $\mu\text{mol/L}$	2.35 (1.60, 3.50)	2.85 (2.10, 3.10)	2.30 (1.60, 3.53)	-0.987			0.324 <sup>a</sup>
IBIL, mean $\pm$ SD, $\mu\text{mol/L}$	4.28 $\pm$ 2.10	4.57 $\pm$ 1.08	4.25 $\pm$ 2.16		-0.360		0.720 <sup>d</sup>
GGT, median (IQR), IU/L	19.00 (11.00, 59.50)	32.00 (15.50, 48.75)	18.00 (11.00, 61.00)	-0.621			0.535 <sup>a</sup>
LDH, mean $\pm$ SD, IU/L	487.58 $\pm$ 281.91	614 $\pm$ 478.83	475.38 $\pm$ 254.56	-0.699			0.485 <sup>a</sup>

**Table 2** continued

Variables	Total (n=203)	CAAs (n=17)	Without CAAs (n=186)	z value	t value	$\chi^2$ value	p value
Alb, mean $\pm$ SD, g/L	37.64 $\pm$ 4.53	38.14 $\pm$ 4.00	37.60 $\pm$ 4.59		-0.445		0.657 <sup>d</sup>
Sodium, mean $\pm$ SD, mmol/L	136.46 $\pm$ 3.46	135.78 $\pm$ 3.11	136.52 $\pm$ 3.49		0.771		0.442 <sup>d</sup>

Data are expressed as mean  $\pm$  SD, median (IQR), or number with percentage.

Alb albumin, ALT alanine aminotransferase, AST aspartate aminotransferase, CAA coronary artery abnormality, CK-MB creatine kinase MB, CRP C-reactive protein, DBIL direct bilirubin, ESR erythrocyte sedimentation rate, GGT  $\gamma$ -glutamyltransferase, Hb hemoglobin, Hct hematocrit, IBIL indirect bilirubin, IL-6 interleukin-6, IVIG intravenous immunoglobulin, LDH lactate dehydrogenase, MO% percentage of monocytes, N% percentage of neutrophils, L% percentage of lymphocytes, EO% percentage of eosinophils, PCT procalcitonin, PLT platelet count, TBIL total bilirubin, TNF tumor necrosis factor, WBC white blood cell counts, NLR neutrophil-to-lymphocyte ratio, PLR platelet-to-lymphocyte ratio.

\*Statistical significance.

<sup>a</sup>Mann–Whitney *U* test.

<sup>b</sup>Chi-squared test.

<sup>c</sup>Fisher's exact test.

<sup>d</sup>Two independent samples *t* test.

**Table 3.** Multivariable predictors of CAAs using the newly established scoring system.

	Logistic coefficient ( $\beta$ )	Standard error	Odds ratio (95% CI)	p value	Score point
Days of illness at initial treatment $\geq 7$	0.462	0.177	1.588 (1.123–2.245)	0.009	2
Redness and swelling of extremities	0.550	0.693	1.733 (1.446–6.736)	0.027	2
Hct $\leq 33\%$	0.063	0.125	1.165 (1.034–1.360)	0.014	1
MO% $\geq 8.89\%$	0.298	0.150	1.347 (1.004–1.808)	0.047	1
PCT $\geq 0.5$ ng/mL	0.079	0.119	1.208 (1.056–1.367)	0.050	1

CAA coronary artery abnormality, CI confidence interval, Hct hematocrit, MO% percentage of monocytes, PCT procalcitonin.

### Assessment of the newly established scoring system for CAAs

Further analyses for the logistic regression model are listed in Table 3. The odds ratio (95% CI) values for days of illness at initial treatment  $\geq 7$ , redness and swelling of extremities, Hct  $\leq 33\%$ , MO%  $\geq 8.89\%$ , and PCT  $\geq 0.5$  ng/mL were 1.588 (1.123–2.245), 1.733 (1.446–6.736), 1.165 (1.034–1.360), 1.347 (1.004–1.808) and 1.208 (1.056–1.367), respectively. Based on the logistic coefficient of significant predictors, two point was assigned for days of illness at initial treatment  $\geq 7$  and redness and swelling of extremities, and one point was assigned for Hct  $\leq 33\%$ , MO%  $\geq 8.89\%$ , and PCT  $\geq 0.5$  ng/mL. After identifying the occurrence of CAAs and overall numbers of patients for individual scores, a scoring system was established, dependent on the sum of points in each patient. For simplicity, 2 risk strata were identified: low risk, with scores of 0–4, encompassing 82.27% of patients; and high risk, with scores of  $\geq 5$ , 17.73% of patients. The AUC value of the newly established scoring system was 0.685 (95% CI 0.526–0.844), with a sensitivity of 41.18% and a specificity of 84.41%.

### Comparisons of the newly established scoring system with the others

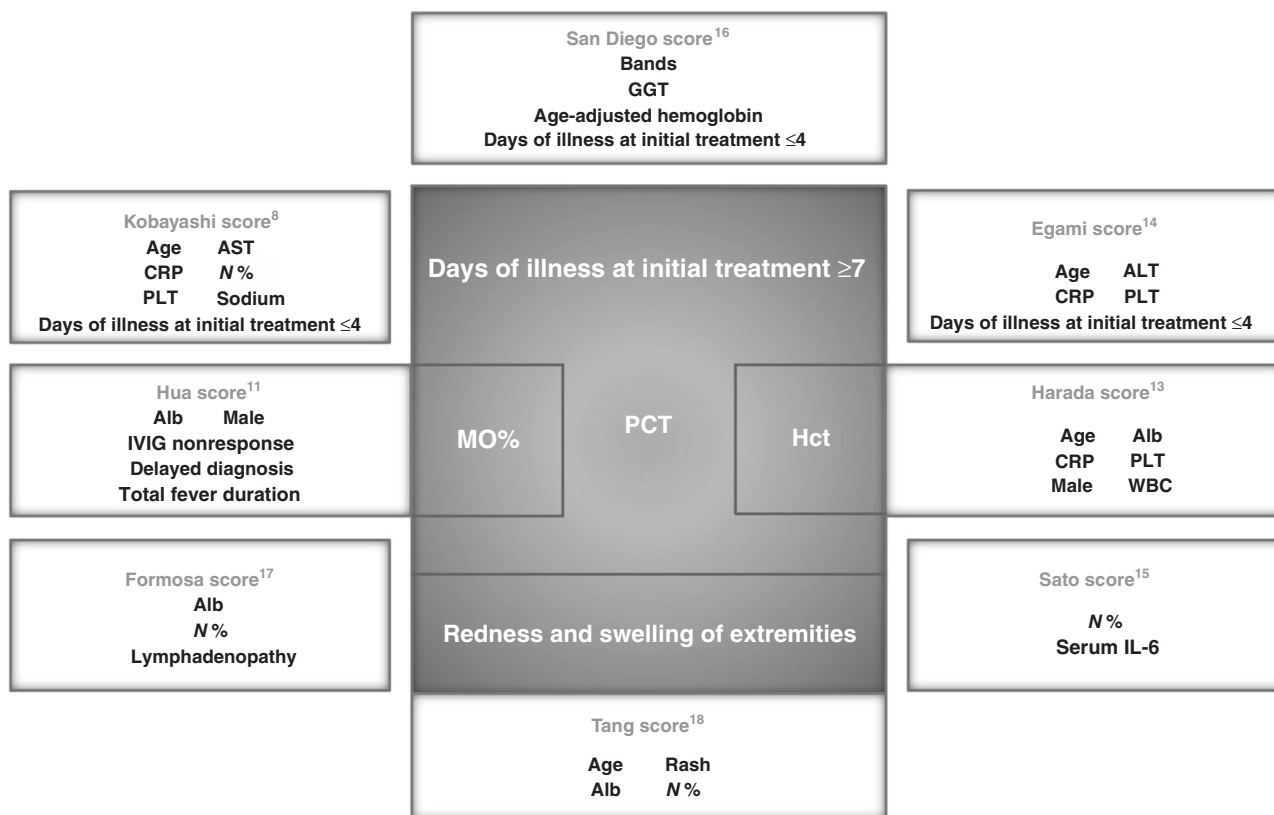
Comparisons of the newly established scoring system with the others are shown in Fig. 2. The newly established scoring system contained five independent predictors: days of illness at initial treatment, redness and swelling of extremities, Hct, MO%, and PCT. Among them, days of illness at initial treatment  $\geq 7$  and PCT were unique for the newly established scoring system; in addition, MO% was included in Hua score<sup>11</sup>; redness and swelling of extremities were included in Tang score<sup>18</sup>; and Hct was included in Harada score<sup>13</sup> simultaneously.

CAA risks of our patients were evaluated by the newly established scoring system and eight other scoring systems (Table 4 and Fig. 3). The newly established scoring system had a sensitivity of 41.18% and a specificity of 84.41%; similarly, based on our patients, the sensitivity of 8 published scoring systems ranged from 5.88 to 52.94% and the specificity ranged from 41.94

to 94.62% for predicting CAAs. The newly established scoring system had a positive predictive value of 19.44% and a negative predictive value of 94.01%, similar to 8 other scoring systems (positive predictive values: 4.76–18.75%, negative predictive values: 89.66–93.57%). The AUC value of newly established scoring system was 0.685, higher than Harada (0.452),<sup>13</sup> Egami (0.479),<sup>14</sup> Kobayashi (0.526),<sup>8</sup> Sato (0.476),<sup>15</sup> San Diego (0.522),<sup>16</sup> Formosa (0.473),<sup>17</sup> and Tang score (0.458),<sup>18</sup> whereas only lower than Hua score (0.722).<sup>11</sup>

### DISCUSSION

Cardiovascular manifestations are prominent during the acute KD episode and a major contributor to the long-term prognosis. CAAs are the most important cardiovascular complications. In the past two decades, several previous epidemiological surveys from mainland China indicated that the overall trend in CAA morbidity appeared to be on the rise from 15.9 to 24.1%.<sup>1,10,11</sup> Taking noncoronary cardiac abnormalities (NCAs) into account, data from multiple centers revealed that mitral regurgitation was the most common (27%), followed by pericardial effusion (15%), cardiac arrhythmia (11%), and myocarditis (9.9%).<sup>19,20</sup> In the present study, 8.37% of KD patients developed CAAs and 65.02% had NCAs in the acute stage; more specifically, valvular regurgitation was the most prevalent type of NCAs, followed by myocarditis, pericardial effusion, and cardiac arrhythmia. It should be pointed out that despite higher incidence of NCAs and KD in children aged  $>5$  years, the present study was totally irrelevant to multisystem inflammatory syndrome in children (MIS-C) triggered by coronavirus disease 2019 (COVID-19). First, the present study between July 2015 and March 2019 was before the date of the first official report of COVID-19 and none of patients were infected by COVID-19. Second, according to the latest American College of Rheumatology Clinical Guidance for MIS-C in children,<sup>21</sup> the incidence of MIS-C is lower in patients of East Asian descent than of African, Afro-Caribbean, and Hispanic descent. Third, patients



**Fig. 2 Comparisons of the newly established scoring system with the others.** Alb albumin, ALT alanine aminotransferase, AST aspartate aminotransferase, CAA coronary artery abnormality, CRP C-reactive protein, GGT  $\gamma$ -glutamyltransferase, Hct hematocrit, IL-6 interleukin-6, IVIG intravenous immunoglobulin, MO%, percentage of monocytes, N% percentage of neutrophils, PLT platelet count, WBC white blood cell count.

with MIS-C have more prominent gastrointestinal and neurologic symptoms and present more frequently in a state of shock. However, in our patients, gastrointestinal and neurologic system involvements were rare. Fourth, in the present study, we analyzed that the higher incidence of myocarditis may be attributed to larger variations in classification criteria, genetic background, and pediatricians' awareness.<sup>22</sup>

In the present study, a new scoring system for CAAs was established based on the data of Chinese population; it contained 5 variables and cut-off values of each variable are as follows: days of illness at initial treatment  $\geq 7$ , redness and swelling of extremities, Hct  $\leq 33\%$ , MO%  $\geq 8.89\%$ , and PCT  $\geq 0.5$  ng/mL. Compared with the other published scoring systems, days of illness at initial treatment  $\geq 7$  and PCT were unique predictors of CAAs. PCT is an acute-phase reactant and displays a quick response to bacterial infection. In a recent study from Chongqing, China, Si et al.<sup>23</sup> noted that PCT underwent a 12-fold increase in KD patients compared with the healthy controls. More persuasively, a multivariable logistic regression model established by Yoshikawa et al.<sup>24</sup> revealed that PCT had a sensitivity of 67% and a specificity of 55% at a cut-off value of  $>0.5$  ng/mL to predict CAAs and possessed the highest odds ratio compared with the other classical inflammatory mediators. However, in a retrospective study, Dominguez et al.<sup>25</sup> found that elevated PCT was not related to CAA development, providing controversial evidence for the role of PCT in predicting CAAs. Therefore, further studies of larger sample size and long-term follow-up will be necessary to consolidate our findings.

As we all know, days of illness at initial treatment  $\leq 10$  can dramatically decrease CAA risks in KD patients.<sup>12</sup> In the present study, more rigorously, days of illness at initial treatment  $< 7$  were identified as an optimal cut-off value for IVIG administration. In a retrospective study from Wenzhou, China, Qiu et al.<sup>26</sup> recruited

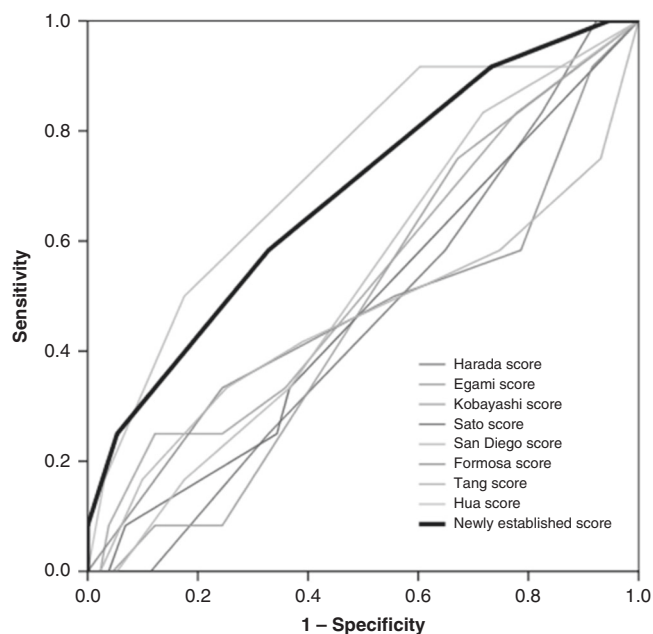
930 KD patients with a mean day at initial treatment of 6.99 and found that CAA risks increased 18% per 1 day delay of IVIG administration at 1 month after illness onset. In 2018, a national epidemiological survey from Japan also highly recommended initial IVIG administration within 7 days of illness for preventing CAAs.<sup>27</sup> However, in Egami,<sup>14</sup> Kobayashi,<sup>8</sup> and San Diego scores,<sup>16</sup> days of illness at initial treatment  $\leq 4$  were regarded as a strong predictor of CAAs. Because days of illness at initial treatment in all patients ranged from 5 to 11 days, it was unlikely to identify whether IVIG treatment before the fourth day of illness was an independent predictor in the present study. Therefore, based on the current evidence from our center and the others, we hypothesize that the time option of IVIG administration should be in parallel with peak time of systemic inflammation for the purpose of reducing CAA risks; in contrast, either earlier or later IVIG administration may increase CAA risks.

The sign of redness and swelling in extremities was the only major clinical feature included in our scoring system and also included in Tang score.<sup>18</sup> According to the 2017 AHA guidelines,<sup>12</sup> in the presence of  $\geq 4$  major clinical features, particularly when redness and swelling of the hands and feet are present, the diagnosis of KD may be made with only 4 days of fever. Both redness/swelling and periungual desquamation are the common extremity changes in KD; the former often occurs within 2 weeks of illness, while the latter usually begins thereafter. A multivariable analysis by Wang et al.<sup>28</sup> indicated that KD patients without periungual desquamation were more likely to develop CAAs. On the other hand, Michie et al.<sup>29</sup> followed up 259 KD patients for  $> 4$  years and found that recurrent desquamation was significantly less frequent in CAA patients. In spite of limited information, pediatricians should pay more attention to the potential contribution of extremity changes to CAA risks in KD patients.

**Table 4.** Comparisons of the newly established scoring system with the others.

Scoring systems	Category	CAAs (n = 17)	Without CAAs (n = 186)	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	AUC	95% CI
Harada score <sup>13</sup>	High risk	9	101	52.94	45.70	8.18	91.40	46.31	0.452	0.303–0.601
	Low risk	8	85							
Egami score <sup>14</sup>	High risk	1	20	5.88	89.25	4.76	91.21	82.27	0.479	0.332–0.625
	Low risk	16	166							
Kobayashi score <sup>8</sup>	High risk	3	24	17.65	87.10	11.11	92.05	81.28	0.526	0.353–0.700
	Low risk	14	162							
Sato score <sup>15</sup>	High risk	1	10	5.88	94.62	9.09	91.67	87.19	0.476	0.310–0.643
	Low risk	16	176							
San Diego score <sup>16</sup>	High risk	5	62	29.41	66.67	7.46	91.18	63.55	0.522	0.369–0.674
	Low risk	12	124							
Formosa score <sup>17</sup>	High risk	8	108	47.06	41.94	6.90	89.66	42.36	0.473	0.277–0.669
	Low risk	9	78							
Tang score <sup>18</sup>	High risk	5	68	29.41	63.44	6.85	90.77	60.59	0.458	0.252–0.664
	Low risk	12	118							
Hua score <sup>11</sup>	High risk	6	26	35.29	86.02	18.75	93.57	81.77	0.722	0.566–0.879
	Low risk	11	160							
≤6 months old (n = 16)	High risk	1	0	50	100	100	93.33	93.75	0.982	0.918–1.046
	Low risk	1	14							
Newly established score	High risk	7	29	41.18	84.41	19.44	94.01	80.79	0.685	0.526–0.844
	Low risk	10	157							

AUC area under the curve, CAA coronary artery abnormality, CI confidence interval, NPV negative predictive value, PPV positive predictive value.



**Fig. 3** ROC curves of multiple scoring systems for predicting CAAs. Thick solid line represents ROC curve of the newly established scoring system. Thin solid lines represent ROC curves of other published scoring systems.

Although the detailed pathogenesis of KD is still unclear, monocyte/macrophage activation, cytokine cascade, and vascular endothelial cell injury have been proven to be the three key steps of CAAs. Infiltration by monocytes/macrophages is ubiquitous in vascular pathology. Consistently, MO% was screened as a candidate predictor of CAAs in both the present and Hua score.<sup>11</sup> Furukawa et al.<sup>30</sup> analyzed the medical records of 27 KD patients and detected that MO% was subjected to a 0.81-fold increase in cases with CAAs compared with their counterparts. In 2018, a multivariable logistic model by Hua et al.<sup>11</sup> showed that KD patients with MO%  $\geq 5.9\%$  were 1.36 times likely to develop CAAs as compared with their counterparts.

Anemia occurs in 87.28% of KD patients and resolves with resolution of systemic inflammation.<sup>31</sup> Bone marrow infiltration by infective agents, alteration of iron metabolism, and reduction in erythropoiesis and erythrocyte survival are speculated to play a combined role in the pathogenic connection between KD and anemia. In the present study, a lower Hct was identified as a powerful predictor of CAAs and also included in Harada score.<sup>13</sup> In an American retrospective study, Ghelani et al.<sup>32</sup> reported that Hct significantly decreased in KD patients with CAAs ( $29.5 \pm 4.7\%$ ) compared to those with normal coronary artery ( $32.0 \pm 3.4\%$ ). In addition, the 15th Japanese epidemiological survey on KD indicated that patients with Hct  $\leq 32.5\%$  had a 0.45-fold increase in CAA risks as compared with their counterparts.<sup>33</sup> As we all know, KD predominantly affects children  $< 5$  years of age. Although several studies have shown an underlying correlation between anemia and CAAs, interpreting hematology analyses in children is challenging and hematology reference intervals for separate age groups exhibits significant variations due to the extensive changes in hematopoiesis that accompany physiological development. A further study after adjustment for age is warranted to more accurately elucidate the predictive role of anemia in CAAs.

To date, at least three different scoring systems for predicting CAAs have been published in Chinese population.<sup>11,17,18</sup> Apart from establishing a new scoring system, the present study undertook a systemic comparison of these Chinese scoring systems for the first time and revealed that the AUC value of

newly established scoring system was 0.685, very similar to Hua score.<sup>11</sup> Because of different identification criteria and analytical methodologies, there is still some heterogeneity among these Chinese scoring systems. First, the JMH criteria was used in Hua<sup>11</sup> and Tang<sup>18</sup> scores to identify CAAs, while Z score was used in Formosa score.<sup>17</sup> Second, the occurrence of incomplete KD was 58.8% in the study of Hua et al.,<sup>11</sup> significantly higher than the findings of Tang et al. (29.7%)<sup>18</sup> and ours (31.5%). We speculate that the other febrile diseases known to mimic KD may not be excluded completely in the study of Hua et al.<sup>11</sup> Third, variables with  $p < 0.1$  in the univariate analysis were entered into the multivariable model in Formosa score,<sup>17</sup> while variables with  $p < 0.05$  were selected in the present study.

## CONCLUSIONS

In the present study, 8.37% of KD patients were identified as having CAAs. Using a multivariable logistic regression analysis, a new scoring system for CAAs was established, including five independent predictors (days of illness at initial treatment, redness and swelling of extremities, Hct, MO%, and PCT). The AUC value of the newly established scoring system was 0.685, with a sensitivity of 41.18% and a specificity of 84.41%. Compared with 8 other published scoring systems, days of illness at initial treatment  $\geq 7$  and PCT were unique predictors of CAAs in our study. The AUC value of newly established scoring system was higher than Harada, Egami, Kobayashi, Sato, San Diego, Formosa, and Tang scores, whereas only lower than Hua score. Moreover, the present study also undertook a systemic comparison of three Chinese scoring systems for the first time and revealed that the AUC value of newly established scoring system was very similar to Hua score. Taking into account different identification criteria and analytical methodologies, there is still some heterogeneity among these Chinese scoring systems. Further multicenter collaborative studies should be advocated to consolidate our preliminary findings prior to widespread application in clinical practice.

## REFERENCES

- Chen, J. J. et al. Epidemiologic features of Kawasaki disease in Shanghai from 2008 through 2012. *Pediatr. Infect. Dis. J.* **35**, 7–12 (2016).
- Holve, T. J. et al. Long-term cardiovascular outcomes in survivors of Kawasaki disease. *Pediatrics* **133**, e305–e311 (2014).
- Tang, Y., Gao, X., Shen, J., Sun, L. & Yan, W. Epidemiological and clinical characteristics of Kawasaki disease and factors associated with coronary artery abnormalities in East China: nine years experience. *J. Trop. Pediatr.* **62**, 86–93 (2016).
- Chen, J., Liu, Y., Liu, W. & Wu, Z. A meta-analysis of the biomarkers associated with coronary artery lesions secondary to Kawasaki disease in Chinese children. *J. Huazhong. Univ. Sci. Technol. Med. Sci.* **31**, 705 (2011).
- Jakob, A. et al. Failure to predict high-risk Kawasaki disease patients in a population-based study cohort in Germany. *Pediatr. Infect. Dis. J.* **37**, 850–855 (2018).
- Davies, S. et al. Predicting IMG resistance in UK Kawasaki disease. *Arch. Dis. Child.* **100**, 366–368 (2015).
- Berdej-Szczot, E. et al. Risk factors of immunoglobulin resistance and coronary complications in children with Kawasaki disease. *Kardiol. Pol.* **75**, 261–266 (2017).
- Kobayashi, T. et al. Prediction of intravenous immunoglobulin unresponsiveness in patients with Kawasaki disease. *Circulation* **113**, 2606–2612 (2006).
- Fabi, M. et al. Inability of Asian risk scoring systems to predict intravenous immunoglobulin resistance and coronary lesions in Kawasaki disease in an Italian cohort. *Eur. J. Pediatr.* **178**, 315–322 (2019).
- Du, Z. D. et al. Epidemiologic study on Kawasaki disease in Beijing from 2000 through 2004. *Pediatr. Infect. Dis. J.* **26**, 449–451 (2007).
- Hua, W. et al. A new scoring system to predict Kawasaki disease with coronary artery lesions. *Clin. Rheumatol.* **38**, 1099–1107 (2019).
- McCrindle, B. W. et al. Diagnosis, treatment, and long-term management of Kawasaki disease: a scientific statement for health professionals from the American Heart Association. *Circulation* **135**, 927–999 (2017).
- Harada, K. Intravenous gamma-globulin treatment in Kawasaki disease. *Acta Paediatr. Jpn.* **33**, 805–810 (1991).



14. Egami, K. et al. Prediction of resistance to intravenous immunoglobulin treatment in patients with Kawasaki disease. *J. Pediatr.* **149**, 237–240 (2006).
15. Sato, S., Kawashima, H., Kashiwagi, Y. & Hoshika, A. Inflammatory cytokines as predictors of resistance to intravenous immunoglobulin therapy in Kawasaki disease patients. *Int. J. Rheum. Dis.* **16**, 168–172 (2013).
16. Tremoulet, A. H. et al. Resistance to intravenous immunoglobulin in children with Kawasaki disease. *J. Pediatr.* **153**, 117–121 (2008).
17. Lin, M. T. et al. Risk factors and derived formosa score for intravenous immunoglobulin unresponsiveness in Taiwanese children with Kawasaki disease. *J. Formos. Med. Assoc.* **115**, 350–355 (2016).
18. Tang, Y. et al. Prediction of intravenous immunoglobulin resistance in Kawasaki disease in an East China population. *Clin. Rheumatol.* **35**, 2771–2776 (2016).
19. Garrido-García, L. M., Peña-Juárez, R. A. & Yamazaki-Nakashimada, M. A. Cardiac manifestations in the acute phase of Kawasaki disease in a third level children's hospital in Mexico City. *Arch. Cardiol.* **88**, 441–446 (2018).
20. Printz, B. F. et al. Noncoronary cardiac abnormalities are associated with coronary artery dilation and with laboratory inflammatory markers in acute Kawasaki disease. *J. Am. Coll. Cardiol.* **57**, 86–92 (2011).
21. Henderson, L. A. et al. American College of Rheumatology Clinical Guidance for multisystem inflammatory syndrome in children associated with SARS-CoV-2 and hyperinflammation in pediatric COVID-19: Version 2. *Arthritis Rheumatol.* **73**, e13–e29 (2021).
22. Liu, F. F. et al. Clinical observation of noncoronary cardiac abnormalities in Chinese children with Kawasaki disease. *Eur. J. Clin. Invest.* **50**, e13210 (2020).
23. Si, F. et al. Relationship between IL-27 and coronary arterial lesions in children with Kawasaki disease. *Clin. Exp. Med.* **17**, 451–457 (2017).
24. Yoshikawa, H. et al. Serum procalcitonin value is useful for predicting severity of Kawasaki disease. *Pediatr. Infect. Dis. J.* **31**, 523–525 (2012).
25. Dominguez, S. R. et al. Procalcitonin (PCT) and Kawasaki disease: does PCT correlate with IVIG-resistant disease, admission to the intensive care unit, or development of coronary artery lesions? *J. Pediatr. Infect. Dis. Soc.* **5**, 297–302 (2016).
26. Qiu, H. et al. Delayed intravenous immunoglobulin treatment increased the risk of coronary artery lesions in children with Kawasaki disease at different status. *Postgrad. Med.* **130**, 442–447 (2018).
27. Kuwabara, M., Yashiro, M., Ae, R., Yanagawa, H. & Nakamura, Y. The effects of early intravenous immunoglobulin therapy for Kawasaki disease: the 22nd nationwide survey in Japan. *Int. J. Cardiol.* **269**, 334–338 (2018).
28. Wang, S., Best, B. M. & Burns, J. C. Periungual desquamation in patients with Kawasaki disease. *Pediatr. Infect. Dis. J.* **28**, 538–539 (2009).
29. Michie, C., Kinsler, V., Tulloh, R. & Davidson, S. Recurrent skin peeling following Kawasaki disease. *Arch. Dis. Child.* **83**, 353–355 (2000).
30. Furukawa, S. et al. Peripheral blood monocyte/macrophages and serum tumor necrosis factor in Kawasaki disease. *Clin. Immunol. Immunopathol.* **48**, 247–251 (1988).
31. Kim, S. & Eun, L. Y. Iron deficiency anemia as a predictor of coronary artery abnormalities in Kawasaki disease. *Korean J. Pediatr.* **62**, 301–306 (2019).
32. Ghelani, S. J., Kwatra, N. S. & Spurney, C. F. Can coronary artery involvement in Kawasaki disease be predicted? *Diagnostics* **3**, 232–243 (2013).
33. Zhang, T., Yanagawa, H., Oki, I. & Nakamura, Y. Factors relating to the cardiac sequelae of Kawasaki disease one month after initial onset. *Acta Paediatr.* **91**, 517–520 (2002).

#### AUTHOR CONTRIBUTIONS

H.H.L. and W.X.C. conceptualized and designed the study, collected data, conducted the analyses, drafted the initial manuscript, and reviewed and revised the manuscript. M.M.N., Q.J., Z.Q., G.Z.F., R.X.L., and G.M. collected data, carried out the initial analyses, and reviewed and revised the manuscript. Y.F.W., H.H.L., and D.D.Z. helped conceptualize and design the project, assisted in the analysis of the data, and reviewed and revised the manuscript. P.H. conceptualized and designed the study, coordinated and supervised data collection, and critically reviewed the manuscript for important intellectual content. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

#### COMPETING INTERESTS

The authors declare no competing interests.

#### ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The informed consent was obtained from the parents of each patient.

#### ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Peng Hu.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.