

# Predicting referral need for febrile children in low-resource community settings in South and Southeast Asia

Received: 12 October 2025

Accepted: 11 March 2026

Published online: 29 April 2026

 Check for updates

A list of authors and their affiliations appears at the end of the paper

In resource-constrained community settings, identifying which febrile children require referral remains a major unmet need. Current World Health Organization (WHO) danger signs have limited accuracy, resulting in missed severe illness and unnecessary referrals. Here we developed and validated clinical prediction models to support referral decisions using data from 3,405 children aged 1–59 months presenting with community-acquired acute febrile illnesses to seven hospitals across Bangladesh, Cambodia, Indonesia, Laos and Vietnam. Cambodian data were held out for external validation. The model using simple clinical parameters (sensitivity 74.7% (95% confidence interval (CI): 59.4–88.1); specificity 99.1% (95% CI: 97.7–99.7)) outperformed WHO criteria (sensitivity 55.5% (95% CI: 39.4–72.7); specificity 82.6% (95% CI: 77.1–87.6)) for identification of children at risk of severe disease (death or organ support within 2 days). Including either pulse oximetry or the host biomarker soluble TREM1 (sTREM1) increased sensitivity to 88.9% (95% CI: 76.7–97.8; pulse oximetry) and 89.2% (95% CI: 76.9–97.5; sTREM1), respectively. The pulse oximetry-based model achieved these gains with a threefold reduction in referral rates. These approaches appear cost-effective (pulse oximetry incremental cost effectiveness ratio (ICER) = \$26.28; sTREM1 ICER = \$196.46) and could improve triage for febrile illness in low-resource settings by enabling more accurate referral decisions. They warrant evaluation in community-based trials.

Infectious diseases account for the majority of the 2.5 million deaths that occur each year among children aged 1–59 months<sup>1</sup>. Many deaths happen in community settings, because a child does not attend health services, a child presents too late, there are financial or logistical barriers to referral or there is failure to recognize impending critical illness<sup>2–4</sup>. Early identification of children at risk of life-threatening infection remains difficult<sup>5,6</sup>. Most Early Warning Scores (EWSs) rely on combinations of abnormal vital signs (for example, fever, tachycardia and tachypnea), but these are dynamic, prone to confounding and poorly sensitive and specific for identifying early stages of sepsis<sup>7,8</sup>.

In resource-constrained community settings, current WHO Integrated Management of Childhood Illnesses (IMCI) guidelines do not

advocate the use of vital signs to assess illness severity<sup>9,10</sup>. Instead, hospital referral is prompted by clinical danger signs, such as convulsions, intractable vomiting, lethargy or prostration. The accuracy of these indicators is suboptimal, and they are prone to considerable interobserver variability<sup>7,11–13</sup>. Neither vital signs nor danger signs reliably stratify risk in common childhood infections, underscoring the need for better prognostic tools<sup>14</sup>. Such tools would be particularly valuable in remote and conflict-affected settings, where capacity for safety-netting is typically minimal while the opportunity cost of referral is high.

One option, spotlighted by the COVID-19 pandemic, is expansion of pulse oximetry to primary care settings. Championed as the

✉ e-mail: [arjun@tropmedres.ac](mailto:arjun@tropmedres.ac)

'fifth vital sign'<sup>15</sup>, hypoxemia predicts poor outcomes in childhood illnesses, including pneumonia, meningitis, malaria and malnutrition, and identifies large numbers of children requiring hospital referral who are missed by IMCI guidelines alone<sup>16–20</sup>. Despite this promise, pulse oximetry remains underutilized in community settings. Barriers include availability, cost and usability of appropriately sized probes, accuracy across skin tones and staff capacity and workload; oximetry also takes time, especially in young children where measurement can be challenged by movement or crying<sup>21–23</sup>.

Another strategy is integration of clinical assessment with host biomarker testing<sup>5,24</sup>. Circulating markers of immune and endothelial activation predict disease severity agnostic to pathogen etiology and, in certain contexts, have outperformed pulse oximetry<sup>25–29</sup>. We recently reported promising prognostic performance of individual host biomarkers at the community level<sup>30</sup>, and an operational evaluation among village malaria workers in Cambodia demonstrated the feasibility of this approach<sup>31</sup>. If effective, such strategies could capitalize on point-of-care testing capacity developed within community healthcare worker networks over preceding decades by control programs for diseases such as malaria and HIV<sup>32</sup>.

Febrile illness refers to the presence of fever (typically  $\geq 38^\circ\text{C}$ )<sup>9,10</sup>, a common clinical feature of many childhood infections. Although most febrile episodes are self-limiting and caused by viral, bacterial or parasitic pathogens, individual clinical features rarely indicate which children will progress to severe disease. Building upon our previous work and using the same multicountry dataset<sup>30</sup>, our objective in this paper was to address a complementary question: how prognostic information can be integrated into risk prediction models to guide referral decisions for febrile children in resource-constrained community contexts and decentralized models of care. Whereas our previous analysis focused on the predictive performance of individual host biomarkers for disease progression, the present study develops and validates multivariable clinical prediction models comprising simple clinical parameters (for example, vital signs and danger signs) and evaluates the added utility of including pulse oximetry and host biomarker testing for referral decision-making.

## Results

### Study population

A total of 11,962 children were screened between 5 March 2020 and 4 November 2022. Of these, 3,998 were eligible (3,998/11,962, 33.4%) and 3,423 were recruited (575/3,998, 14.4% refusal rate). There were two withdrawals and 16 children with incomplete follow-up data, leaving 3,405 participants for further analyses (Extended Data Fig. 1).

Median age was 16.8 months (interquartile range (IQR), 8.7–31.0), and 59.6% (2,029/3,405) of the cohort were male. Malnutrition was prevalent: 17.2% (585/3,393) of children were wasted (weight-for-height  $z$ -score  $< -2$ ), and 19.5% (664/3,401) were stunted (height-for-age  $z$ -score  $< -2$ ). Of the 585 children with Global Acute Malnutrition, 248 (42.4%) had Severe Acute Malnutrition (weight-for-height  $z$ -score  $< -3$ ). Median symptom duration prior to presentation was 3 days (IQR, 2–4). Acute respiratory infections were the most common reason for presentation, followed by diarrheal syndromes and undifferentiated febrile illnesses. A quarter of participants had a microbiological cause for their infection identified (898/3,405, 26.4%; Supplementary Table 1). Most children lived close to the hospital (2,777/3,405, 81.6% within 1 hour). In total, 1,342 participants (1,342/3,405, 39.4%) had received care in the community at an earlier point in their illness; 193 (193/3,405, 5.7%) had received parenteral treatment, and none had been admitted.

Baseline characteristics were largely balanced across derivation and validation cohorts (Table 1), although the validation cohort was younger (median age, 13.1 months versus 18.7 months), and a higher proportion had received parenteral treatment in the community at an earlier point in their illness (109/824, 13.2% versus 84/2,581, 3.3%). Baseline biomarker concentrations were similar across the cohorts.

There was a greater proportion of confirmed viral infections in the derivation cohort (709/2,581, 27.4% versus 155/824, 18.7%). Bacteremia rates were similar (12/781, 1.5% versus 7/411, 1.7%), although blood cultures were collected more frequently from inpatients in the validation cohort (411/644, 63.8% versus 781/2,080, 37.5%).

### Progression to severe febrile illness

Overall, 133 children met the primary outcome (defined as death or organ support within 2 days of enrollment; 133/3,405, 3.9%): 111 survivors who required organ support (non-invasive ventilation,  $n = 59$ ; mechanical ventilation,  $n = 43$ ; inotropic therapy,  $n = 19$ ) and 22 deaths (derivation cohort = 97/2,581, 3.8% and validation cohort = 36/824, 4.4%). Weighted outcome prevalence was 0.34% (95% CI: 0.28–0.41), which was similar across derivation (0.36%, 95% CI: 0.29–0.45) and validation (0.30%, 95% CI: 0.20–0.42) cohorts.

Among the candidate clinical predictors (Supplementary Table 2), association with the outcome was similar across derivation and validation cohorts (Table 1), although a weaker association was observed for altered mental state, prolonged capillary refill time and presence of WHO danger signs in the validation cohort. Most candidate biomarker predictors were associated with the outcome across derivation and validation cohorts, apart from C-reactive protein (CRP), IL-1ra, IL-6, IL-10, glucose and hemoglobin, which were not associated with the outcome in the validation cohort.

### Clinical prediction models

When each of the 17 biomarkers was included in turn alongside the candidate clinical predictors, 11 were retained after backward stepwise selection in the derivation dataset. Similar results were obtained with and without the inclusion of peripheral oxygen saturation ( $\text{SpO}_2$ ) as a candidate predictor (Supplementary Tables 3 and 4). In a sensitivity analysis excluding the northern Vietnam site (Methods and Supplementary Table 5), a biomarker was retained in seven of the models. Across all analyses, the clinical biomarker models containing sTREM1 consistently demonstrated the highest discrimination. Thus, four models were taken forward for validation: the clinical model (clinical parameters only), the pulse oximetry model (clinical parameters plus  $\text{SpO}_2$ ), the sTREM1 model (clinical parameters plus sTREM1) and the combined model (clinical parameters plus  $\text{SpO}_2$  and sTREM1).

Heart rate, respiratory rate, prostration and intractable vomiting were consistently selected across all models ( $n = 4$ ); altered mental state was selected in two models; and prolonged capillary refill time and convulsions were selected in one model each (Table 2). Model equations and odds ratios are reported in the appendix (Supplementary Table 6).

Discrimination and calibration of the models in the validation cohort are presented (Table 2). Discrimination ranged from a weighted area under the receiver operating characteristic curve (AUC) of 0.95–0.98. Calibration was best at lower predicted probabilities, with some underestimation of risk in individuals with the highest predicted probabilities of severe disease.

### Clinical utility

Clinical utility of the models for triage of febrile children was assessed within a hypothetical traffic light framework whereby any child with a predicted probability of developing severe disease lower than 0.5% (the rule-out threshold) is discharged (green); those with a predicted probability higher than 2% (the rule-in threshold) are referred for higher-level care (red); and those with predicted probabilities between 0.5% and 2% are monitored (amber)—for example, at the rural clinic or via telephone or outreach follow-up (Fig. 1).

Within this framework, using the clinical model, one in 10 children recommended for referral would have developed severe disease. However, the clinical model would have failed to identify 25% of children who progressed to severe disease, in whom discharge would have been recommended (sensitivity 74.7% (95% CI: 59.4–88.1)). Compared to the

**Table 1 | Baseline characteristics of the derivation and validation cohorts, stratified by whether a child progressed to develop severe febrile illness**

Characteristic	Derivation cohort				Validation cohort			
	Overall n=2,581	Non-severe n=2,484	Severe n=97	Pvalue	Overall n=824	Non-severe n=788	Severe n=36	Pvalue
Demographics and background								
Age (months)	18.7 (9.3, 32.9)	18.9 (9.5, 33.1)	6.9 (3.2, 25.8)	<0.0001	13.1 (7.5, 24.4)	13.5 (8.1, 24.9)	3.4 (1.8, 6.2)	<0.0001
Male sex	1,577 (61%)	1,513 (61%)	64 (66%)	0.32	452 (55%)	428 (54%)	24 (67%)	0.15
Known comorbidity	65 (2.5%)	63 (2.5%)	2 (2.1%)	1.00	37 (4.5%)	37 (4.7%)	0 (0%)	0.40
Recent hospitalization <sup>a†</sup>	277 (11%)	265 (11%)	12 (13%)	0.58	152 (19%)	147 (19%)	5 (14%)	0.46
Anthropometrics								
Wasted (WHZ < -2) <sup>‡</sup>	474 (18%)	445 (18%)	29 (30%)	0.0030	111 (14%)	106 (14%)	5 (14%)	0.80
Stunted (HAZ < -2) <sup>‡</sup>	484 (19%)	462 (19%)	22 (23%)	0.31	180 (22%)	167 (21%)	13 (36%)	0.035
MUAC-for-age z-score <sup>b*</sup>	-0.2 (-1.1, 0.7)	-0.2 (-1.1, 0.7)	-0.5 (-1.4, 0.7)	0.12	-0.5 (-1.2, 0.2)	-0.5 (-1.1, 0.2)	-0.8 (-1.7, 0.3)	0.40
Illness history								
Duration of illness (days)	3.0 (2.0, 4.0)	3.0 (2.0, 4.0)	3.0 (2.0, 5.0)	0.020	3.0 (2.0, 4.0)	3.0 (2.0, 4.0)	3.5 (3.0, 5.0)	0.051
Care prior to presentation <sup>c</sup>	864 (33%)	829 (33%)	35 (36%)	0.58	478 (58%)	457 (58%)	21 (58%)	0.97
Travel time ≤1 hour <sup>d</sup>	2,227 (86%)	2,150 (87%)	77 (79%)	0.044	550 (67%)	532 (68%)	18 (50%)	0.029
Presenting syndrome <sup>e</sup>								
URTI	795 (31%)	773 (31%)	22 (23%)	0.077	326 (40%)	309 (39%)	17 (47%)	0.34
LRTI	930 (36%)	877 (35%)	53 (55%)	0.0001	417 (51%)	384 (49%)	33 (92%)	<0.0001
Diarrheal	397 (15%)	386 (16%)	11 (11%)	0.26	249 (30%)	245 (31%)	4 (11%)	0.011
No focus	372 (14%)	366 (15%)	6 (6.2%)	0.019	155 (19%)	153 (19%)	2 (5.6%)	0.037
Neurological	372 (14%)	359 (14%)	13 (13%)	0.77	58 (7.0%)	57 (7.2%)	1 (2.8%)	0.51
WHO danger signs								
Any danger sign present <sup>†</sup>	1,246 (48%)	1,171 (47%)	75 (78%)	<0.0001	361 (44%)	341 (43%)	20 (56%)	0.15
Prostration	199 (7.7%)	155 (6.2%)	44 (45%)	<0.0001	41 (5.0%)	37 (4.7%)	4 (11%)	0.098
Intractable vomiting <sup>†</sup>	604 (23%)	564 (23%)	40 (41%)	<0.0001	83 (10%)	81 (10%)	2 (5.6%)	0.57
Convulsions <sup>†</sup>	373 (14%)	360 (15%)	13 (14%)	0.79	60 (7.3%)	59 (7.5%)	1 (2.8%)	0.51
Lethargy <sup>†</sup>	520 (20%)	478 (19%)	42 (44%)	<0.0001	305 (37%)	286 (36%)	19 (53%)	0.045
Vital signs								
Heart rate (bpm) <sup>†</sup>								
1-12 months	154.0 (140.0, 171.0)	153.0 (140.0, 170.0)	170.0 (152.0, 184.0)	<0.0001	159.0 (143.0, 176.0)	157.0 (142.0, 173.0)	177.0 (161.0, 186.0)	<0.0001
12-60 months	140.0 (126.0, 157.0)	140.0 (126.0, 156.0)	159.0 (140.0, 179.5)	<0.0001	140.0 (128.0, 162.0)	140.0 (128.0, 162.0)	167.0 (146.0, 172.0)	0.17
Respiratory rate (bpm) <sup>†</sup>								
1-12 months	45.0 (38.0, 55.0)	44.0 (38.0, 53.0)	60.0 (50.0, 68.0)	<0.0001	44.0 (37.0, 54.0)	44.0 (36.0, 52.0)	58.0 (46.0, 64.0)	<0.0001
12-60 months	35.0 (30.0, 41.0)	35.0 (30.0, 41.0)	39.0 (31.5, 55.0)	0.016	38.0 (32.0, 45.0)	38.0 (32.0, 45.0)	54.0 (50.0, 67.0)	0.032
SpO <sub>2</sub> in room air (%) <sup>†</sup>	98.0 (97.0, 99.0)	98.0 (97.0, 99.0)	97.0 (95.0, 98.0)	0.0004	99.0 (98.0, 100.0)	99.0 (98.0, 100.0)	96.0 (94.5, 97.5)	<0.0001
Axillary temperature (°C) <sup>†</sup>	37.7 (37.0, 38.4)	37.7 (37.0, 38.4)	38.0 (37.2, 38.6)	0.098	37.3 (36.7, 38.1)	37.3 (36.7, 38.1)	37.0 (36.5, 37.5)	0.019
CRT > 2 seconds	180 (7.0%)	153 (6.2%)	27 (28%)	<0.0001	14 (1.7%)	12 (1.5%)	2 (5.6%)	0.12
Not alert <sup>†</sup>	58 (2.2%)	41 (1.7%)	17 (18%)	<0.0001	28 (3.4%)	26 (3.3%)	2 (5.6%)	0.35
Endothelial activation markers								
ANG-1 (pg ml <sup>-1</sup> ) <sup>†</sup>	6,396.5 (3,387.5, 11,645.0)	6,411.0 (3,387.0, 11,607.0)	6,049.0 (3,426.0, 13,382.0)	1.00	5,686.5 (3,454.0, 9,139.0)	5,713.5 (3,454.0, 9,240.0)	4,621.0 (3,440.5, 7,421.5)	0.24
ANG-2 (pg ml <sup>-1</sup> ) <sup>†</sup>	1,330.0 (998.5, 1,844.0)	1,320.0 (993.0, 1,802.0)	2,343.0 (1,317.0, 4,608.0)	<0.0001	1,455.5 (1,090.0, 2,051.0)	1,427.0 (1,081.0, 1,965.0)	3,008.0 (2,119.0, 4,254.5)	<0.0001
sFLT-1 (pg ml <sup>-1</sup> ) <sup>†</sup>	183.0 (149.0, 231.0)	181.0 (148.0, 227.0)	258.0 (196.0, 369.0)	<0.0001	196.0 (156.0, 245.0)	194.5 (154.0, 242.0)	260.5 (227.0, 415.5)	<0.0001
Immune activation markers								
CHI3L1 (ng ml <sup>-1</sup> ) <sup>†</sup>	23.8 (15.7, 37.7)	23.5 (15.5, 36.9)	37.6 (22.8, 62.6)	<0.0001	28.9 (20.2, 43.1)	28.5 (20.1, 42.5)	38.2 (26.9, 72.8)	0.0037
CRP (mg l <sup>-1</sup> ) <sup>†</sup>	13.2 (3.2, 49.3)	12.8 (3.2, 47.8)	31.4 (6.8, 150.5)	0.0001	17.3 (4.8, 55.3)	17.5 (4.8, 55.3)	11.8 (2.1, 55.5)	0.31
IL-1ra (pg ml <sup>-1</sup> ) <sup>†</sup>	2,036.0 (980.0, 4,875.0)	1,983.0 (967.0, 4,764.0)	4,117.0 (1,756.0, 14,784.0)	<0.0001	1,755.0 (789.0, 4,429.0)	1,781.0 (806.0, 4,440.0)	1,505.0 (728.5, 3,509.5)	0.32

**Table 1 (continued) | Baseline characteristics of the derivation and validation cohorts, stratified by whether a child progressed to develop severe febrile illness**

Characteristic	Derivation cohort				Validation cohort			
	Overall n=2,581	Non-severe n=2,484	Severe n=97	Pvalue	Overall n=824	Non-severe n=788	Severe n=36	Pvalue
IL-6 (pg ml <sup>-1</sup> ) <sup>a</sup>	18.4 (7.7, 46.3)	17.6 (7.6, 43.5)	73.7 (20.9, 320.0)	<0.0001	16.9 (7.2, 46.9)	16.8 (7.2, 46.9)	17.8 (7.7, 48.7)	0.80
IL-8 (pg ml <sup>-1</sup> ) <sup>a</sup>	15.4 (9.2, 27.5)	15.1 (9.1, 26.8)	35.7 (13.1, 129.0)	<0.0001	15.6 (9.3, 27.6)	15.2 (9.1, 26.4)	23.7 (16.4, 38.8)	0.0035
IL-10 (pg ml <sup>-1</sup> ) <sup>a</sup>	19.3 (10.5, 43.6)	19.1 (10.3, 42.6)	34.7 (15.5, 83.3)	<0.0001	20.5 (10.8, 43.6)	20.1 (10.8, 43.6)	23.8 (12.0, 43.4)	0.62
IP-10 (pg ml <sup>-1</sup> ) <sup>a</sup>	852.5 (387.5, 1,882.0)	857.0 (394.0, 1,895.0)	745.0 (248.0, 1,381.0)	0.023	727.0 (368.0, 1,858.0)	754.5 (383.0, 1,929.0)	371.0 (218.0, 692.0)	0.0001
PCT (ng ml <sup>-1</sup> ) <sup>a</sup>	0.3 (0.2, 0.7)	0.3 (0.2, 0.7)	1.0 (0.4, 6.1)	<0.0001	0.4 (0.2, 1.1)	0.4 (0.2, 1.1)	0.7 (0.4, 1.6)	0.0072
sTNF-R1 (pg ml <sup>-1</sup> ) <sup>a</sup>	1,559.0 (1,244.0, 2,022.0)	1,546.0 (1,240.0, 1,992.0)	2,166.0 (1,549.0, 3,542.0)	<0.0001	1,643.5 (1,306.0, 2,153.0)	1,627.0 (1,297.0, 2,129.0)	1,947.0 (1,542.5, 2,905.0)	0.0012
sTREM1 (pg ml <sup>-1</sup> ) <sup>a</sup>	227.0 (165.0, 333.0)	224.0 (163.0, 322.0)	415.0 (267.0, 615.0)	<0.0001	238.0 (182.0, 331.0)	235.0 (181.0, 325.0)	359.0 (279.5, 453.0)	<0.0001
suPAR (ng ml <sup>-1</sup> ) <sup>a</sup>	4.1 (3.3, 5.0)	4.0 (3.3, 5.0)	5.1 (3.8, 7.9)	<0.0001	4.6 (3.7, 5.8)	4.5 (3.7, 5.7)	5.6 (4.6, 7.2)	0.0001
Other laboratory biomarkers								
Lactate (mmol l <sup>-1</sup> ) <sup>a</sup>	1.0 (0.7, 1.3)	1.0 (0.7, 1.3)	1.1 (0.7, 1.8)	0.0060	1.2 (0.9, 1.6)	1.1 (0.8, 1.6)	1.8 (1.3, 2.8)	<0.0001
Glucose (mmol l <sup>-1</sup> ) <sup>a</sup>	5.4 (4.8, 6.2)	5.4 (4.8, 6.1)	6.0 (5.1, 7.1)	<0.0001	5.5 (4.9, 6.4)	5.5 (4.9, 6.4)	5.9 (5.0, 7.1)	0.14
Hb (g dl <sup>-1</sup> ) <sup>a</sup>	11.5 (10.6, 12.4)	11.5 (10.6, 12.4)	10.9 (9.5, 12.2)	0.0019	10.8 (9.9, 11.6)	10.8 (9.9, 11.6)	10.5 (9.4, 11.5)	0.23
Recruitment strata								
Inpatient	2,080 (81%)	1,983 (80%)	97 (100%)	<0.0001	644 (78%)	608 (77%)	36 (100%)	0.0012
Outpatient	501 (19%)	501 (20%)	0		180 (22%)	180 (23%)	0	

CRP, C-reactive protein; HAZ, height-for-age z-score; Hb, hemoglobin; LRTI, lower respiratory tract infection; SpO<sub>2</sub>, oxygen saturation; URTI, upper respiratory tract infection; WHZ, weight-for-height z-score. Data are shown as median (Q1, Q3) or n (%) unless otherwise indicated. P values (all two-sided) were obtained by Wilcoxon rank-sum test (continuous variables), Pearson's  $\chi^2$  test or Fisher's exact test (categorical variables). <sup>a</sup>Missing data: recent hospitalization, n=11 (7 derivation cohort: 6 non-severe and 1 severe; 4 validation cohort: 4 non-severe); wasted, n=12 (8 derivation cohort: 8 non-severe; 4 validation cohort: 3 non-severe and 1 severe); stunted, n=4 (2 derivation cohort: 2 non-severe; 2 validation cohort: 2 non-severe); MUAC-for-age z-score, n=216 (158 derivation cohort: 134 non-severe and 24 severe; 58 validation cohort: 41 non-severe and 17 severe); WHO danger signs, n=7 (6 derivation cohort: 5 non-severe and 1 severe; 1 validation cohort: 1 non-severe); intractable vomiting, n=4 (3 derivation cohort: 3 non-severe; 1 validation cohort: 1 non-severe); convulsions, n=5 (4 derivation cohort: 3 non-severe and 1 severe; 1 validation cohort: 1 non-severe); lethargy, n=6 (6 derivation cohort: 5 non-severe and 1 severe); heart rate, n=1 (1 derivation cohort: 1 non-severe); respiratory rate, n=2 (1 derivation cohort: 1 non-severe; 1 validation cohort: 1 non-severe); oxygen saturation, n=205 (150 derivation cohort: 110 non-severe and 40 severe; 55 validation cohort: 35 non-severe and 20 severe); axillary temperature, n=1 (1 derivation cohort: 1 non-severe); ANG-1, ANG-2, sFLT-1, IL-6, IL-8, IL-10, IP-10, PCT, sTNF-R1 and sTREM1, n=91 (57 derivation cohort: 51 severe and 6 non-severe; 34 validation cohort: 34 non-severe); CHI3L1, n=93 (59 derivation cohort: 52 non-severe and 7 severe; 34 validation cohort: 34 non-severe); CRP, n=96 (62 derivation cohort: 56 non-severe and 6 severe; 34 validation cohort: 34 non-severe); IL-1ra, n=92 (58 derivation cohort: 52 non-severe and 6 severe; 34 validation cohort: 34 non-severe); suPAR, n=124 (90 derivation cohort: 82 non-severe and 8 severe; 34 derivation cohort: 34 non-severe); lactate, glucose, n=98 (98 validation cohort: 91 non-severe and 7 severe); and Hb, n=574 (565 derivation cohort: 552 non-severe and 13 severe; 9 validation cohort: 9 non-severe). <sup>b</sup>For 125/205, immediate routinely collected SpO<sub>2</sub> on room air measurements were available and utilised for the clinical prediction models. <sup>c</sup>Overnight admission in the last 6 months. <sup>d</sup>Calculated in children aged 3–60 months (R package: zscorer)<sup>57</sup>. <sup>e</sup>Receipt of any treatment for the current illness in the community prior to presentation at the study site (of these, none had been admitted and 193 had received parenteral treatment). <sup>f</sup>Travel time to the study site  $\leq$ 1 hour. <sup>g</sup>Presenting syndromes were not mutually exclusive. <sup>h</sup>Assessed using the Alert Voice Pain Unresponsive (AVPU) scale.

clinical model, the pulse oximetry model was better able to rule in children at risk of disease progression (positive likelihood ratio (PLR) 150.7 (95% CI: 72.9–278.2) versus 37.5 (95% CI: 12.2–121.1)), with one in three children who would have been recommended for referral developing severe disease. It also demonstrated superior rule-out performance (sensitivity 88.9% (95% CI: 76.7–97.8) and negative likelihood ratio (NLR) 0.11 (95% CI: 0.02–0.25)). Overall, the pulse oximetry model demonstrated superior rule-out properties at the rule-out threshold (difference in sensitivity = 13.7% (95% CI: 3.4–27.3)) and rule-in properties at the rule-in threshold (difference in specificity = 0.7% (95% CI 0.1–2.1)), compared to the clinical model (Extended Data Table 1).

Compared to the clinical model, the sTREM1 model improved rule-out performance to a similar extent as the pulse oximetry model (sensitivity 89.2% (95% CI: 76.9–97.5) and NLR 0.12 (95% CI: 0.03–0.25)) but did not improve rule-in performance (PLR 18.2 (95% CI: 7.2–78.9) versus 37.5 (95% CI: 12.2–121.1)). Differences in sensitivity and specificity between the clinical model and the sTREM1 model were 13.9% (95% CI: 3.4–26.7) and –1.1% (95% CI: –2.8 to –0.03), respectively

(Extended Data Table 1). The combined model (including both SpO<sub>2</sub> and sTREM1) did not offer advantage over the pulse oximetry model, in terms of neither discharge (rule-out) nor referral (rule-in) performance.

All four models outperformed WHO danger signs in terms of rule-out (sensitivity 55.5% (95% CI: 39.4–72.7 and NLR 0.54 (95% CI: 0.32–0.74)) and rule-in (specificity 82.6% (95% CI: 77.1–87.6)) and PLR 3.2 (95% CI: 2.1–4.9)) performance. Differences in sensitivity and specificity between the WHO danger signs and the clinical model were 19.0% (95% CI: 0.0–41.1) and 16.4% (95% CI: 11.7–21.7), respectively (Extended Data Table 1).

### Impact analysis

In a deterministic impact analysis applied to a hypothetical population of 10,000 children, using fixed point estimates of sensitivity, specificity and disease prevalence obtained from the validation cohort, we projected that, compared to using the clinical model, either the pulse oximetry model or the sTREM1 model would have prevented one additional child who would progress to life-threatening infection

**Table 2 | Selected variables, discrimination and calibration of the clinical prediction models in the validation cohort**

Model	Constituent variables	Discrimination		Calibration	
		AUC (95% CI)	Intercept (95% CI)	Slope (95% CI)	
Clinical	Heart rate (bpm) Respiratory rate (bpm) Prostration Intractable vomiting Altered mental state Prolonged capillary refill time	0.95 (0.92–0.97)	0.31 (–0.08 to 0.71)	1.46 (1.19–1.73)	
Pulse oximetry	Heart rate (bpm) Respiratory rate (bpm) Prostration Intractable vomiting Convulsions Oxygen saturation (%)	0.97 (0.96–0.99)	0.60 (0.21–0.99)	1.74 (1.41–2.06)	
sTREM1	Heart rate (bpm) Respiratory rate (bpm) Prostration Intractable vomiting Altered mental state sTREM1 (pg ml <sup>-1</sup> )	0.96 (0.93–0.98)	0.18 (–0.26 to 0.62)	1.33 (1.04–1.63)	
Combined	Heart rate (bpm) Respiratory rate (bpm) Prostration Intractable vomiting Oxygen saturation (%) sTREM1 (pg ml <sup>-1</sup> )	0.98 (0.96–0.99)	0.48 (0.05–0.91)	1.58 (1.21–1.96)	

Variables selected in the different clinical prediction models (clinical model: clinical parameters only; pulse oximetry model: clinical parameters plus SpO<sub>2</sub>; sTREM1 model: clinical parameters plus sTREM1; combined model: clinical parameters plus SpO<sub>2</sub> and sTREM1). Performance of each model in the validation cohort is summarized using its weighted discrimination and calibration. Coefficients and odds ratios for the constituent variables are reported in the appendix (Supplementary Table 6).

being discharged for every approximately 2,300 children tested (number needed to test (NNT)). For the pulse oximetry model, these gains would have been achieved in addition to a threefold reduction in the referral rate (0.3% versus 1.0%), fewer children being recommended for monitoring (5.0% versus 8.4%) and more recommended for discharge (94.7% versus 90.6%). For the sTREM1 model, the gains would have been partially offset by twice the number of referrals (2.1% versus 1.0%) compared to the clinical model.

Compared to WHO danger signs, the prediction models could prevent one additional child who would progress to life-threatening infection being discharged for every 1,000 (pulse oximetry, sTREM1 or combined models) or 1,750 (clinical model) children tested (NNT) while simultaneously substantially reducing referral rates from 17.0% to 0.3–2.1%. As WHO danger signs provide a binary classification (refer or not), any benefits gained through use of the prediction models would be partially offset by the costs of monitoring up to 8.4% of all presentations.

### Categorical outcome scale

All four models predicted increasing probability of progression to severe febrile illness at higher levels of the categorical outcome scale (Supplementary Table 7) in the validation cohort (Fig. 2), suggesting that they would be able to provide discrimination across the severity spectrum rather than only identify the most severely unwell patients.

Using the categorical outcome scale to assess triage correctness indicated that improvement in sensitivity of the other models compared to the clinical model was driven by increases in both referrals and monitoring of the most severe (category IV) patients (Fig. 3). The reduction in referrals achieved by the pulse oximetry model was a consequence of both more discharges and fewer cases being recommended for monitoring among the least unwell children (category I).

### Missed cases of severe febrile illness

Among the 133 participants who progressed to severe febrile illness, the same children would have been missed by all four models in both

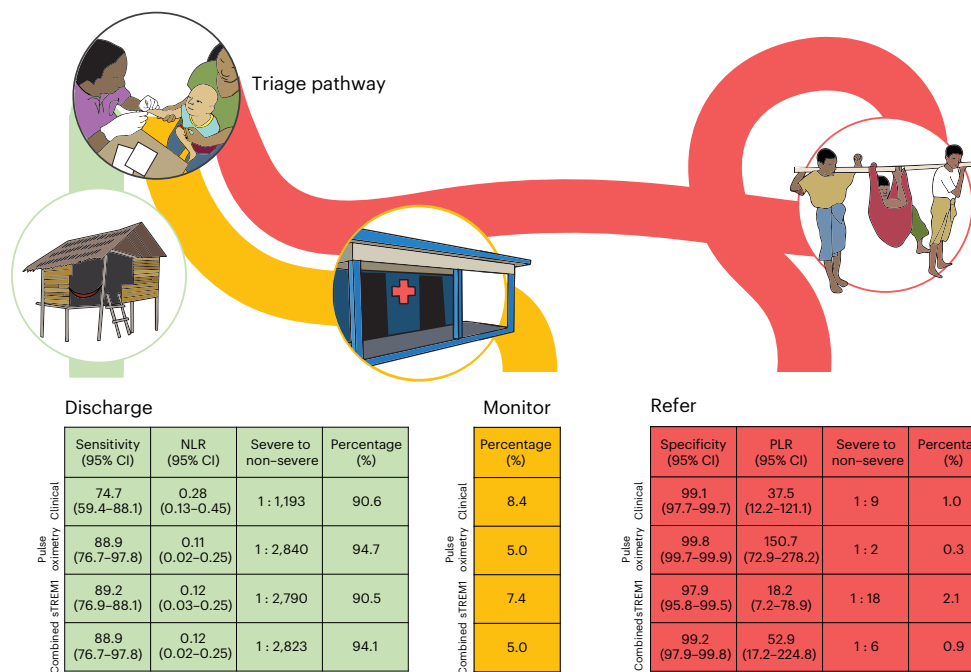
the derivation and validation cohorts (Extended Data Fig. 2 and Supplementary Table 8). Furthermore, the additional participants who would have been identified by the pulse oximetry, sTREM1 and combined models were broadly consistent across these three models.

Older children, males, those with a shorter illness prior to presentation and those without pneumonia may be more likely to be missed by the models (Supplementary Table 8). Baseline biomarker concentrations were typically less deranged in children whom the models would have failed to identify, apart from ANG-1, CRP, IL-6 and IP-10 concentrations, which were more abnormal in children who progressed to severe disease but who would have been missed by the models.

The models may be more likely to identify children who progress to severe illness quickly. Extending the definition of severe febrile illness to include all cases that developed in the 28 days after enrollment identified an additional 10 participants (two deaths and eight survivors who required organ support). Among the 42 children who progressed to develop severe febrile illness more than 24 hours after enrollment, 12 (28.6%) would have been missed by the clinical model, and the pulse oximetry and sTREM1 models would have missed nine (21.4%) and 10 (23.8%) participants, respectively.

### Cost-effectiveness

All four models were predicted to dominate (that is, were cost-saving and more effective than) WHO danger signs. Compared to the clinical model, the pulse oximetry, sTREM1 and combined models were all predicted to be cost-effective using both cost-effectiveness thresholds (CETs): \$2,551 per disability-adjusted life year (DALY) averted and \$459 per DALY averted (Table 3)<sup>33–35</sup>. Cost-effectiveness for the combined and pulse oximetry models was predicted to improve in contexts with higher referral costs, as a result of their better specificity compared to the clinical model. Although specificity for the sTREM1 model was inferior to the clinical model, it was predicted to remain cost-effective up to a referral cost of approximately \$600 per patient, even when using the conservative CET (Supplementary Table 9). Neither the sTREM1 model nor the combined model was predicted to be more cost-effective than the pulse oximetry model at either of the analyzed CETs.



**Fig. 1 | Utility of the clinical prediction models within a hypothetical traffic light triage framework.** Among patients recommended for discharge (green; predicted probability (pp) of severe disease <0.5%), the sensitivity and NLR indicate the ability of the model to rule out severe disease. The percentage of patients recommended for monitoring (amber; pp 0.5–2%) indicates the potential burden on the health system created by the resources required for observation or proactive safety-netting. Among patients recommended

for referral (red; pp >2%), the specificity and PLR indicate the ability of the model to rule in severe disease, and the percentage and PLR indicate the potential burden on the health system created by the resources required for referral. Clinical model: clinical parameters only; pulse oximetry model: clinical parameters plus SpO<sub>2</sub>; sTREM1 model: clinical parameters plus sTREM1; combined model: clinical parameters plus SpO<sub>2</sub> and sTREM1. NLR, negative likelihood ratio; PLR, positive likelihood ratio.

## Discussion

This multicountry prospective study, conducted in seven locations across Asia, developed and validated clinical prediction models that outperform the current standard of care for triage of febrile children in resource-constrained community settings. Although all models were superior to WHO danger signs, integration of pulse oximetry or host biomarker testing alongside measurement of simple clinical parameters was required to adequately rule out impending critical illness, a key consideration for effective risk stratification at the community level. Notably, our analyses indicate that both strategies are likely to be cost-effective across a range of CETs, including in conflict-affected settings and remote contexts, where referral costs are often high.

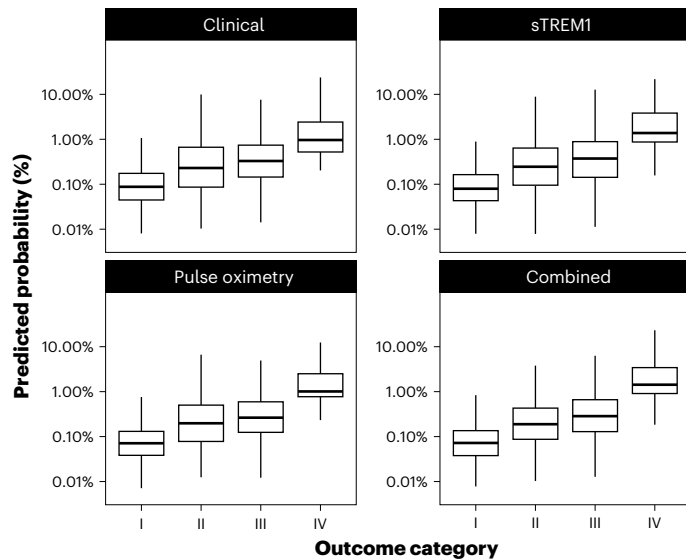
Our findings align with previous studies, which identified the prognostic potential of specific clinical parameters in pediatric febrile illness<sup>36–42</sup>, and build upon emerging evidence supporting hypoxemia and host biomarkers, such as sTREM1, as indicators of poor outcome<sup>17,20,25,43,44</sup>. The superior performance of the prediction models compared to previous univariate analyses underscores the importance of a multivariate approach to risk stratification<sup>30,41,45</sup>, whereby combinations of multiple ‘nearly’ signs can prompt action even if no ‘absolute’ red flags are present<sup>46</sup>.

The pulse oximetry model outperformed the sTREM1 model and may be the preferred option in many settings. Its adoption would be particularly suited to primary health centers and rural clinics that have access to safe storage, secure supply chains, supplemental oxygen therapy and staff trained to manage other conditions for which a pulse oximeter may provide benefit. By contrast, for lesser-trained community health workers, who typically have a limited scope of practice and are accustomed to screening all febrile children for malaria, a triage tool incorporating host biomarker testing may be more practical, particularly if multiplexed with a rapid diagnostic test for malaria<sup>31,47</sup>.

It is important to understand which patients at risk of severe illness may be missed by a particular triage approach. Our models appear to perform best in younger children with pneumonia—a group that carries a disproportionate share of the disease burden, indicating opportunity for substantial public health impact<sup>1</sup>. Conversely, children presenting earlier in their illness or deteriorating later may be more likely to be missed, suggesting a need for adjusted approaches to identify these children. Previous work showed that host biomarkers may be particularly valuable for identifying children whose illness severity is not clinically evident at presentation and who are at risk for delayed clinical deterioration<sup>30,48</sup>. In our study, both the pulse oximetry and sTREM1 models improved identification of such children relative to the clinical model, although small event numbers in this subgroup mean that further work is needed.

To minimize outcome misclassification, we applied a stringent definition of severe illness—restricted to children who died or received organ support within 2 days of enrollment—to develop the prediction models. Consequently, the reported NNTs and severe to non-severe referral ratios represent conservative estimates, emphasizing model performance in identifying children who progressed to the most severe illness. In practice, however, a substantially larger proportion of children would merit hospital referral. Secondary analyses using the categorical outcome indicate that the models provide discrimination across the severity spectrum, and further work is underway to quantify this performance.

This study developed and validated risk stratification tools for febrile children presenting directly from the community. We overcame weaknesses of previous studies by adhering to best-practice methodology for clinical prediction model building, prespecifying candidate predictors, avoiding overfitting and using a held-out geographic validation dataset to assess model performance<sup>49</sup>. Other strengths include eligibility criteria and a recruitment strategy that ensured a



**Fig. 2 | Predicted probability of severe febrile illness across the categorical outcome scale for each prediction model in the validation cohort.** Box plots show the median (center line), IQR (box; 25th–75th percentiles) and whiskers extending to the most extreme values within 1.5× the IQR. Observations beyond this range (outliers) are not shown. Clinical model: clinical parameters only; pulse oximetry model: clinical parameters plus SpO<sub>2</sub>; sTREM1 model: clinical parameters plus sTREM1; combined model: clinical parameters plus SpO<sub>2</sub> and sTREM1. Category I: managed as an outpatient and recovered by day 28 (*n* = 638). Category II: admission to any health facility for two or fewer nights between enrollment and day 28 (*n* = 779). Category III: admission to any health facility for more than two nights between enrollment and day 28 (*n* = 1,829). Category IV: death or organ support within 2 days of enrollment (*n* = 133).

study population and outcome prevalence reflective of community care settings<sup>50,51</sup>. Our analytical framework prioritized clinical utility, health system impact and cost-effectiveness over summary measures of model performance, such as the AUC<sup>52</sup>.

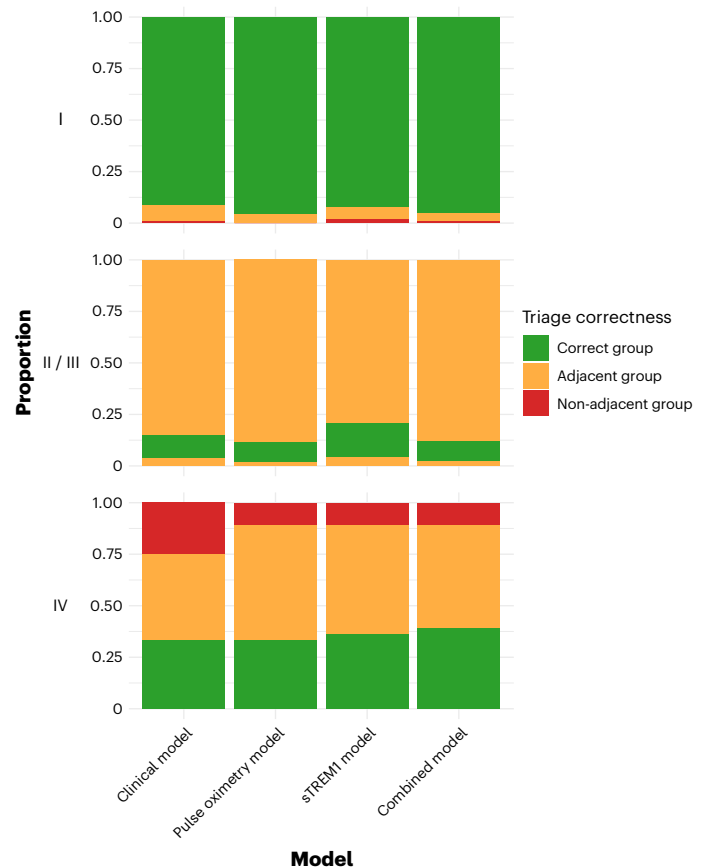
Several limitations must be discussed. Despite efforts to approximate community settings, our study population was recruited at rural hospital outpatient departments, and differences between these patients and those seeking care in more peripheral locations may remain. The relatively short interval to developing severe disease may indicate a higher baseline severity than is typical in some community care settings. Use of trained research assistants may have inflated the apparent performance of clinical parameters; in routine care settings where interobserver variability is high, it may be that incremental benefit of pulse oximetry and host biomarker testing could be even greater<sup>13</sup>. Respiratory presentations were highly represented in our cohort, and most children who developed severe disease required respiratory support. This may partly explain the observed benefits of pulse oximetry but, nonetheless, reflects the prevalent spectrum of common childhood infections.

Model calibration was suboptimal, likely due to differences in the prevalence of key clinical predictors, such as intractable vomiting and convulsions, between the derivation and validation cohorts. These discrepancies may reflect the well-documented interobserver variability in assessing WHO danger signs, suggesting that excluding these variables could improve model performance<sup>13</sup>. Although standardized assessments might mitigate this issue, in our study, trained research assistants used precise definitions; under routine care, variability is likely to be greater. Miscalibration primarily involved underestimation of risk among the highest-risk individuals; critically, the models would still have advised referral in these cases. Nonetheless, further external validation is needed, particularly given the relatively few events in the validation dataset (*n* = 36), which contributed to imprecision in

the effect estimates. We have published our full models to encourage independent validation.

In our cohort, malaria was uncommon, consistent with the evolving epidemiology of febrile illness across Asia<sup>53</sup>. Circulating markers of immune and endothelial activation, such as sTREM1, appear to capture convergent pathways of host response and enable risk stratification of pediatric fever syndromes irrespective of underlying etiology, including both malarial and non-malarial infections<sup>25,54,55</sup>. Similarly, hypoxemia has emerged as a robust predictor of adverse outcome across a spectrum of common childhood infections<sup>17,20</sup>. Nonetheless, evaluation of model performance in malaria-endemic regions will be essential to establish generalizability and clinical utility.

Among the clinical biomarker models, we selected the sTREM1 model for external validation because it consistently performed best across development and sensitivity analyses, retained face validity



**Fig. 3 | Proportional distribution of triage groups by patient category for each prediction model.** Proportional disposition of patients in each categorical outcome category is color coded to reflect triage appropriateness: green indicates triage to correct group; amber indicates triage to group adjacent to correct group; and red indicates triage to group non-adjacent to correct group. Category I (least severe): discharged patients (predicted probability (pp) <0.5%), green; observed patients (pp 0.5–2%), amber; referred patients (pp >2%), red. Categories II and III (moderately severe): discharged patients, green; observed patients, green; referred patients, amber. Category IV (most severe): discharged patients, red; observed patients, amber; referred patients, green. Within each outcome category (horizontal facet), patients are ordered as discharged, observed and referred, from top to bottom. Clinical model: clinical parameters only; pulse oximetry model: clinical parameters plus SpO<sub>2</sub>; sTREM1 model: clinical parameters plus sTREM1; combined model: clinical parameters plus SpO<sub>2</sub> and sTREM1. Category I: managed as an outpatient and recovered by day 28. Category II: admission to any health facility for two or fewer nights between enrollment and day 28. Category III: admission to any health facility for more than two nights between enrollment and day 28. Category IV: death or organ support within 2 days of enrollment.

**Table 3 | Cost effectiveness of the pulse oximetry, sTREM1 and combined models compared to the clinical model**

Model	Costs (\$)	DALYs	Incremental costs (\$)	DALYs averted	ICER (\$ per DALY averted)
Clinical	7.27	0.084	-	-	-
Pulse oximetry	7.95	0.058	0.69	0.026	\$26.28
sTREM1	12.39	0.058	5.13	0.026	\$196.46
Combined	12.11	0.058	4.85	0.026	\$185.76

Clinical model: clinical parameters only; pulse oximetry model: clinical parameters plus SpO<sub>2</sub>; sTREM1 model: clinical parameters plus sTREM1; combined model: clinical parameters plus SpO<sub>2</sub> and sTREM1.

from previous studies<sup>25,27,30,44</sup> and is under active development as a point-of-care assay<sup>36</sup>. Nonetheless, optimism-adjusted performance was similar across several candidate biomarkers, and other clinical biomarker models may prove equally or more effective. Limiting models to a single biomarker reduced overfitting and improves feasibility for application on the field, but future work should examine combinations of biomarkers, balancing potential gains in accuracy against increased cost and complexity. Including biomarkers found to be elevated in children who progressed to severe illness but who would have been missed by the current models (for example, ANG-1, CRP, IL-6 or IP-10) may be a fruitful approach.

We envisage these models being applied selectively to children for whom referral decisions are borderline and no overt signs of severe illness are evident at presentation. The exact 'red flags' triggering referral will differ across settings and are influenced by health worker capacity, safety-netting options, treatments available in the community, opportunity costs of referral and the quality of care at the higher-level facility. In our cohort, some children may be considered as presenting with features mandating referral and would not be appropriate candidates for model-based triage in all settings. Careful specification of referral criteria and further refinement and contextualization of the models will be essential prior to evaluation in clinical trials or integration into routine care.

The traffic light framework we used to illustrate clinical utility reflects the reality that, in most settings, some sort of safety-netting is possible<sup>46</sup>. Decision thresholds were selected pragmatically based on field experience but will require adaptation to local contexts. Future work should evaluate how different decision thresholds influence model performance and health system burden. Our results indicate that incorporating either pulse oximetry or host biomarker testing are likely to be cost-effective approaches to triage; however, the modeling outputs are inherently limited by assumptions—for example, that children recommended for monitoring and who develop severe disease can be recognized and escalated quickly—and assessing real-world impact will require interventional trials.

These findings should be interpreted in the context of the ongoing epidemiological transition in South and Southeast Asia. As malaria incidence declines, childhood febrile illness has become increasingly heterogeneous and is seldom attributable to a single, easily diagnosable and treatable cause. At first presentation, whether in primary care clinics, rural facilities, community-based services or conflict-affected settings, triage strategies must adapt to better identify children at risk of clinical deterioration. In such contexts, referral decisions entail substantial opportunity costs, yet failure to recognize severe illness early can be fatal. Strengthening prognostic capacity is, therefore, central to reinforcing primary care and advancing universal health coverage, with the potential to improve equity, reduce avoidable referrals and ensure timely escalation for the children most likely to benefit.

In summary, we developed and validated clinical prediction models that substantially improve risk stratification of febrile children compared to current practice. Incorporation of pulse oximetry or

sTREM1 measurements alongside simple clinical parameters achieves triage performance that may be acceptable to patients, providers and policymakers. Both approaches appear cost-effective across diverse contexts, including those with high referral costs. Further external validation and interventional studies will be essential to confirm clinical impact and support policy adoption.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41591-026-04338-1>.

## References

1. United Nations Inter-agency Group for Child Mortality Estimation. Levels and trends in child mortality. <https://data.unicef.org/resources/levels-and-trends-in-child-mortality-2023/> (2024).
2. Knappett, M. et al. Pediatric post-discharge mortality in resource-poor countries: a systematic review and meta-analysis. *EClinicalMedicine* **67**, 102380 (2024).
3. Debarre, A. *Hard to Reach: Providing Healthcare in Armed Conflict* [https://www.ipinst.org/wp-content/uploads/2018/12/1812\\_Hard-to-Reach.pdf](https://www.ipinst.org/wp-content/uploads/2018/12/1812_Hard-to-Reach.pdf) (International Peace Institute, 2018).
4. Kruk, M. E. et al. Mortality due to low-quality health systems in the universal health coverage era: a systematic analysis of amenable deaths in 137 countries. *Lancet* **392**, 2203–2212 (2018).
5. McDonald, C. R., Weckman, A., Richard-Greenblatt, M., Leligowicz, A. & Kain, K. C. Integrated fever management: disease severity markers to triage children with malaria and non-malarial febrile illness. *Malar. J.* **17**, 353 (2018).
6. Buntinx, F., Mant, D., Van den Bruel, A., Donner-Banzhof, N. & Dinant, G. J. Dealing with low-incidence serious diseases in general practice. *Br. J. Gen. Pract.* **61**, 43–46 (2011).
7. van den Brink, D. A. et al. Predicting clinical deterioration and mortality at differing stages during hospitalization: a systematic review of risk prediction models in children in low- and middle-income countries. *J. Pediatr.* **260**, 113448 (2023).
8. Eisenberg, M. A. & Balamuth, F. Pediatric sepsis screening in US hospitals. *Pediatr. Res.* **91**, 351–358 (2022).
9. Integrated management of childhood illnesses. [https://www.who.int/publications/m/item/integrated-management-of-childhood-illness---chart-booklet-\(march-2014\)](https://www.who.int/publications/m/item/integrated-management-of-childhood-illness---chart-booklet-(march-2014)) (World Health Organization, 2014).
10. WHO/UNICEF Joint Statement. Integrated community case management (iCCM). <https://cdn.who.int/media/docs/default-source/mca-documents/child/who-unicef-joint-statement-child-services-access.pdf> (World Health Organization, 2012).
11. Keitel, K., Kilowoko, M., Kyungu, E., Genton, B. & D'Acremont, V. Performance of prediction rules and guidelines in detecting serious bacterial infections among Tanzanian febrile children. *BMC Infect. Dis.* **19**, 769 (2019).
12. Tareke, A. A. et al. Adherence to IMCI guidelines for key symptoms in Ethiopian children: a 2021–2022 national service provision survey. *PLoS ONE* **19**, e0312138 (2024).
13. Hansoti, B. et al. Reliability and validity of pediatric triage tools evaluated in Low resource settings: a systematic review. *BMC Pediatr.* **17**, 37 (2017).
14. Biomarkers for acute febrile illness at the point-of-care in low-resource settings. [https://unitaid.org/uploads/Biomarkers-for-acute-febrile-illness-at-the-point-of-care-in-low-resource-settings\\_Meeting-report.pdf](https://unitaid.org/uploads/Biomarkers-for-acute-febrile-illness-at-the-point-of-care-in-low-resource-settings_Meeting-report.pdf) (World Health Organization, 2021).
15. Zar, H. J. & McCollum, E. D. Pulse oximetry to detect paediatric hypoxaemia—the fifth vital sign. *Lancet Glob. Health* **11**, e1684–e1685 (2023).

16. Wilkes, C., Bava, M., Graham, H. R., Duke, T. & group, A. R. I. R. What are the risk factors for death among children with pneumonia in low- and middle-income countries? A systematic review. *J Glob. Health* **13**, 05003 (2023).
17. Chew, R., Zhang, M., Chandna, A. & Lubell, Y. The impact of pulse oximetry on diagnosis, management and outcomes of acute febrile illness in low-income and middle-income countries: a systematic review. *BMJ Glob. Health* **6**, e007282 (2021).
18. Karunaratne, R., Sturgeon, J. P., Patel, R. & Prendergast, A. J. Predictors of inpatient mortality among children hospitalized for severe acute malnutrition: a systematic review and meta-analysis. *Am. J. Clin. Nutr.* **112**, 1069–1079 (2020).
19. McCollum, E. D. et al. Pulse oximetry for children with pneumonia treated as outpatients in rural Malawi. *Bull. World Health Organ.* **94**, 893–902 (2016).
20. McCollum, E. D. et al. Risk and accuracy of outpatient-identified hypoxaemia for death among suspected child pneumonia cases in rural Bangladesh: a multifacility prospective cohort study. *Lancet Respir. Med.* **11**, 769–781 (2023).
21. Hooli, S. et al. Pulse oximetry accuracy in children with dark skin tones: relevance to acute lower respiratory infection care in low- and middle-income countries. *Am. J. Trop. Med. Hyg.* **111**, 736–739 (2024).
22. King, C. et al. Opportunities and barriers in paediatric pulse oximetry for pneumonia in low-resource clinical settings: a qualitative evaluation from Malawi and Bangladesh. *BMJ Open* **8**, e019177 (2018).
23. Baker, K. et al. Performance of five pulse oximeters to detect hypoxaemia as an indicator of severe illness in children under five by frontline health workers in low resource settings—a prospective, multicentre, single-blinded, trial in Cambodia, Ethiopia, South Sudan, and Uganda. *EClinicalMedicine* **38**, 101040 (2021).
24. Balanza, N. et al. Host-based prognostic biomarkers to improve risk stratification and outcome of febrile children in low- and middle-income countries. *Front. Pediatr.* **8**, 552083 (2020).
25. Leligdowicz, A. et al. Risk-stratification of febrile African children at risk of sepsis using sTREM-1 as basis for a rapid triage test. *Nat. Commun.* **12**, 6832 (2021).
26. McDonald, C. R. et al. Immune and endothelial activation markers and risk stratification of childhood pneumonia in Uganda: a secondary analysis of a prospective cohort study. *PLoS Med.* **19**, e1004057 (2022).
27. Jullien, S. et al. Performance of host-response biomarkers to risk-stratify children with pneumonia in Bhutan. *J. Infect.* **85**, 634–643 (2022).
28. Singla, M. et al. Immune response to dengue virus infection in pediatric patients in New Delhi, India—association of viremia, inflammatory mediators and monocytes with disease severity. *PLoS Negl. Trop. Dis.* **10**, e0004497 (2016).
29. Carrol, E. D. et al. Procalcitonin as a diagnostic marker of meningococcal disease in children presenting with fever and a rash. *Arch. Dis. Child.* **86**, 282–285 (2002).
30. Chandna, A. et al. Risk stratification of childhood infection using host markers of immune and endothelial activation in Asia (Spot Sepsis): a multi-country, prospective, cohort study. *Lancet Child Adolesc. Health* **9**, 634–645 (2025).
31. Visser, M. T. et al. Operational evaluation of the deployment of Malaria/CRP Duo and Dengue Duo rapid diagnostic tests for the management of febrile illness by village malaria workers in rural Cambodia. *BMC Infect. Dis.* **25**, 679 (2025).
32. Ebbs, D., Taricia, M., Funaro, M. C., O’Daniel, M. & Cappello, M. Prehospital use of point-of-care tests by community health workers: a scoping review. *Front. Public Health* **12**, 1360322 (2024).
33. Kazibwe, J. et al. The use of cost-effectiveness thresholds for evaluating health interventions in low- and middle-income countries from 2015 to 2020: a review. *Value Health* **25**, 385–389 (2022).
34. Woods, B., Revill, P., Sculpher, M. & Claxton, K. Country-level cost-effectiveness thresholds: initial estimates and the need for further research. *Value Health* **19**, 929–935 (2016).
35. World Bank Group. GDP per capita (current US\$) - Bangladesh. <https://data.worldbank.org/indicator/NY.GDP.PCAP.CD?locations=BD> (accessed 7 May 2025).
36. Aramburo, A. et al. Lactate clearance as a prognostic marker of mortality in severely ill febrile children in East Africa. *BMC Med.* **16**, 37 (2018).
37. Mtove, G. et al. Point-of-care measurement of blood lactate in children admitted with febrile illness to an African District Hospital. *Clin. Infect. Dis.* **53**, 548–554 (2011).
38. Conroy, A. L. et al. Prospective validation of pediatric disease severity scores to predict mortality in Ugandan children presenting with malaria and non-malaria febrile illness. *Crit. Care* **19**, 47 (2015).
39. van Nassau, S. C. et al. Translating Sepsis-3 criteria in children: prognostic accuracy of age-adjusted Quick SOFA score in children visiting the emergency department with suspected bacterial infection. *Front. Pediatr.* **6**, 266 (2018).
40. Scott, H. F., Donoghue, A. J., Gaieski, D. F., Marchese, R. F. & Mistry, R. D. Effectiveness of physical exam signs for early detection of critical illness in pediatric systemic inflammatory response syndrome. *BMC Emerg. Med.* **14**, 24 (2014).
41. Chandna, A. et al. Predictors of disease severity in children presenting from the community with febrile illnesses: a systematic review of prognostic studies. *BMJ Glob. Health* **6**, e003451 (2021).
42. George, E. C. et al. Predicting mortality in sick African children: the FEAST Paediatric Emergency Triage (PET) Score. *BMC Med.* **13**, 174 (2015).
43. Graham, H. R. et al. The prevalence of hypoxaemia in paediatric and adult patients in health-care facilities in low-income and middle-income countries: a systematic review and meta-analysis. *Lancet Glob. Health* **13**, e222–e231 (2025).
44. Balanza, N. et al. Prognostic accuracy of biomarkers of immune and endothelial activation in Mozambican children hospitalized with pneumonia. *PLoS Glob. Public Health* **3**, e0001553 (2023).
45. Van den Bruel, A., Haj-Hassan, T., Thompson, M., Buntinx, F. & Mant, D. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet* **375**, 834–845 (2010).
46. Graham, H. R. et al. Hypoxaemia and risk of death among children: rethinking oxygen saturation, risk-stratification, and the role of pulse oximetry in primary care. *Lancet Glob. Health* **12**, e1359–e1364 (2024).
47. Lubell, Y. et al. Economic considerations support C-reactive protein testing alongside malaria rapid diagnostic tests to guide antimicrobial therapy for patients with febrile illness in settings with low malaria endemicity. *Malar. J.* **18**, 442 (2019).
48. Chandna, A. et al. Defining the role of host biomarkers in the diagnosis and prognosis of the severity of childhood pneumonia: a prospective cohort study. *Sci. Rep.* **13**, 12024 (2023).
49. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann. Intern. Med.* **162**, 55–63 (2015).
50. Holtman, G. A. et al. Development of practical recommendations for diagnostic accuracy studies in low-prevalence situations. *J. Clin. Epidemiol.* **114**, 38–48 (2019).
51. Leeflang, M. M., Bossuyt, P. M. & Irwig, L. Diagnostic test accuracy may vary with prevalence: implications for evidence-based diagnosis. *J. Clin. Epidemiol.* **62**, 5–12 (2009).

52. de Hond, A. A. H., Steyerberg, E. W. & van Calster, B. Interpreting area under the receiver operating characteristic curve. *Lancet Digit. Health* **4**, e853–e855 (2022).
53. Shrestha, P. et al. Non-malarial febrile illness: a systematic review of published aetiological studies and case reports from Southern Asia and South-eastern Asia, 1980–2015. *BMC Med.* **18**, 299 (2020).
54. Jolly, L. et al. sTREM-1 is a specific biomarker of TREM-1 pathway activation. *Cell Mol. Immunol.* **18**, 2054–2056 (2021).
55. Leligdowicz, A., Richard-Greenblatt, M., Wright, J., Crowley, V. M. & Kain, K. C. Endothelial activation: the Ang/Tie axis in sepsis. *Front. Immunol.* **9**, 838 (2018).
56. Echilibrist Project. ECHILIBRIST: enhancing children's lives with biomarkers for risk stratification and triage. <https://echilibrist.org> (accessed 5 September 2025).
57. Myatt, M. & Guevarra, E. zscorer: Child anthropometry z-score calculator. R package version 0.3.1. CRAN <https://cran.r-project.org/web/packages/zscorer/index.html> (2019).

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

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2026

**Arjun Chandna**<sup>1,2,3</sup>✉, **Constantinos Koshiaris**<sup>4,5</sup>, **Raman Mahajan**<sup>6</sup>, **Riris Adono Ahmad**<sup>7</sup>, **Dinh Thi Van Anh**<sup>8</sup>, **Khalid Shams Choudhury**<sup>6</sup>, **Suy Keang**<sup>1,9</sup>, **Nguyen The Nguyen Phung**<sup>10</sup>, **Sayaphet Rattanavong**<sup>11</sup>, **Souphaphone Vannachone**<sup>11</sup>, **Spot Sepsis Investigator Group\***, **Chris Painter**<sup>12,11,12</sup>, **Mikhael Yosia**<sup>6</sup>, **Naomi Waithira**<sup>2,12</sup>, **Mohammad Yazid Abdad**<sup>2,11,15</sup>, **Janjira Thaipadungpanit**<sup>12,13</sup>, **Paul Turner**<sup>1,2</sup>, **Phan Huu Phuc**<sup>8</sup>, **Dinesh Mondal**<sup>14</sup>, **Mayfong Mayxay**<sup>2,11,15</sup>, **Bui Thanh Liem**<sup>10</sup>, **Elizabeth A. Ashley**<sup>2,11</sup>, **Eggi Arguni**<sup>7</sup>, **Rafael Perera-Salazar**<sup>4</sup>, **Melissa Richard-Greenblatt**<sup>2,16,17</sup>, **Yoel Lubell**<sup>2,12,18,29</sup> & **Sakib Burza**<sup>3,6,19,29</sup>

<sup>1</sup>Cambodia Oxford Medical Research Unit, Angkor Hospital for Children, Siem Reap, Cambodia. <sup>2</sup>Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, UK. <sup>3</sup>Clinical Research Department, London School of Hygiene and Tropical Medicine, London, UK. <sup>4</sup>Department of Primary Care Health Sciences, University of Oxford, Oxford, UK. <sup>5</sup>Department of Primary Care and Population Health, University of Nicosia Medical School, Nicosia, Cyprus. <sup>6</sup>Médecins Sans Frontières Operational Centre Barcelona, Barcelona, Spain. <sup>7</sup>Centre for Tropical Medicine, Universitas Gadjah Mada, Yogyakarta, Indonesia. <sup>8</sup>Viet Nam National Children's Hospital, Hanoi, Vietnam. <sup>9</sup>Angkor Hospital for Children, Siem Reap, Cambodia. <sup>10</sup>University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam. <sup>11</sup>Lao-Oxford-Mahosot Hospital-Wellcome Trust Research Unit, Mahosot Hospital, Vientiane, Laos. <sup>12</sup>Mahidol Oxford Tropical Medicine Research Unit, Mahidol University, Bangkok, Thailand. <sup>13</sup>Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. <sup>14</sup>International Centre for Diarrhoeal Disease Research, Dhaka, Bangladesh. <sup>15</sup>Institute for Research and Education Development, University of Health Sciences, Vientiane, Laos. <sup>16</sup>Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada. <sup>17</sup>Department of Pediatric Laboratory Medicine, The Hospital for Sick Children, Toronto, Ontario, Canada. <sup>18</sup>Amsterdam Institute of Global Health and Development, Amsterdam, The Netherlands. <sup>19</sup>Health In Harmony, Portland, OR, USA. <sup>29</sup>These authors contributed equally: Yoel Lubell, Sakib Burza. \*A list of authors and their affiliations appears at the end of the paper. ✉e-mail: [arjun@tropmedres.ac](mailto:arjun@tropmedres.ac)

### Spot Sepsis Investigator Group

**Mohammad Yazid Abdad**<sup>2,12</sup>, **Riris Adono Ahmad**<sup>7</sup>, **Dinh Thi Van Anh**<sup>8</sup>, **Eggi Arguni**<sup>7</sup>, **Elizabeth A. Ashley**<sup>2,11</sup>, **Elizabeth M. Batty**<sup>2,12</sup>, **Stuart D. Blacksell**<sup>2,12</sup>, **Latsaniphone Boutthasavong**<sup>11</sup>, **Sakib Burza**<sup>3,6,19,29</sup>, **Arjun Chandna**<sup>1,2,3</sup>, **Ngoun Chanpheaktra**<sup>9</sup>, **Khalid Shams Choudhury**<sup>6</sup>, **Tran Quoc Dat**<sup>8</sup>, **Vu Quoc Dat**<sup>20</sup>, **Nicholas P. J. Day**<sup>2,12</sup>, **Arjen M. Dondorp**<sup>2,12</sup>, **Prakash Ghosh**<sup>14,21,22</sup>, **Carolina Jimenez**<sup>6</sup>, **Kevin Kain**<sup>16</sup>, **Muhammad Karyana**<sup>23</sup>, **Suy Keang**<sup>1,9</sup>, **Sommay Keomany**<sup>24</sup>, **Rungnapa Khamboocha**<sup>12</sup>, **Constantinos Koshiaris**<sup>4,5</sup>, **Khamfong Kunlaya**<sup>11</sup>, **Estrella Lasry**<sup>6</sup>, **Bui Thanh Liem**<sup>10</sup>, **Nguyen Huy Luan**<sup>10</sup>, **Yoel Lubell**<sup>2,12,18,29</sup>, **Raman Mahajan**<sup>6</sup>, **Saysamone Malavong**<sup>25</sup>, **Mayfong Mayxay**<sup>2,11,15</sup>, **Chonticha Menggred**<sup>12</sup>, **Dinesh Mondal**<sup>14</sup>, **Phung Nguyen The Nguyen**<sup>10</sup>, **Chris Painter**<sup>2,11,12</sup>, **Rafael Perera-Salazar**<sup>5</sup>, **Chom Phaiphichit**<sup>11</sup>, **Chanthala Phamisith**<sup>25</sup>, **Phan Huu Phuc**<sup>8</sup>, **Tiengkham Pongvongsa**<sup>26</sup>, **Sayaphet Rattanavong**<sup>11</sup>, **Michael Rekart**<sup>6</sup>, **Melissa Richard-Greenblatt**<sup>2,16,17</sup>, **Bran Sambou**<sup>1</sup>, **Mohammad Shomik**<sup>14</sup>, **Phouthalavanh Souvannasing**<sup>24</sup>, **Phattaranit Tanunchai**<sup>12</sup>, **Janjira Thaipadungpanit**<sup>12,13</sup>, **Watcharintorn Thongpiam**<sup>12</sup>, **Bang Huyen Tran**<sup>27</sup>, **Claudia Turner**<sup>1,2,9</sup>, **Paul Turner**<sup>1,2</sup>, **Souphaphone Vannachone**<sup>11</sup>, **Asama Vinitorn**<sup>12</sup>, **Ranitha Vongpromek**<sup>12</sup>, **Manivanh Vongsouvath**<sup>11</sup>, **Naomi Waithira**<sup>2,12</sup>, **James A. Watson**<sup>2,12</sup>, **Mikhael Yosia**<sup>6</sup> & **Asri Yuniastuti**<sup>28</sup>

<sup>20</sup>Hanoi Medical University, Hanoi, Vietnam. <sup>21</sup>Institute of Animal Hygiene and Veterinary Public Health, Leipzig University, Leipzig, Germany. <sup>22</sup>Technische Universität Berlin, Berlin, Germany. <sup>23</sup>INA-RESPOND, Ministry of Health Republic Indonesia, Jakarta, Indonesia. <sup>24</sup>Salavan Provincial Hospital, Salavan, Laos. <sup>25</sup>Savannakhet Provincial Hospital, Savannakhet, Laos. <sup>26</sup>Savannakhet Provincial Health Office, Savannakhet, Laos. <sup>27</sup>Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam. <sup>28</sup>Wates District Hospital, Yogyakarta, Indonesia.

## Methods

### Study design and setting

Spot Sepsis was a multicountry, prospective cohort study conducted in seven hospitals across Bangladesh, Cambodia, Indonesia, Laos and Vietnam<sup>58</sup>. Study sites were purposefully chosen to represent facilities serving rural populations and providing a first point of access to the formal healthcare sector, as a proxy for the ultimate intended use-case in community care settings. Previous analyses reported the prognostic performance of individual host biomarkers<sup>30</sup>. In the analyses described herein, data from six sites were used to derive the clinical prediction models, with the site in Cambodia prespecified a priori for held-out external geographic validation. This split was pragmatic, based on the ability to conduct recruitment for longer in Cambodia.

### Study site profiles

Sites located outside major urban cities were proactively approached. As part of the shortlisting process, sites were asked to indicate the proportion of their patient population residing in a rural location and the proportion of children using the hospital as a first point of contact with the formal healthcare sector. Where possible, routinely collected data were used to inform site selection. If these were not available, local clinicians and hospital administrators were asked to provide their best estimate.

Due to disruptions caused by the COVID-19 pandemic, site activation was delayed, and recruitment proceeded slower than anticipated at all sites. After 18 months of recruitment, a decision was taken to identify an additional site to boost recruitment and ensure sufficient outcome events. The second Vietnam site was subsequently identified and brought online in December 2021, acknowledging that it departed from the rural site target profile established at the outset of the study. A sensitivity analysis excluding this site was prespecified to address this.

#### **Bangladesh—recruitment period: 18 March 2021 to 27 April 2022.**

Goyalpara Mother and Child Hospital is a non-governmental hospital managed by Médecins sans Frontières (MSF) located in Cox's Bazaar, Chattogram Division, in eastern Bangladesh, which predominantly provides health services to the forcibly displaced Rohingya refugee population. The hospital provides primary and secondary care, with approximately 20,000 outpatient attendances and 4,000 inpatient admissions annually. The 12-bed high-dependency unit provides non-invasive ventilation and inotropic therapy. There is a basic on-site diagnostic laboratory.

#### **Cambodia—recruitment period: 5 March 2020 to 24 February 2022.**

Angkor Hospital for Children is a non-governmental pediatric hospital located in Siem Reap province, northern Cambodia. The hospital provides primary to tertiary care, with approximately 80,000 outpatient attendances and 3,000 inpatient admissions annually. The 14-bed intensive care unit provides non-invasive ventilation, mechanical ventilation, inotropic therapy and peritoneal dialysis. There is an on-site diagnostic microbiology laboratory (ISO15189 accredited since 23 November 2023).

#### **Indonesia—recruitment period: 22 March 2021 to 22 April 2021.**

Rumah Sakit Umum Daerah Wates is a government district hospital located in Yogyakarta province, Indonesia. The hospital provides primary and secondary care, with approximately 40,000 outpatient attendances and 4,000 inpatient admissions annually. The eight-bed intensive care unit provides non-invasive ventilation, mechanical ventilation, inotropic therapy and renal replacement therapy. Basic microscopy is available on-site. Culture-based microbiology is available via nearby private laboratories.

#### **Laos 1—recruitment period: 10 September 2020 to 30 August 2021.**

Salavan Provincial Hospital is the government provincial hospital for Salavan, a predominantly rural province in southern Laos.

The hospital provides primary to tertiary care, with approximately 65,000 outpatient attendances and 12,000 inpatient admissions annually. The three-bed pediatric intensive care unit provides non-invasive ventilation, mechanical ventilation and inotropic therapy. There is an on-site diagnostic microbiology laboratory.

#### **Laos 2—recruitment period: 21 January 2021 to 26 August 2021.**

Savannakhet Provincial Hospital is the government provincial hospital for Savannakhet, a predominantly rural province in southern Laos. The hospital provides primary to tertiary care, with approximately 7,000 outpatient attendances and 3,000 inpatient admissions annually. The 22-bed intensive care unit provides non-invasive ventilation, mechanical ventilation and inotropic therapy. There is an on-site diagnostic microbiology laboratory.

#### **Vietnam 1—recruitment period: 10 May 2021 to 28 October 2022.**

Dong Nai Children's Hospital is the government provincial pediatric hospital for Dong Nai province in southern Vietnam. The hospital provides primary to tertiary care, with approximately 80,000 outpatient attendances and 7,000 inpatient admissions annually. The 30-bed intensive care unit provides non-invasive ventilation, mechanical ventilation, inotropic therapy and renal replacement therapy. There is an on-site diagnostic microbiology laboratory.

#### **Vietnam 2—recruitment period: 8 December 2021 to 4 November 2022.**

Vietnam National Children's Hospital is the government national pediatric hospital, located in Ha Noi. The hospital provides primary to quaternary care, with approximately 1,200,000 outpatient attendances and 110,000 inpatient admissions annually. The 40-bed intensive care unit provides non-invasive ventilation, mechanical ventilation, inotropic therapy and renal replacement therapy. There is an on-site diagnostic microbiology laboratory.

### Study participants

Children (aged >28 days and <60 months) presenting with a febrile illness (axillary temperature  $\geq 37.5^\circ\text{C}$  or  $< 35.5^\circ\text{C}$  or history of fever in the preceding 24 hours)<sup>59</sup> of  $\leq 14$  days were eligible for inclusion. Exclusion criteria were prior admission to any health facility during the current illness, receipt of >15 minutes of parenteral treatment (intravenous, intramuscular or nebulized medication, intravenous fluids or supplemental oxygen) prior to screening<sup>60</sup>, presentation within 3 days of routine immunizations, trauma as the reason for attendance or specific known comorbidities (chronic infection (for example, viral hepatitis or tuberculosis), immunosuppression (for example, HIV or oncological conditions) or active (symptomatic or currently medicated) cardiorespiratory conditions). Participants could be enrolled only once.

### Screening and enrollment

Patients were screened during daytime working hours. Screening was stratified by admission status. Admissions could occur via the emergency or outpatient departments. Inpatients were screened consecutively upon arrival at the emergency department or inpatient ward. Children admitted from the outpatient department typically did not receive treatment before arrival on the inpatient ward. Children attending the emergency department typically received treatment promptly and were either admitted or kept for a period of observation. Thus, all patients presenting to the emergency department (both those admitted and those kept for a period of observation) were considered inpatients.

Due to high outpatient numbers, outpatient screening was randomized. Outpatients to be approached were randomly allocated by computer-generated random number tables, with the preceding week's routinely collected hospital attendance data providing the sampling frame. Screening of outpatients occurred prior to assessment by the clinical team, with three outpatients targeted for recruitment each

week at each site. On rare occasions when the health worker decided to admit a child just recruited into the outpatient strata, the participant was transferred to the inpatient strata, and the next patient specified on the random number table was approached for screening.

### Data collection

Trained study personnel measured vital signs (including SpO<sub>2</sub> via pulse oximetry) and anthropometrics, assessed clinical signs (including WHO danger signs) and collected venous blood samples and nasopharyngeal swabs at enrollment (Supplementary Table 10). Demographic information and perinatal, past medical and illness histories were collected via interview with the participant's caregiver. All data were entered onto electronic case record forms using Android tablets via Open Data Kit Collect software (version 1.25.1). Participants were provided with routine care by their treating clinician. When feasible, the study supported collection and processing of peripheral blood cultures at the discretion of the clinical team.

Participants were followed-up on days 2 and 28 after enrollment, with additional follow-up on day 1 and at discharge for inpatients. Follow-up was conducted in-person if the participant remained admitted at the study site or by telephone if they had been discharged. Study monitoring was conducted by the Clinical Trials Support Group at the Mahidol Oxford Tropical Medicine Research Unit (MORU) in Bangkok, Thailand.

### Clinical parameters and host biomarkers

Clinical parameters useful for the prognostication of pediatric febrile illness were identified through systematic review of the literature<sup>41</sup>. Feasibility of measurement in resource-constrained community settings was considered<sup>61</sup>, with only the most practicable parameters retained. Prioritization and standardization followed guidance set out by the Pediatric Sepsis Predictors Standardization working group<sup>62</sup>, operationalized through standard operating procedures (SOPs) on which all study personnel were trained prior to beginning data collection. Adherence to SOPs was assessed at site initiation visits and regular monitoring visits. Study equipment was procured centrally and distributed to sites.

Host response biomarkers were shortlisted after review of the literature and consultation with domain experts (Supplementary Table 11). In primary care, where the cause of infection is typically unknown at the time of presentation, biomarkers that are predictive across a range of pathogens are essential for risk stratification. Therefore, biomarkers implicated in final common pathways to severe febrile illness and sepsis were prioritized. Endothelial activation markers included ANG-1 (refs. 26,55,63–66), ANG-2 (refs. 25,26,48,55,63–68) and soluble FLT-1 (sFLT-1; also known as soluble VEGFR-1)<sup>25,26,44,67,68</sup>. Immune activation markers included CHI3L1 (refs. 25,26,68), CRP<sup>69</sup>, IL-1ra<sup>70–72</sup>, IL-6 (refs. 28,71,73), IL-8 (refs. 28,44,70), IL-10 (ref. 28), IP-10 (also known as CXCL10)<sup>67</sup>, procalcitonin (PCT)<sup>29,74</sup>, soluble TNF-R1 (sTNF-R1)<sup>25,26,68</sup>, sTREM1 (refs. 25–27,44,67,68,75–77) and soluble uPAR (suPAR)<sup>78–81</sup>. Lactate, glucose and hemoglobin were also included because they are easily measured using inexpensive rapid tests, are well known to clinicians, have prognostic value<sup>41</sup> and are recommended in pediatric sepsis guidelines<sup>59,82,83</sup>.

### Laboratory procedures

Venous blood samples and nasopharyngeal swabs were processed immediately. Complete blood counts were conducted on-site, and peripheral blood cultures were processed at in-country laboratories. Aliquots of whole blood, EDTA plasma, fluoride/oxalate plasma and universal transport medium were stored at –20 °C or lower. These samples were then transported at –80 °C to the MORU laboratories in Bangkok, Thailand, for further analysis and biobanking.

Biomarker concentrations in EDTA plasma were measured using the Simple Plex Ella microfluidic platform (ProteinSimple)

and the suPARnostic ELISA (ViroGates), as outlined in the appendix (Supplementary Table 12). Glucose (GLUC3; Roche Diagnostics) and lactate (LACT2; Roche Diagnostics) concentrations were quantified in fluoride/oxalate plasma.

Nucleic acid was extracted from whole blood using the MagNA Pure 24 instrument and Total NA Isolation Kit (Roche Diagnostics), following the manufacturer's instructions. Real-time polymerase chain reaction (RT-PCR) multiplex assays were used to detect viral (chikungunya, dengue, Japanese encephalitis and Zika) and bacterial (*Leptospira* spp., *Orientia tsutsugamushi* and *Rickettsia* spp.) targets in whole blood. Respiratory pathogen targets were detected directly from nasopharyngeal swabs using the FilmArray RP2 panel (BioFire Diagnostics), except for samples from Cambodia, which were processed for influenza A/B and respiratory syncytial virus (RSV) using the FTD FLU/HRSV assay (Siemens). Specimens from all sites were tested using an in-house multiplex RT-PCR assay to detect SARS-CoV-2 in nasopharyngeal swabs, based on the E and N genes, as previously described<sup>84</sup>.

### Outcome measures

The primary outcome was development of severe febrile illness within 2 days of enrollment, defined as death or receipt of organ support (mechanical ventilation, non-invasive ventilation, inotropic therapy or renal replacement therapy).

Recognizing that illness severity is a continuum, participant outcomes were also classified on a categorical scale (Supplementary Table 7). This was used to further assess the performance of the prediction models and provide insight into the relative importance of misclassifications.

### Sample size

The methods of Riley et al.<sup>85</sup> were followed, and, using an anticipated outcome prevalence of 1%, conservative  $R^2$  Nagelkerke of 0.15 and shrinkage factor of 0.9, it was estimated that six events per parameter (EPP) would be required to derive the prediction models. The COVID-19 pandemic, declared in the same week that recruitment began at the first site, delayed initiation of other sites and slowed enrollment. Study duration (and, thus, sample size) was determined by available resources, with recruitment continuing for as long as funding allowed. Accordingly, as there were 97 events in the derivation dataset, a maximum of 16 candidate parameters could be included for development of the prediction models while minimizing the risk of overfitting. The combined model (including clinical parameters, SpO<sub>2</sub> and one host biomarker) comprised the largest number of candidate predictors (13 parameters). Thirty-six events were accrued in the validation dataset. Although clinically and statistically significant results were obtained, effect estimates are relatively imprecise. We provide our model equations to encourage further independent external validation.

### Alteration to the sample size calculation

Sample size calculations for clinical prediction model development depend, in part, on outcome prevalence, with more events required in high-prevalence settings to avoid overfitting<sup>86</sup>. In the published protocol, the raw (unweighted) anticipated outcome prevalence (13%) was erroneously used in the original sample size calculation, resulting in an estimate of 14 EPP for derivation of the prediction models<sup>58</sup>. This mistake was recognized while the study was still recruiting. At this point, and before any data were analyzed, the sample size was recalculated using the correct (weighted) anticipated outcome prevalence of 1%. This resulted in an estimate of six EPP for derivation of the prediction models.

### Statistical methods

Categorical and continuous variables were summarized using descriptive statistics for both the derivation and validation cohorts and

compared between participants according to their outcome status, using the Wilcoxon rank-sum test, Pearson's  $\chi^2$  test or Fisher's exact test as appropriate. Site-specific outpatient weights were applied to adjust for unequal probabilities of selection in the sample, arising due to random sampling of outpatients (Supplementary Table 13), to ensure that the study population was representative of all eligible children presenting to the hospital and that the outcome prevalence reflected that which might be observed in community care settings<sup>50,51</sup>.

Candidate predictors were prespecified using domain knowledge and literature review (Supplementary Table 2). The relationship between each continuous candidate predictor and the primary outcome was explored by visually inspecting a smoothing curve to identify nonlinear patterns and by examining the overall calibration of the model. Transformations were considered only for serious violations of linearity that were substantially affecting model calibration. There were few missing data ( $\leq 3.6\%$ ) for all predictors except hemoglobin, where 16.9% (574/3,405) of participants had missing data, as complete blood counts were collected at the discretion of the treating clinical team. For all predictors, missing observations were replaced with their median value, conditional on outcome status.

To build the clinical model, all 11 prespecified candidate clinical predictors (excluding SpO<sub>2</sub>) were entered into the model, and backward stepwise weighted logistic regression was used to identify the most parsimonious model. This process was repeated to build the pulse oximetry model, starting with all prespecified candidate clinical predictors, this time including SpO<sub>2</sub>. To identify the most promising biomarker, clinical biomarker models were developed using backward stepwise weighted logistic regression with all prespecified candidate clinical predictors (with and without SpO<sub>2</sub>) and each of the 17 biomarkers in turn. The models were internally validated using bootstrapping, with 800 samples being drawn with replacement, to adjust for optimism. A sensitivity analysis excluding the northern Vietnam site ( $n = 612$ ), which departed from the ideal rural target site profile and where outpatient weighting was derived using different methodology (Supplementary Table 13), was performed. The clinical biomarker model with highest discrimination was carried forward for external validation. Overall, a total of four models (clinical model, pulse oximetry model, a clinical biomarker model and a combined model including the biomarker and SpO<sub>2</sub>) were taken forward for validation in the validation cohort.

Models were applied to the validation dataset and the predicted probabilities estimated. Model performance was evaluated in terms of discrimination and calibration. Discrimination—the ability of a model to distinguish between individuals who did and did not experience the outcome—was assessed using the weighted AUC, with values ranging from 0.5 (no discrimination) to 1.0 (perfect discrimination). Calibration—the agreement between predicted and observed risks—was assessed using the weighted calibration intercept and slope. An ideal model has an intercept of 0 (with positive values indicating overall risk underestimation and negative values indicating overestimation) and a slope of 1 (where slopes  $>1$  suggest underestimation in low-risk individuals and overestimation in high-risk individuals, and slopes  $<1$  indicate the opposite pattern).

Clinical utility of each model was assessed within a traffic light triage framework by evaluating classification (sensitivity, specificity, NLR and PLR) at clinically relevant decision thresholds (predicted probabilities of severe febrile illness): any child with a predicted probability of developing severe disease  $<0.5\%$  (the rule-out threshold) is discharged (green); those with a predicted probability  $>2\%$  (the rule-in threshold) are referred for higher-level care (red); and those with predicted probabilities between 0.5% and 2% are monitored (amber). Differences in classification indices were estimated with 95% CI derived using bootstrap resampling procedures.

A deterministic impact analysis was applied to a hypothetical population of 10,000 children, using fixed point estimates of

sensitivity, specificity and disease prevalence obtained from the validation cohort, to calculate projected triage dispositions and NNT to identify one additional child who would progress to develop severe disease.

Finally, for each model, the distribution of predicted probabilities and triage groups was examined in relation to the categorical outcome scale, to better understand the potential implications of misclassifications. Characteristics of participants who developed severe febrile illness but were missed by the models were explored in order to understand model vulnerabilities and determine what adjustments may lead to improved model performance.

Data management was performed using R version 4.3.2, and all analyses were performed using STATA version 18.

### Cost-effectiveness analyses

A cost-effectiveness analysis of the different models for the triage of febrile children was conducted, set in the context of Bangladesh, which was the study site most closely aligned with the intended use setting for the referral tool (Extended Data Fig. 3). The analyses were parameterized using primary data from the study and published estimates available in the literature (Supplementary Table 14)<sup>87–92</sup>. Two CETs were explored: \$2,551 per DALY averted (approximating Bangladesh's 2023 gross domestic product per capita)<sup>33,35</sup> and a more conservative CET of \$459 per DALY averted obtained from Woods et al.<sup>34</sup> (inflated to 2024 values). A sensitivity analysis was conducted to determine the maximum referral cost at which the different models might remain cost-effective, for both of the analyzed CETs.

### Inclusion and ethics statement

The study was designed collaboratively by investigators from all participating countries during an inaugural meeting co-convened by MSF and MORU and held in Bangkok, Thailand, in May 2019. During the meeting, the study hypotheses and design were discussed to ensure relevance of the research to the study settings. The draft protocol was then iteratively improved based on feedback from stakeholders, healthcare providers and the Young Persons Advisory Group at Angkor Hospital for Children, Cambodia. Throughout the study, capacity building activities were a priority and included training of research teams, tailored support for blood culture infrastructure and opportunities for junior team members to present findings at internal and external meetings. Regional capacity for biomarker assays was developed at the MORU laboratories in Bangkok, Thailand. Study progress was reviewed at monthly co-investigator meetings, where feedback from the site teams was discussed. The diversity of the study team is reflected by the authorship.

### Ethical approvals

Caregivers of all participants provided informed written consent. The study was prospectively registered on ClinicalTrials.gov (NCT04285021) and received ethical approval from the sponsors (MSF Ethical Review Board, MSF ERB 1967, and Oxford Tropical Medicine Research Committee, OxtREC 59-19) and ethical review boards in all participating countries (International Centre for Diarrhoeal Disease Research, Bangladesh, PR-200006; Angkor Hospital for Children Research Committee, Cambodia, 01296/AHC; National Ethics Committee for Health Research, Cambodia, 264/NECHR; Medical and Health Research Ethics Committee, Indonesia, KE/FK/1397/EC/2019; National Ethics Committee for Health Research, Laos, 051/NECHR; University of Medicine and Pharmacy at Ho Chi Minh City, Vietnam, 818/HDDD-DHYD; and Ethics Committee for Biomedical Research, Vietnam, VNCH-RICH-2021-77).

### Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

## Data availability

Deidentified, individual participant data from this study will be available to researchers whose proposed purpose of use is aligned with the study objectives and approved by the data access committees at MSF and MORU. Inquiries or requests for data can be sent to [data.sharing@london.msf.org](mailto:data.sharing@london.msf.org) and [dasharing@tropmedres.ac](mailto:dasharing@tropmedres.ac), with an anticipated timeline of 1 month between submission of a request and committee decision. Researchers interested in accessing biobanked samples should contact the corresponding author, who will coordinate with the Spot Sepsis Sample Use Committee, with an anticipated timeline of 1 month between submission of a request and committee decision.

## Code availability

Researchers interested in accessing the code should contact the corresponding author, A.C. ([arjun@tropmedres.ac](mailto:arjun@tropmedres.ac)), who will coordinate with the study statisticians, with an anticipated timeline of 1 month between submission of a request and response.

## References

58. Chandna, A. et al. Prediction of disease severity in young children presenting with acute febrile illness in resource-limited settings: a protocol for a prospective observational study. *BMJ Open* **11**, e045826 (2021).
59. Suspected sepsis: recognition, diagnosis and early management. NICE Guideline, No. 51. <https://www.ncbi.nlm.nih.gov/books/NBK553314/> (National Institute for Health and Care Excellence, 2024).
60. Scott, H. F., Donoghue, A. J., Gaieski, D. F., Marchese, R. F. & Mistry, R. D. The utility of early lactate testing in undifferentiated pediatric systemic inflammatory response syndrome. *Acad. Emerg. Med.* **19**, 1276–1280 (2012).
61. Fung, J. S. T. et al. Determining predictors of sepsis at triage among children under 5 years of age in resource-limited settings: a modified Delphi process. *PLoS ONE* **14**, e0211274 (2019).
62. Mawji, A. et al. Common data elements for predictors of pediatric sepsis: a framework to standardize data collection. *PLoS ONE* **16**, e0253051 (2021).
63. Parikh, S. M. Dysregulation of the angiotensin–Tie-2 axis in sepsis and ARDS. *Virulence* **4**, 517–524 (2013).
64. Mankambo, L. A. et al. The role of angiogenic factors in predicting clinical outcome in severe bacterial infection in Malawian children. *Crit. Care* **14**, R91 (2010).
65. Gutbier, B. et al. Prognostic and pathogenic role of angiotensin-1 and -2 in pneumonia. *Am. J. Respir. Crit. Care Med.* **198**, 220–231 (2018).
66. Zhang, R. et al. Dysregulation of angiotensin–Tie-2 axis in Ugandan children hospitalized with pneumonia. *Cytokine* **133**, 155175 (2020).
67. Conroy, A. L. et al. Host biomarkers are associated with response to therapy and long-term mortality in pediatric severe malaria. *Open Forum Infect. Dis.* **3**, ofw134 (2016).
68. Richard-Greenblatt, M. et al. Prognostic accuracy of soluble triggering receptor expressed on myeloid cells (sTREM-1)-based algorithms in febrile adults presenting to Tanzanian outpatient clinics. *Clin. Infect. Dis.* **70**, 1304–1312 (2020).
69. Florin, T. A. et al. Biomarkers and disease severity in children with community-acquired pneumonia. *Pediatrics* **145**, e20193728 (2020).
70. Carrol, E. D., Thomson, A. P. J. & Mobbs, K. J. The role of RANTES in meningococcal disease. *J. Infect. Dis.* **182**, 363–366 (2000).
71. Fuchs, A. et al. Cytokine kinetic profiles in children with acute lower respiratory tract infection: a post hoc descriptive analysis from a randomized control trial. *Clin. Microbiol. Infect.* **24**, 1341.e1–1341.e7 (2018).
72. Zhao, Y. et al. Longitudinal COVID-19 profiling associates IL-1RA and IL-10 with disease severity and RANTES with mild disease. *JCI Insight* **5**, e139834 (2020).
73. Van Singer, M. et al. COVID-19 risk stratification algorithms based on sTREM-1 and IL-6 in emergency department. *J. Allergy Clin. Immunol.* **147**, 99–106 e104 (2021).
74. Pourakbari, B. et al. Evaluation of procalcitonin and neopterin level in serum of patients with acute bacterial infection. *Braz. J. Infect. Dis.* **14**, 252–255 (2010).
75. Balanza, N. et al. Soluble triggering receptor expressed on myeloid cells 1 (sTREM-1) predicts mortality in patients with febrile illness in southern Mozambique. *Commun. Med. (Lond.)* **5**, 310 (2025).
76. Chen, H. L., Hung, C. H., Tseng, H. I. & Yang, R. C. Soluble form of triggering receptor expressed on myeloid cells-1 (sTREM-1) as a diagnostic marker of serious bacterial infection in febrile infants less than three months of age. *Jpn. J. Infect. Dis.* **61**, 31–35 (2008).
77. Wright, S. W. et al. sTREM-1 predicts mortality in hospitalized patients with infection in a tropical, middle-income country. *BMC Med.* **18**, 159 (2020).
78. Stefanova, V. et al. suPAR as a prognostic marker of Ugandan children at risk of severe and fatal malaria. *Clin. Infect. Dis.* **76**, e1079–e1086 (2023).
79. Wrotek, A., Jackowska, T. & Pawlik, K. Soluble urokinase plasminogen activator receptor: an indicator of pneumonia severity in children. *Adv. Exp. Med. Biol.* **835**, 1–7 (2015).
80. Raggam, R. B. et al. Soluble urokinase plasminogen activator receptor predicts mortality in patients with systemic inflammatory response syndrome. *J. Intern. Med.* **276**, 651–658 (2014).
81. Donadello, K., Scolletta, S., Covajes, C. & Vincent, J. L. suPAR as a prognostic biomarker in sepsis. *BMC Med.* **10**, 2 (2012).
82. Khilnani, P. et al. Pediatric sepsis guidelines: summary for resource-limited countries. *Indian J. Crit. Care Med.* **14**, 41–52 (2010).
83. Sanchez-Pinto, L. N. et al. Development and validation of the Phoenix criteria for pediatric sepsis and septic shock. *JAMA* **331**, 675–686 (2024).
84. Schilling, W. H. K. et al. Evaluation of hydroxychloroquine or chloroquine for the prevention of COVID-19 (COPCOV): a double-blind, randomised, placebo-controlled trial. *PLoS Med.* **21**, e1004428 (2024).
85. Riley, R. D. et al. Calculating the sample size required for developing a clinical prediction model. *BMJ* **368**, m441 (2020).
86. Riley, R. D. et al. Minimum sample size for developing a multivariable prediction model: PART II - binary and time-to-event outcomes. *Stat. Med.* **38**, 1276–1296 (2019).
87. World Population Prospects 2024. <https://population.un.org/wpp/> (United Nations, 2024).
88. Chew, R., Greer, R. C., Tasak, N., Day, N. P. J. & Lubell, Y. Modelling the cost-effectiveness of pulse oximetry in primary care management of acute respiratory infection in rural northern Thailand. *Trop. Med. Int. Health* **27**, 881–890 (2022).
89. Calarco, S. et al. Analytical performance of 17 commercially available point-of-care tests for CRP to support patient management at lower levels of the health system. *PLoS ONE* **18**, e0267516 (2023).
90. WHO-CHOICE estimates of cost for inpatient and outpatient health service delivery. <https://www.who.int/publications/m/item/who-choice-estimates-of-cost-for-inpatient-and-outpatient-health-service-delivery> (World Health Organization, 2021).
91. Purba, A. K. R. et al. The burden and costs of sepsis and reimbursement of its treatment in a developing country: an observational study on focal infections in Indonesia. *Int. J. Infect. Dis.* **96**, 211–218 (2020).

92. Nanyonjo, A. et al. Estimating the cost of referral and willingness to pay for referral to higher-level health facilities: a case series study from an integrated community case management programme in Uganda. *BMC Health Serv. Res.* **15**, 347 (2015).

## Acknowledgements

We are extremely grateful to all the participants and their caregivers and to the laboratory and research staff at the participating sites and MORU laboratories in Bangkok, Thailand, for specimen processing, data management and study monitoring. We acknowledge the support of bioMérieux, which provided laboratory consumables and equipment for some of the respiratory specimen assays. We also acknowledge the valuable support of William Robertson (MSF), without whom the development of the study would not have been possible, and guidance of N. J. White (MORU) at key points during the study and R. Western (Health In Harmony) for assistance with the production of Fig. 1. The study was co-funded by MSF Spain (S.B.) and Wellcome (219644/Z/19/Z; Y.L.). MSF maintained a sponsor/investigator role for the study. Wellcome had no role in study design, data collection, data analysis, data interpretation, writing of the report or decision to submit for publication.

## Author contributions

A.C., C.K., R.A.A., P.T., M.M., E.A.A., E.A., R.P.-S., M.R.-G., Y.L. and S.B. conceptualized the study. A.C., D.T.V.A., K.S.C., S.K., P.N.T.N., S.R., S.V., P.H.P., D.M., B.T.L. and E.A. acquired the data. A.C., R.M., M.Y., N.W. and M.R.-G. curated the data. M.Y.A., J.T. and M.R.-G. performed the

laboratory assays. A.C., C.K., R.M. and R.P.-S. did the formal analysis. Y.L. and S.B. acquired funding. A.C. wrote the original draft of the manuscript. All authors reviewed and edited the manuscript. A.C., C.K., R.M., M.R.-G., Y.L. and S.B. verified the underlying data. A.C., C.K., Y.L. and S.B. had full access to all of the data in the study and had final responsibility for the decision to submit for publication.

## Competing interests

The authors declare no competing interests.

## Additional information

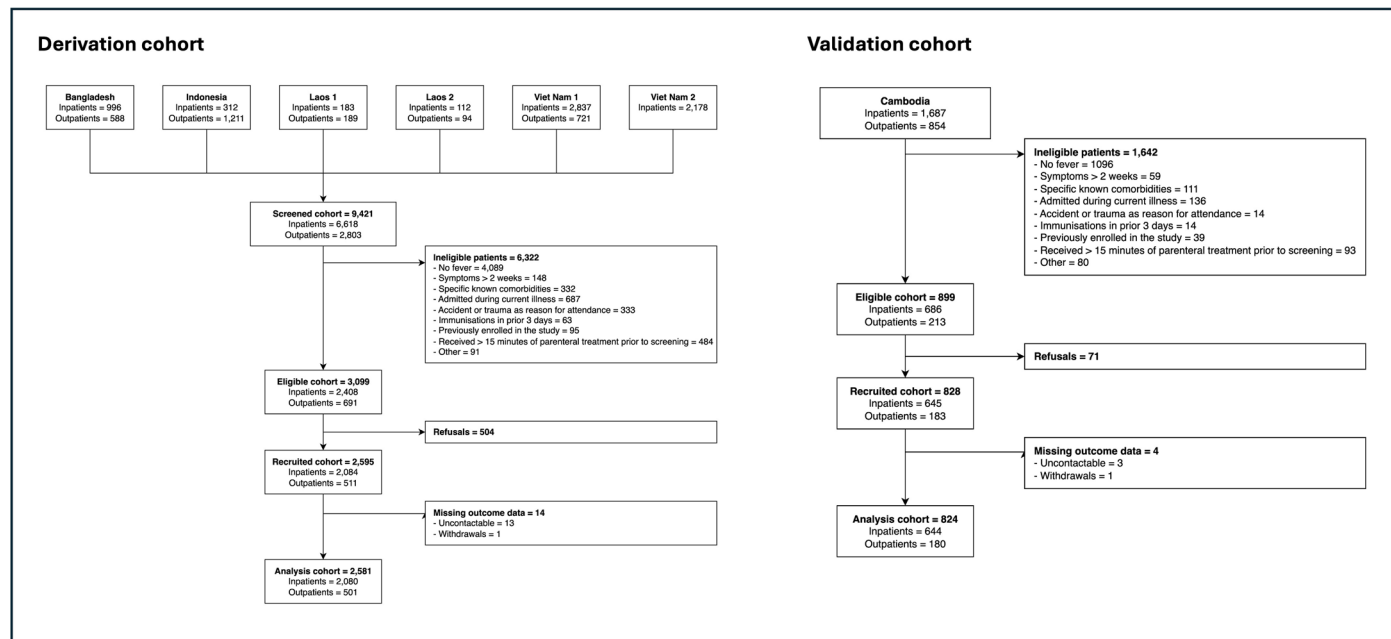
**Extended data** is available for this paper at <https://doi.org/10.1038/s41591-026-04338-1>.

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41591-026-04338-1>.

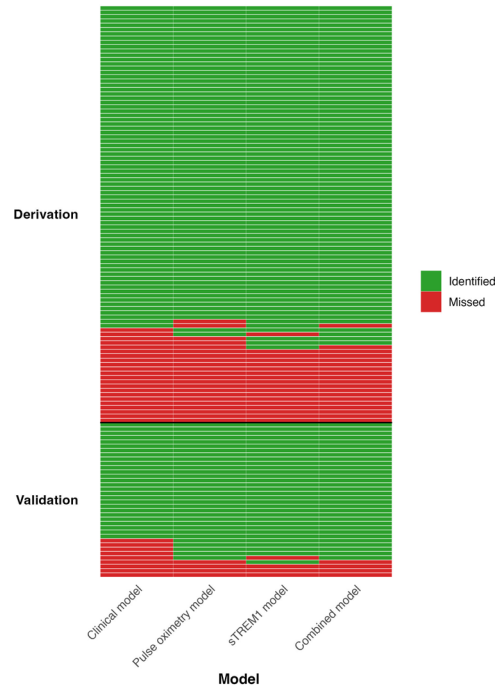
**Correspondence and requests for materials** should be addressed to Arjun Chandna.

**Peer review information** *Nature Medicine* thanks Mihir Atreya, Sumanth Gandra and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Ming Yang, in collaboration with the *Nature Medicine* team.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

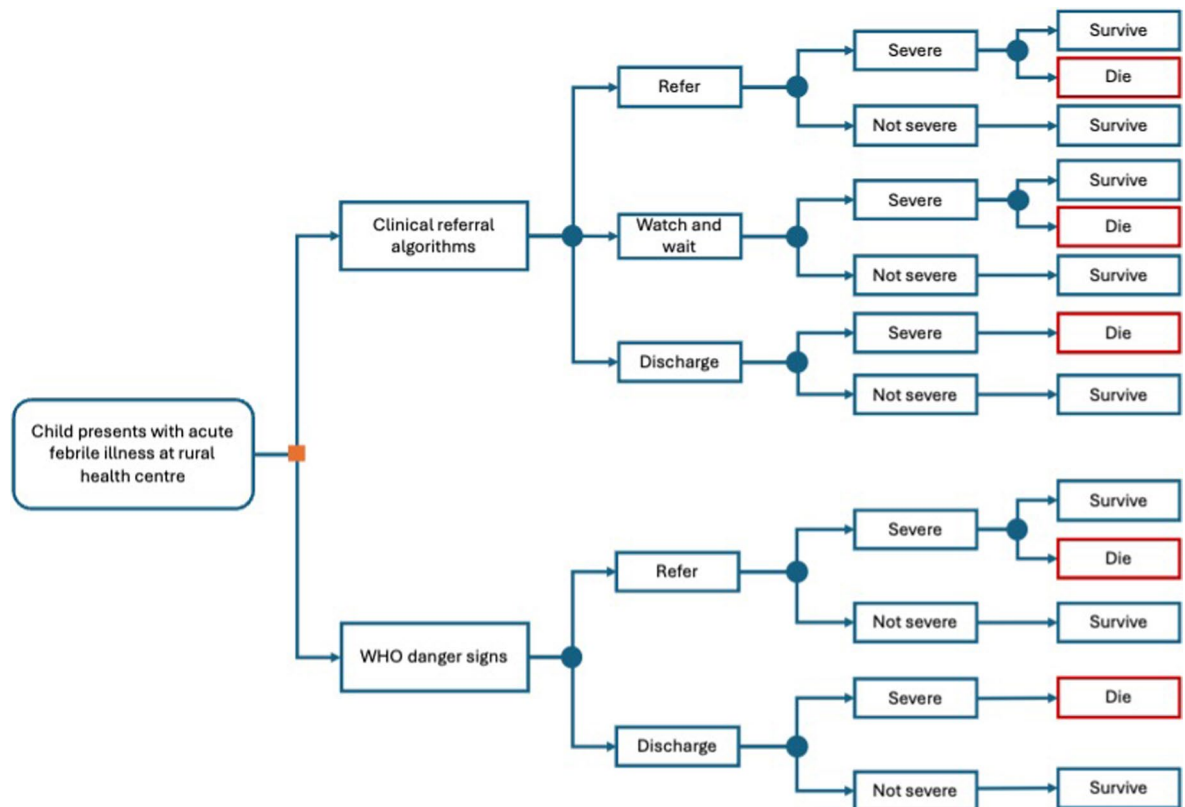


**Extended Data Fig. 1 | Study flowchart.** One reason for ineligibility is provided per patient, according to the hierarchy listed in the figure. 1,245 ineligible children had more than one reason for ineligibility (1,245/7,964; 15.6%).



**Extended Data Fig. 2 | Identification of participants who progressed to severe febrile illness for each of the models, split by derivation and validation cohort.** Classification for each of the 133 participants who progressed to develop severe febrile illness according to whether they were identified (green;  $pp > 0.5\%$ )

or missed (red;  $pp \leq 0.5\%$ ) by each of the models. Clinical model = clinical parameters only; pulse oximetry model = clinical parameters plus SpO<sub>2</sub>; sTREM1 model = clinical parameters plus sTREM1; combined model = clinical parameters plus SpO<sub>2</sub> and sTREM1.  $pp$  = predicted probability of severe febrile illness.



**Extended Data Fig. 3 | Cost-effectiveness analysis decision tree model structure.** Analytical framework to evaluate the comparative cost-effectiveness of the WHO danger signs vs. the clinical referral algorithms for triage of febrile children presenting from the community. Children triaged by the WHO danger signs can either be referred or discharged. Children triaged by the referral

algorithms can either be referred, monitored for a period, or discharged. All children who develop non-severe febrile illness are assumed to survive. All discharged children who develop severe febrile illness are assumed to die, whereas a proportion of referred or monitored children who develop severe febrile illness are assumed to survive.

Extended Data Table 1 | Comparison of the different triage approaches

		Clinical model (95% CI)	Pulse oximetry model (95% CI)	Difference or Ratio (95% CI)
Rule out	Sensitivity	74.7 (59.4 to 88.1)	88.9 (76.7 to 97.8)	13.7 (3.4 to 27.3)
	NLR	0.28 (0.13 to 0.45)	0.11 (0.02 to 0.25)	0.43 (0.16 to 0.78)
Rule in	Specificity	99.1 (97.7 to 99.7)	99.8 (99.7 to 99.9)	0.7 (0.11 to 2.13)
	PLR	37.5 (12.2 to 121.1)	150.7 (72.9 to 278.2)	3.9 (1.3 to 12.7)

		Clinical model (95% CI)	sTREM1 model (95% CI)	Difference or Ratio (95% CI)
Rule out	Sensitivity	74.7 (59.4 to 88.1)	89.2 (76.9 to 97.5)	13.9 (3.4 to 26.7)
	NLR	0.28 (0.13 to 0.45)	0.12 (0.03 to 0.25)	0.44 (0.15 to 0.81)
Rule in	Specificity	99.1 (97.7 to 99.7)	97.9 (97.7 to 99.7)	-1.1 (-2.8 to -0.03)
	PLR	37.5 (12.2 to 121.1)	18.2 (7.2 to 78.9)	0.50 (0.16 to 1.19)

		WHO danger signs (95% CI)	Clinical model (95% CI)	Difference or Ratio (95% CI)
Rule out	Sensitivity	55.5 (39.4 to 72.7)	74.7 (59.4 to 88.1)	19.0 (0 to 41.1)
	NLR	0.54 (0.32 to 0.74)	0.28 (0.13 to 0.45)	0.51 (0.22 to 0.93)
Rule in	Specificity	82.6 (77.1 to 87.6)	99.1 (97.7 to 99.7)	16.4 (11.7 to 21.7)
	PLR	3.2 (2.1 to 4.9)	37.5 (12.2 to 121.1)	12.5 (3.85 to 50.0)

For sensitivity and specificity, the absolute difference between the two models with 95% CI for the difference is reported. A CI that does not cross 0, indicates a statistically significant result. For the NLR and PLR, the ratio between the two models with 95% CI for the ratio is reported. A CI that does not cross 1, indicates a statistically significant result. CI = confidence interval; NLR = negative likelihood ratio; PLR = positive likelihood ratio.

## Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

### Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- |                                     |   |
|-------------------------------------|---|
| n/a                                 | Confirmed   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The exact sample size ( $n$ ) for each experimental group/condition, given as a discrete number and unit of measurement  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided<br><i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of all covariates tested   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons  |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For null hypothesis testing, the test statistic (e.g. $F$ , $t$ , $r$ ) with confidence intervals, effect sizes, degrees of freedom and $P$ value noted<br><i>Give <math>P</math> values as exact values whenever suitable.</i>                            |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes   |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's $d$ , Pearson's $r$ ), indicating how they were calculated   |

*Our web collection on [statistics for biologists](#) contains articles on many of the points above.*

### Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

### Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

De-identified, individual participant data from this study will be available to researchers whose proposed purpose of use is aligned with the study objectives and approved by the data access committees at MSF and MORU. Enquiries or requests for data can be sent to [data.sharing@london.msf.org](mailto:data.sharing@london.msf.org) and [datasharing@tropmedres.ac](mailto:datasharing@tropmedres.ac), with an anticipated timeline of one month between submission of a request and committee decision. Researchers interested in

accessing biobanked samples should contact the corresponding author, who will coordinate with the Spot Sepsis Sample Use Committee, with an anticipated timeline of one month between submission of a request and committee decision.

## Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender	Proportion of male and female participants recruited reported in Table 1
Reporting on race, ethnicity, or other socially relevant groupings	No reporting on race or ethnic groups performed as these vary considerably across and within different study locations
Population characteristics	3,423 children aged 1-59 months presenting with acute febrile illnesses. Screening occurred during daytime working hours. Inpatients were recruited consecutively. A random selection of outpatients were recruited, informed by computer-generated random number tables which utilised the previous week's routinely collected hospital data for the sampling frame. The main biases that may have occurred as a result of the recruitment process is that the study sample may not be representative of patients presenting out of hours and that the outpatient sample may be more representative of patients attending the hospital earlier in the working day.
Recruitment	5 March 2020 to 4 November 2022
Ethics oversight	The study received ethical approval from the sponsors (Médecins Sans Frontières Ethical Review Board MSF ERB 1967; Oxford Tropical Medicine Research Committee OxtREC 59-19) and ethical review boards in all participating countries (International Centre for Diarrhoeal Disease Research, Bangladesh PR-200006; Angkor Hospital for Children Research Committee, Cambodia 01296/AHC; National Ethics Committee for Health Research, Cambodia 264/NECHR; Medical and Health Research Ethics Committee, Indonesia KE/FK/1397/EC/2019; National Ethics Committee for Health Research, Laos 051/NECHR; University of Medicine and Pharmacy at Ho Chi Minh City, Viet Nam; 818/HDDD-DHYD; Ethics Committee for Biomedical Research, Viet Nam VNCH-RICH-2021-77).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

## Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences  Behavioural & social sciences  Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://nature.com/documents/nr-reporting-summary-flat.pdf)

## Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The methods of Riley et al. were followed and using an anticipated outcome prevalence of 1%, conservative R2 Nagelkerke of 0.15, and shrinkage factor of 0.9, it was estimated that six events per parameter would be required to derive the prediction models. The Covid-19 pandemic, declared in the same week that recruitment began at the first site, delayed initiation of other sites and slowed enrolment. Study duration (and thus sample size) was determined by available resources, with recruitment continuing for as long as funding allowed. Accordingly, as there were 97 events in the derivation dataset, a maximum of 16 candidate parameters could be included for development of the prediction models, whilst minimising the risk of overfitting.
Data exclusions	18 participants. Lost to follow up and so no data available available on outcome status.
Replication	Cross-validation during the development of the prediction models. External validation of the prediction models in a held-out dataset.
Randomization	Not applicable - observational study.
Blinding	Not applicable - observational study therefore no group allocation.

## Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

## Materials &amp; experimental systems

## Methods

- n/a | Involved in the study
- Antibodies
- Eukaryotic cell lines
- Palaeontology and archaeology
- Animals and other organisms
- Clinical data
- Dual use research of concern
- Plants

- n/a | Involved in the study
- ChIP-seq
- Flow cytometry
- MRI-based neuroimaging

## Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration	<input type="text" value="ClinicalTrials.gov NCT04285021"/>
Study protocol	<input type="text" value="https://osf.io/v594s/ and https://bmjopen.bmj.com/content/11/1/e045826"/>
Data collection	<input type="text" value="Goyalmara Mother and Child Hospital, Bangladesh: 18/03/2021 to 27/04/2022&lt;br/&gt;Angkor Hospital for Children, Cambodia: 05/03/2020 to 24/02/2022&lt;br/&gt;Rumah Sakit Umum Daerah Wates, Indonesia: 22/03/2021 to 22/04/2021&lt;br/&gt;Salavan Provincial Hospital, Lao PDR: 10/09/2020 to 30/08/2021&lt;br/&gt;Savannakhet Provincial Hospital, Lao PDR: 21/01/2021 to 26/08/2021&lt;br/&gt;Dong Nai Children's Hospital, Viet Nam: 10/05/2021 to 28/10/2022&lt;br/&gt;Viet Nam National Children's Hospital, Viet Nam: 08/12/2021 to 04/11/2022"/>
Outcomes	<input type="text" value="Primary outcome: death or receipt of organ support (mechanical ventilation, non-invasive ventilation, inotropic therapy, or renal replacement therapy) within two days of enrolment. Assessed by research staff via observation at the bedside and/or interview with caregiver.&lt;br/&gt;&lt;br/&gt;Secondary outcome: Category IV = Death or organ support ≤ 2 days after enrolment; Category III = Admission to any health facility for &gt; 2 nights between enrolment and D28 OR death or organ support between D2 and D28; Category II = Admission to any health facility for ≤ 2 nights between enrolment and D28 OR symptoms not resolved by D28; Category I = Not admitted to any health facility between enrolment and D28 AND recovered by D28. Assessed by research staff via observation at the bedside and/or interview with caregiver."/>

## Plants

Seed stocks	<input type="text" value="NA"/>
Novel plant genotypes	<input type="text" value="NA"/>
Authentication	<input type="text" value="NA"/>