BMJ Global Health

Predictors of disease severity in children presenting from the community with febrile illnesses: a systematic review of prognostic studies

Arjun Chandna ,[1,2] Rainer Tan,[3,4,5] Michael Carter,[6] Ann Van Den Bruel,[7] Jan Verbakel,[7,8] Constantinos Koshiaris,[8] Nahya Salim,[9,10] Yoel Lubell,[2,11] Paul Turner ,[1,2] Kristina Keitel[5,12]

To cite: Chandna A, Tan R, Carter M, *et al*. Predictors of disease severity in children presenting from the community with febrile illnesses: a systematic review of prognostic studies. *BMJ Global Health* 2021;**6**:e003451. doi:10.1136/bmjgh-2020-003451

Handling editor Soumyadeep Bhaumik

► Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/bmjgh-2020-003451).

AC and RT contributed equally. PT and KK contributed equally.

Received 15 July 2020
Revised 26 November 2020
Accepted 19 December 2020

For numbered affiliations see end of article.

Correspondence to
Dr Arjun Chandna;
arjun@tropmedres.ac;
arjunchandna@gmail.com

ABSTRACT

Introduction Early identification of children at risk of severe febrile illness can optimise referral, admission and treatment decisions, particularly in resource-limited settings. We aimed to identify prognostic clinical and laboratory factors that predict progression to severe disease in febrile children presenting from the community.

Methods We systematically reviewed publications retrieved from MEDLINE, Web of Science and Embase between 31 May 1999 and 30 April 2020, supplemented by hand search of reference lists and consultation with an expert Technical Advisory Panel. Studies evaluating prognostic factors or clinical prediction models in children presenting from the community with febrile illnesses were eligible. The primary outcome was any objective measure of disease severity ascertained within 30 days of enrolment. We calculated unadjusted likelihood ratios (LRs) for comparison of prognostic factors, and compared clinical prediction models using the area under the receiver operating characteristic curves (AUROCs). Risk of bias and applicability of studies were assessed using the Prediction Model Risk of Bias Assessment Tool and the Quality In Prognosis Studies tool.

Results Of 5949 articles identified, 18 studies evaluating 200 prognostic factors and 25 clinical prediction models in 24 530 children were included. Heterogeneity between studies precluded formal meta-analysis. Malnutrition (positive LR range 1.56–11.13), hypoxia (2.10–8.11), altered consciousness (1.24–14.02), and markers of acidosis (1.36–7.71) and poor peripheral perfusion (1.78–17.38) were the most common predictors of severe disease. Clinical prediction model performance varied widely (AUROC range 0.49–0.97). Concerns regarding applicability were identified and most studies were at high risk of bias.

Conclusions Few studies address this important public health question. We identified prognostic factors from a wide range of geographic contexts that can help clinicians assess febrile children at risk of progressing to severe disease. Multicentre studies that include outpatients are required to explore generalisability and develop data-driven tools to support patient prioritisation and triage at the community level.

PROSPERO registration number CRD42019140542.

Key questions

What is already known?

- An increasing number of clinical decision-support algorithms and risk stratification tools integrate clinical and laboratory predictors to guide healthcare workers in their assessment of febrile children.
- Which prognostic factors—alone or as components of clinical prediction models—best identify children at risk of developing severe febrile illness is not clear.
- Previous systematic reviews have focused on diagnostic studies and used imperfect reference standards for severe disease.

What are the new findings?

- Malnutrition, hypoxia, altered consciousness, and bedside markers of acidosis and poor peripheral perfusion were the most commonly identified predictors of severe disease.
- Clinical prediction model performance varied—the best performing models being those evaluated in similar settings and using similar outcomes as the original derivation studies.
- The prognostic factors and clinical prediction models identified in this study reflect children with relatively advanced illnesses and hence the degree to which they can inform community triage and prioritisation strategies is unclear.

INTRODUCTION

Acute febrile illnesses are among the most common reasons that parents seek medical care for their children.[1 2] While most episodes are mild, an important minority of children progress to severe disease. Early recognition of low-incidence serious disease is challenging,[3] especially in many tropical settings where health workers receive limited training, patient volumes are high, diagnostic capacity is poor and different acute febrile syndromes are often clinically indistinguishable.[4 5]

Key questions

What do the new findings imply?

- The studies included in this systematic review, together with other studies, highlight the importance of not over interpreting prognostic performance of individual predictors, which vary across different epidemiological contexts.
- If prediction models and decision-support algorithms are to be used as an adjunct to clinical assessment, they must be derived and validated using populations and outcomes appropriate to the clinical problem.
- To improve identification of children at risk of developing severe febrile illness, this will require multiple, large, collaborative research initiatives, which collect harmonised yet contextualised data on predictors and outcomes, and include unselected children presenting from the community.

Clinical and laboratory prognostic factors that enable early and accurate identification of children at risk of developing severe disease could improve patient outcomes and reduce resource misallocation.[6 7] An increasing number of clinical decision-support algorithms and risk stratification tools integrate clinical and laboratory predictors to guide referral, admission and treatment decisions.[8] While no unified strategy exists to guide selection of candidate predictors, those already reported as prognostic should normally be considered.[9]

Previous reviews have evaluated predictors of 'serious bacterial infections'.[10 11] However, these studies are diagnostic rather than prognostic.[9] Furthermore, 'serious bacterial infection' is an imperfect measure of disease severity: microbiological tests for bacterial infections lack sensitivity, especially in settings with high antibiotic consumption; 'serious bacterial infections' are not always severe (eg, children with enteric fever are often successfully managed as outpatients) and severe febrile illnesses are frequently caused by non-bacterial pathogens, especially in low/middle-income countries (LMICs),[4 12] in part secondary to the introduction of widespread vaccination against prevalent bacterial pathogens of childhood.[13]

We performed a systematic review to identify which clinical and laboratory factors—alone or as part of clinical prediction models—predict progression to severe disease in febrile children presenting from the community to a community health worker, primary health centre or hospital outpatient or emergency department. Our aim was to understand which prognostic factors might support health workers faced with this difficult and common clinical question and to inform variable selection for future prospective studies aiming to develop data-driven triage tools.

METHODS

Protocol and registration

The methods for this systematic review were specified in advance (PROSPERO protocol: CRD42019140542) and adhere to the Checklist for Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS),[14] a modification of CHARMS for prognostic factor studies (CHARMS-PF),[15] Quality In Prognosis Studies (QUIPS)[16] and Prediction Model Risk of Bias Assessment Tool (PROBAST) guidelines.[17] The report has been prepared in accordance with Preferred Reporting Items for Systematic Review and Meta-Analysis guidelines.[18]

Eligibility criteria

All prognostic studies (prognostic factor and clinical prediction model) including ≥20 patients were eligible. Our target population was children aged >28 days and <19 years, presenting from the community with an acute febrile illness (documented abnormal temperature (fever or hypothermia) or history of fever) or suspected sepsis. While sepsis is not always well defined in children,[19] 'suspected sepsis' was included along with febrile children so as to include all children with suspected infection. Studies were excluded if disaggregated paediatric data were not presented or patients were recruited partway through receipt of inpatient treatment, as the aim of the review was to identify *prognostic* variables measured at presentation. Studies that only evaluated specific clinical syndromes (eg, neurological presentations, acute respiratory infections and so on) or particular pathogens (eg, *Plasmodium* spp, influenza and so on) were not included.

Studies measuring predictors at presentation to care were included. Studies where authors identified that a substantial proportion of participants were recruited following transfer from another health facility were excluded. Demographic, anthropometric, socioeconomic, clinical and historical variables were considered, as well as laboratory parameters measured at presentation to care. Studies only reporting variables that would not be available at the time of presentation to care (eg, blood culture results) were excluded.

The primary outcome was any objective measure of disease severity occurring within 30 days of measurement of the predictors or during hospitalisation. Studies assessing outcome at the same time point as baseline predictor measurements (diagnostic studies) were excluded.

Search strategy and selection criteria

We searched MEDLINE, Embase and Web of Science databases, without language restriction, for publications between 31 May 1999 and 30 April 2020 (initial search to 31 May 2019; updated search to 30 April 2020). We followed Cochrane Prognosis Methods Group recommendations to build our search strategy (online supplemental appendix S1), structured according to the 'populations, interventions, comparators, outcomes, timing and setting' (PICOTS) framework and adapted published search strings as appropriate.[20–22] The search

strategy was peer-reviewed by an independent Technical Advisory Panel (online supplemental appendix S2).

Study selection

Title, abstract and full-text screening were performed independently by two reviewers (AC and RT). Agreement was checked after the first 20 and 250 articles. Discrepancies were resolved by discussion or independent assessment by a third reviewer (KK).

Eligible studies and relevant review articles were 'snowballed' (forward and reverse crosschecking of reference lists) to identify additional studies. The list of eligible studies was presented to the Technical Advisory Panel who were asked to identify obvious omissions and suggest key authors whose publication lists were subsequently reviewed for additional eligible studies (online supplemental appendix S2).

Data collection process

Data extraction sheets were developed based on the CHARMS and CHARMS-PF checklists (online supplemental appendix S3).[14 15] Data were extracted independently by one reviewer (AC or RT) and checked by the other. Discrepancies were discussed and resolved between the two reviewers. Authors of studies not reporting likelihood ratios (LRs) (prognostic factors) or area under the receiver operating characteristic curves (AUROCs) (clinical prediction models), or the data to allow their calculation, were contacted. Seven authors responded to requests for clarifications and six provided additional data not available in the published manuscript. All predictors were harmonised using the Systematised Nomenclature of Medicine Clinical Terms (SNOMED-CT).

Data analysis: prognostic factors

Contingency tables were constructed and positive likelihood ratio (PLR) and negative likelihood ratio (NLR) calculated for each prognostic factor. In the case of an empty cell, 0.5 was added to each cell (Haldane-Anscombe correction). CIs were calculated on the basis of the SE of a proportion (Stata V.16.0). LRs were selected as the principal effect estimate as they allow estimation of post-test probabilities, are independent of prevalence, are intuitive for clinicians and are frequently used to compare performance of predictors in diagnostic and prognostic studies.[10 11 23 24] Prognostic factors are presented in the main analysis if at least one study reported a PLR ≥5.0 (ie, a rule-in test), or a NLR ≤0.2 (ie, a rule-out test).[23] To contextualise the results, we used the outcome prevalence of individual studies to calculate the pre-test probability, and display positive and negative post-test probabilities on dumbbell plots (R V.3.6.1).

Data analysis: clinical prediction models

For clinical prediction models, AUROCs are presented on forest plots (Stata V.16.0). When available, we present LRs for different thresholds of the models in online supplemental appendix S4.

Synthesis of results

Due to expected heterogeneity between studies (as a result of variations in case-mix and baseline risk), few common predictors for comparison and absence of well-defined subgroups, no formal meta-analysis nor comparison of variability and bias between studies was planned, as these comparisons are recognised as being prone to bias.[25] Qualitative comparisons are described considering major differences between populations and study design. Prevalence of severe disease was used to group studies into low (<2.5%), moderate (2.5%–7.5%) and high (>7.5%) prevalence settings, as a proxy for the case-mix and level of care.

Quality assessment

Risk of bias and applicability of studies were assessed using the QUIPS tool for prognostic factor studies,[16] and PROBAST for studies developing, validating or updating prediction models.[17] Each study was independently assessed using QUIPS or PROBAST by two reviewers (AC and RT), as well as an independent senior reviewer (MC, AVDB or JV). All discrepancies were resolved by discussion. For prognostic factor studies (QUIPS), risk of bias was categorised as low, medium or high, while in clinical prediction model studies (PROBAST) risk was categorised as low, high or unclear. For all studies, applicability was assessed as being of high, low or unclear concern.

Role of the funding source

The funders had no role in study design, data collection, data analysis, data interpretation or writing of the report. The co-primary authors (AC and RT) had full access to the data and final responsibility for the decision to submit for publication.

Patient and public involvement

Neither patients nor members of the public were directly involved in the conduct of this work.

RESULTS

The electronic search retrieved 5930 articles, and 19 additional articles were identified through snowballing and expert consultation (figure 1). Eighteen studies were included in the review: 16 studies evaluated 200 prognostic factors, from 75 SNOMED-CT categories,[12 26–38] and eight evaluated 33 clinical prediction model/outcome pairs, using 25 distinct models.[27 29 31 38–42]

In total 24 530 children were included, with overlap across eight studies.[26 31 32 34–37 40] The majority (11/18) included only hospitalised patients. Two studies recruited children from primary care,[29 33] and five recruited both children admitted and those sent home directly from hospital outpatient or emergency departments.[28 35–37 40] Seven studies included children aged 5 years and under,[27 30 32–34 39 42] with the remainder including patients up to 19 years of age. Definition of fever varied between studies, ranging from an axillary temperature (or equivalent) of ≥37.5°C to >38.1°C. Five

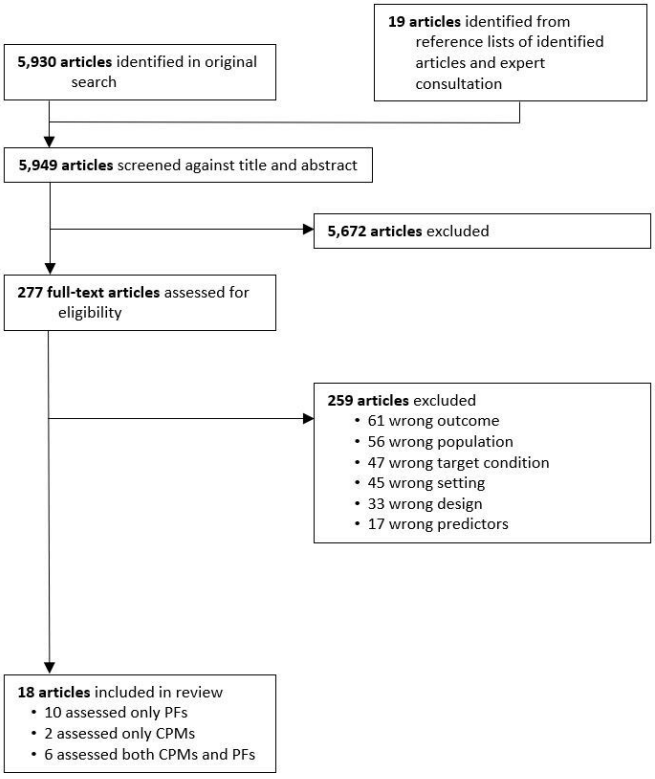

Figure 1 Selection of studies. Only one reason for exclusion per study is listed. CPM, clinical prediction model; PF, prognostic factor.

studies did not include a temperature measurement in their eligibility criteria and enrolled children on the basis of suspected infection or sepsis.[35 38 40–42] Eight studies were conducted in sub-Saharan Africa,[26 27 31–34 41 42] four in North America,[35–37 40] three in Europe,[29 30 38] two in Asia[12 39] and one in Latin America.[28] Six were multicentre studies.[12 26 31 33 40 42] Most used 'hard' outcomes to define severe disease, such as mortality, organ dysfunction or need for organ support, while four used 'softer' outcomes, such as prolonged length of stay or persistence of symptoms.[29 30 33 38] Characteristics of the 18 studies are summarised in table 1.

Prognostic factors

Figures 2–4 present prognostic factors identified as having rule-in (PLR ≥5.0) or rule-out (NLR ≤0.2) value in at least one study. Prognostic factors that met neither of these pre-specified cut-offs are presented in online supplemental appendix S5. In settings with moderate prevalence of severe disease, both high lactate (PLR range 4.97–5.13) and hypoglycaemia (PLR range 12.63–13.36) were useful for ruling in severe disease,[32 34 37] whereas a lactate ≤5 mM was more useful as a rule-out test (NLR 0.13) among a population in whom prevalence of severe disease was high (febrile children with signs of poor organ perfusion).[26]

Hypoxia was most useful to rule-in severe disease in moderate prevalence settings (PLR range 8.11–9.49).[27 34] Some studies found hypotension and bedside markers of poor peripheral perfusion to have useful rule-in value, but this was inconsistent (PLR range 1.89–9.57 and 1.78–17.38, respectively).[26 27 31 32 34–36 38] Bradycardia was evaluated in a multicentre study conducted across three East African countries and found to have useful rule-in value (PLR range 5.95–14.59) for severe disease in those high prevalence settings.[26 31] Impaired consciousness, assessed using bedside coma scales, was a useful predictor of severe disease, particularly in low and moderate prevalence settings (PLR range 3.38–14.02), with the post-test probability of poor outcome increasing with the degree of neurological impairment.[27 32 34 36 38 41 42]

In sub-Saharan African settings, severe malnutrition (PLR range 1.56–11.23),[26 27 32 34 41] HIV positive status (PLR range 2.32–12.48)[26 27 41 42] and bedside correlates of metabolic derangement such as deep breathing and jaundice (PLR range 3.57–7.71) were useful rule-in predictors, across a range of prevalence settings.[27 32 34]

Very few prognostic factors were satisfactorily able to rule-out progression to severe disease: presence of comorbidities (NLR range 0.12–1.04), sepsis at admission (NLR 0.19) and prostration (NLR range 0.18–1.23) were each identified in only one study.[27 28 35]

Clinical prediction models

Figure 5 illustrates the discrimination (AUROC) of 25 clinical prediction models for 33 different outcomes assessed in eight studies: most (18/33) were external validations of existing models[27 31 38 39]; 13 were newly derived models[29 31 40–42] and two were updates and external validations of an existing model.[38] Components of the clinical prediction models are summarised in table 2.

Three models, Lambaréné Organ Dysfunction Score (LODS), Paediatric Early Death Index for Africa (early death score) (PEDIA-e) and Signs of Inflammation in Children that Kill (SICK), showed good (AUROC ≥0.80) discrimination in a Ugandan setting where in-hospital mortality occurred at a prevalence of 4.7% (AUROC range 0.85–0.90).[27] Two of these (LODS and PEDIA-e) were also assessed in a multicentre study in East Africa where discrimination was lower (AUROCs of 0.77 and 0.70).[31] This study also derived two models, the FEAST-Paediatric Emergency Triage (FEAST-PET) and FEAST-Paediatric Emergency Triage and Laboratory (FEAST-PETaL) scores, which showed good discrimination (AUROCs of 0.86 and 0.82).[31] Two other East African studies used combinations of simple clinico-demographic variables to derive a number of prediction models, four of which had AUROCs ≥0.80.[41 42]

One North American study derived a model to predict hypotensive shock in unselected children presenting with suspected sepsis, which showed good discrimination in an external geographic validation (AUROC 0.87).[40] The Yale Observation Score also showed high discrimination for mortality (AUROC 0.97) and mechanical ventilation (AUROC 0.89) in India, however, the small sample size (n=100) renders the results difficult to interpret.[39] In general, models assessed against 'softer' outcomes (eg,

Table 1 Characteristics of included studies

Study (year); setting, country	Cohort	Design	Quality assessment		CPM or PF study*	Sample size	Population		Outcome	Outcome prevalence (n/N)
			Risk of bias	Applicability concern			Inclusion criteria	Exclusion criteria		
Outcomes including death, organ dysfunction, organ support and PICU admission										
Scott (2020)[40]; Secondary and tertiary care hospitals, USA	Hospital OPD/ED	Retrospective cohort	High	Low	CPM	2464	Age 60 days to 18 years; Clinician-suspected sepsis	Hypotensive septic shock on arrival;† transfer to another centre; leaving ED before formal evaluation; incorrect registration	Hypotensive septic shock‡≤24 hour	11.4% (282/2464)
Walia (2016)[39]; Tertiary care hospital, India	Hospitalised§	Prospective cohort	High	High	CPM	100	Age 3–36 months; Axillary temperature >36.9°C (early morning) or >37.4°C	Non-infectious cause of fever; immunisation ≤2 days; immunodeficiency, autoimmune disorder	In-hospital mortality; Mechanical ventilation	11.0% (11/100); 17.0% (17/100)
Aramburo (2018)[26]; Secondary and tertiary care hospitals, Kenya, Tanzania and Uganda	Hospitalised§	Randomised controlled trial	Moderate	High	PF	3008	Age 60 days to 12 years; history of fever or axillary temperature ≥37.5°C or <36°C; severe febrile illness¶	Non-infectious cause of illness; SAM, gastroenteritis, burns, chronic kidney disease, pulmonary oedema, intoxication, surgical conditions, receipt of isotonic fluids during the same illness	In-hospital mortality (72 hours)	10.3% (309/3008)
George (2015)[31]; Secondary and tertiary care hospitals, Kenya, Tanzania and Uganda	Hospitalised§	Randomised controlled trial	High	High	CPM	3121	Age 60 days to 12 years; history of fever or axillary temperature ≥37.5°C or <36°C; severe febrile illness¶	Non-infectious cause of illness; SAM, gastroenteritis, burns, chronic kidney disease, pulmonary oedema, intoxication, surgical conditions, receipt of isotonic fluids during the same illness	In-hospital mortality (48 hours)	9.8% (306/3121)
Scott (2012)[37]; Tertiary care hospital, USA	Hospital OPD/ED	Prospective cohort	High	High	PF	239	Age <19 years; temperature >38.5°C or <36°C and heart rate >2 SD above normal for age; underwent phlebotomy as part of usual care	Transfer from another health facility; known inborn errors of metabolism; receipt of >15 min of intravenous therapy	24 hours organ dysfunction	5.4% (13/239)

Continued

Table 1 Continued

Study (year); setting, country	Cohort	Design	Quality assessment		CPM or PF study*	Sample size	Population		Outcome	Outcome prevalence (n/N)
			Risk of bias	Applicability concern			Inclusion criteria	Exclusion criteria		
Scott (2014)[36]; Tertiary care hospital, USA	Hospital OPD/ED	Prospective cohort	High	High	PF	239	Age <19 years; temperature >38.5°C or <36°C and heart rate >2 SD above normal for age; undergoing phlebotomy as part of routine care	Transfer from another health facility; known inborn errors of metabolism; receipt of >15 min of intravenous therapy	24 hours organ dysfunction	5.4% (13/239)
Nadjm (2013)[34]; Secondary care hospital, Tanzania	Hospitalised§	Prospective cohort	Moderate	High	PF	3319	Age 2 months to 5 years; history of fever in last 48 hours or axillary temperature ≥37.5°C	Chronic illness (excluding HIV and malnutrition); trauma; surgical conditions	In-hospital mortality	5.1% (170/3319)
Mtove (2011)[32]; Secondary care hospital, Tanzania	Hospitalised§	Prospective cohort	Moderate	High	PF	3248	Age 2 months to 13 years; history of fever in last 48 hours or axillary temperature ≥37.5°C	Chronic illness (excluding HIV and malnutrition); trauma; surgical conditions	In-hospital mortality	5.0% (164/3248)
Lowlaavar (2016)[42]; Secondary and tertiary care hospitals, Uganda	Hospitalised§	Prospective cohort	High	High	CPM	1307	Age 6–60 months; admitted during study working hours or within 8 hours of study shift with a proven or suspected infection	Previous enrolment; residence outside study catchment area	In-hospital mortality	5.0% (65/1307)
Conroy (2015)[27]; Tertiary care hospital, Uganda	Hospitalised§	Prospective cohort	High	High	CPM	2502	Age 2 months to 5 years; history of fever in last 48 hours or axillary temperature >37.5°C	None reported	In-hospital mortality	4.7% (99/2089)
van Nassau (2018)[38]; Secondary care hospital, The Netherlands	Hospitalised§	Retrospective cohort	High	High	CPM	864	Age <18 years; suspected bacterial infection**	Surgical conditions	PICU transfer and/or in-hospital mortality	2.7% (24/864)

Continued

Table 1 Continued

Study (year); setting, country	Cohort	Design	Quality assessment		CPM or PF study*	Sample size	Population		Outcome	Outcome prevalence (n/N)
			Risk of bias	Applicability concern			Inclusion criteria	Exclusion criteria		
Scott (2017)[35]; Tertiary care hospital, USA	Hospital OPD/ED	Retrospective cohort	Low	High	PF	1299	Age 60 days to 18 years; suspected sepsis††; measurement of venous lactate as part of routine care within 8 hours of ED arrival	Transfer from another health facility	30-day mortality	1.9% (25/1299)
SEAIDCRN (2017)[12]; Tertiary care hospitals, Indonesia, Thailand and Vietnam	Hospitalised§	Prospective cohort	High	High	PF	763	Age 30 days to 18 years; modified SIRS criteria‡‡	Suspicion of hospital-acquired infection; admission to hospital within previous 30 days; transfer from another health facility after >72 hours admission; weight <3 kg; enrolment in another clinical study	28-day mortality	1.9% (14/731)
Costa de Santana (2017)[28]; Tertiary care hospital, Brazil	Hospital OPD/ED	Retrospective cohort	High	High	PF	254	Age <13 years; axillary temperature >38.5°C; measurement of respiratory rate and heart rate on three occasions in absence of fever; measurement of leucocyte count as part of routine care	Congenital malformations; bronchopulmonary dysplasia; medullary aplasia; cardiac, renal or hepatic insufficiency	In-hospital mortality	1.6% (4/254)
Kwizera (2019)[41]; Secondary care hospital, Rwanda	Hospitalised§	Prospective cohort	High	High	CPM	949	Age 28 days to 18 years; confirmed acute infectious disease; symptom onset <14 days prior to hospital admission	Allergy to antimicrobials to treat sepsis (antibiotics, artesunate, artemether-lumefantrine); terminal disease	In-hospital mortality	1.5% (14/949)
Outcomes including length of stay and persistence of symptoms										
Freyne (2013)[30]; Secondary care hospital, Ireland	Hospitalised§	Prospective cohort	High	High	PF	46	Age 6–36 months; axillary temperature >38.1°C	Chronic illness; immunisation ≤2 days, antipyretic use ≤2 hours	Length of stay >96 hours	26.1% (12/46)

Continued

Table 1 Continued

Study (year); setting, country	Cohort	Design	Quality assessment		CPM or PF study*	Sample size	Population		Outcome	Outcome prevalence (n/N)
			Risk of bias	Applicability concern			Inclusion criteria	Exclusion criteria		
van Nassau (2018)[38]; Secondary care hospital, The Netherlands	Hospitalised§	Retrospective cohort	High	High	CPM	864	Age <18 years; suspected bacterial infection**	Surgical conditions	Length of stay ≥7 days	22.2% (179/806)
Elshout (2015)[29]; General Practice (out of hours), The Netherlands	Primary care	Prospective cohort	High	High	PF	480	Age 3 months to 6 years; history of fever	Communication in Dutch not possible; enrolment in last 2 weeks; direct referral to hospital required	Persistent fever at D3	13.1% (63/480)
Mwandama (2016)[33]; Community Health Workers, Malawi	Primary care	Prospective cohort	High	High	PF	285	Age 2–59 months; history of fever in last 48 hours or temperature ≥37.5°C; negative malaria rapid diagnostic test	Receipt of antimalarial in last 2 weeks; presence of danger signs§§	Persistent symptoms at D7	10.4% (19/182)

Studies are grouped according to the type of outcome they used: 'hard' (death, organ dysfunction, organ support, PICU admission) or 'soft' (length of stay, persistence of symptoms).
*Studies evaluating both PFs and CPMs were categorised on the basis of their primary analysis to facilitate review using the appropriate quality assessment tool.
†Hypotensive systolic blood pressure on arrival with receipt of a fluid bolus or vasoactive agent within 30 min.
‡Hypotension plus receipt of ≥30 mL/kg isotonic crystalloids or vasoactive medication.
§Only children the treating physician decided to admit were eligible but recruitment occurred at the time of admission to the health facility.
¶Respiratory distress (increased work of breathing or deep breathing) and/or impaired consciousness (coma or prostration) AND evidence of poor peripheral perfusion (capillary refill time >2 s or lower limb temperature gradient or weak radial pulse or severe tachycardia).
**Initiation of antibiotics within 24 hours of arrival in the emergency department.
††Decreased mental status or perfusion in the setting of suspected infection.
‡‡Rectal temperature >38.5°C or <35°C (or equivalent) AND heart rate >2 SD above normal for age (unless hypothermic) AND respiratory rate >2 SD above normal for age AND altered mental status OR systolic blood pressure <2 SD below normal for age OR pulse pressure <20 mm Hg OR capillary refill time >2 s OR SpO_2 <95% OR leucocyte count >12×10^3 cells/µL or <5×10^3 cells/µL.
§§Convulsions, repeated vomiting, lethargy, severe anaemia or loss of consciousness.
CPM, clinical prediction model; ED, emergency department; n, number of outcomes; n, number of cases; OPD, outpatient department; PF, prognostic factor; PICU, paediatric intensive care unit; SAM, severe acute malnutrition; SEAIDCRN, Southeast Asia Infectious Disease Clinical Research Network; SIRS, systemic inflammatory response syndrome.;

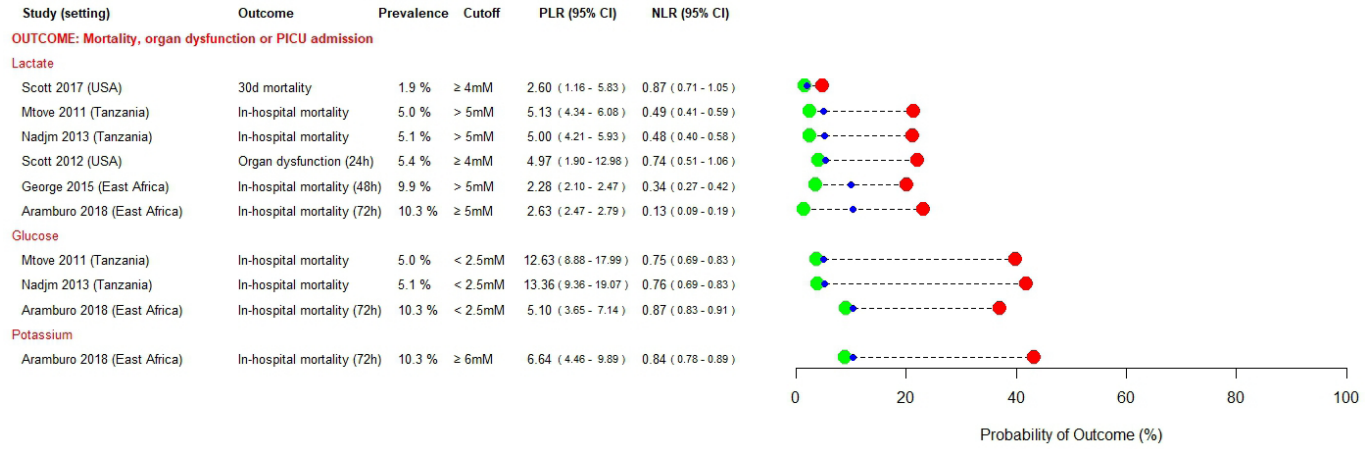

Figure 2 Prognostic factors identified as having rule-in (PLR ≥5.0) or rule-out (NLR ≤0.2) value for severe disease in at least one study—laboratory tests. mM, millimolar; NLR, negative likelihood ratio; PICU, paediatric intensive care unit; PLR, positive likelihood ratio.

persistence of symptoms or length of stay) had poorer discrimination, and a more distal temporal relationship between measurement of predictors and ascertainment of outcome.

Quality assessment

Only one prognostic factor study was at low risk of bias,[35] while another was judged to be at low risk of bias in all but one domain.[26] The domains at highest risk of bias were study confounding, related to omission of important covariates; study participants, often due to requirement for the measurement of specific laboratory parameters (eg, leucocyte count); and statistical analysis, as a result of inadequate reporting or inappropriate exclusion of participants from the analysis (figure 6).

Each clinical prediction model/outcome pair was assessed independently and all judged to be at high risk

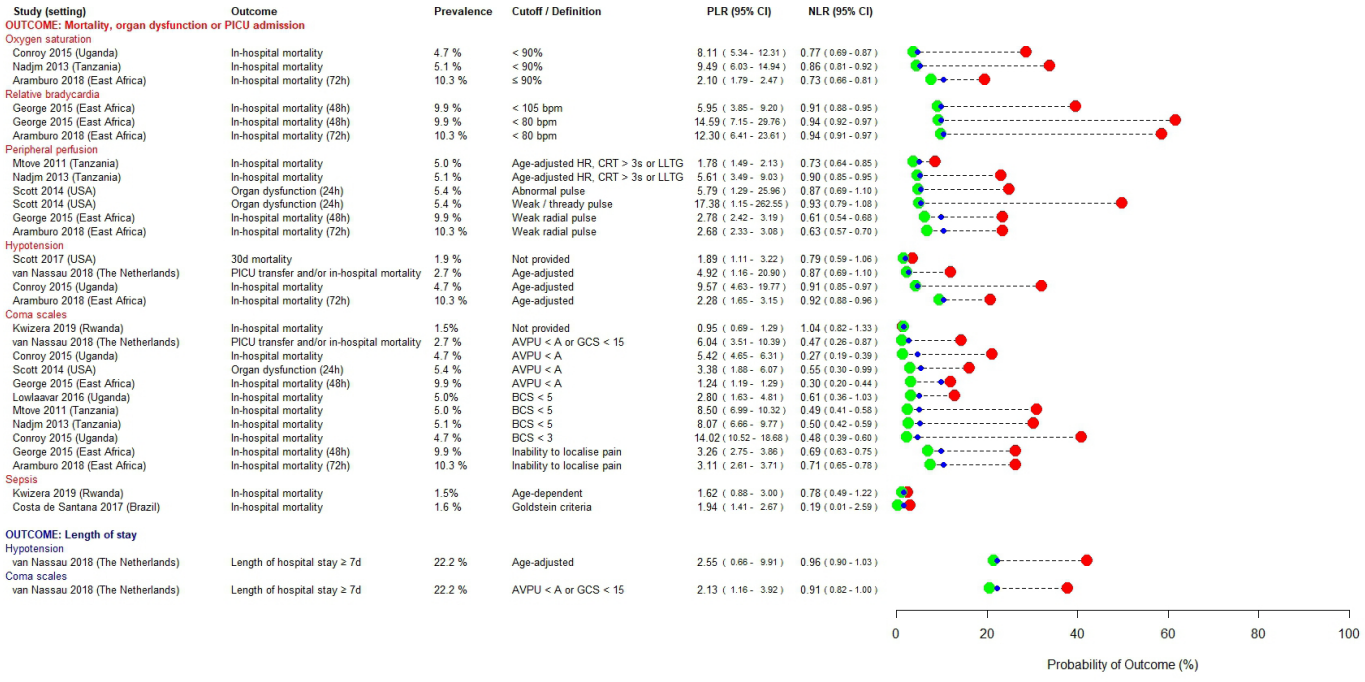

Figure 3 Prognostic factors identified as having rule-in (PLR ≥5.0) or rule-out (NLR ≤0.2) value for severe disease in at least one study—cardiovascular, respiratory or neurological signs. in the study by Costa *et al*. 'sepsis' was defined according to the systemic inflammatory response syndrome (SIRS), requiring measurement of heart rate, respiratory rate, temperature and leucocyte count. In the study by Kwizera *et al*, 'sepsis' was defined according to the qSOFA Score in children aged ≥15 years, and using a combination of temperature, mental status, respiratory distress, prostration and seizures in children aged <15 years. AVPU, alert, voice, pain or unresponsive; BCS, Blantyre Coma Score; bpm, beats per minute; CRT, capillary refill time; GCS, Glasgow Coma Score; HR, heart rate; LLTG, lower limb temperature gradient; NLR, negative likelihood ratio; PICU, paediatric intensive care unit; PLR, positive likelihood ratio; qSOFA, quick Sequential Organ Failure Assessment.

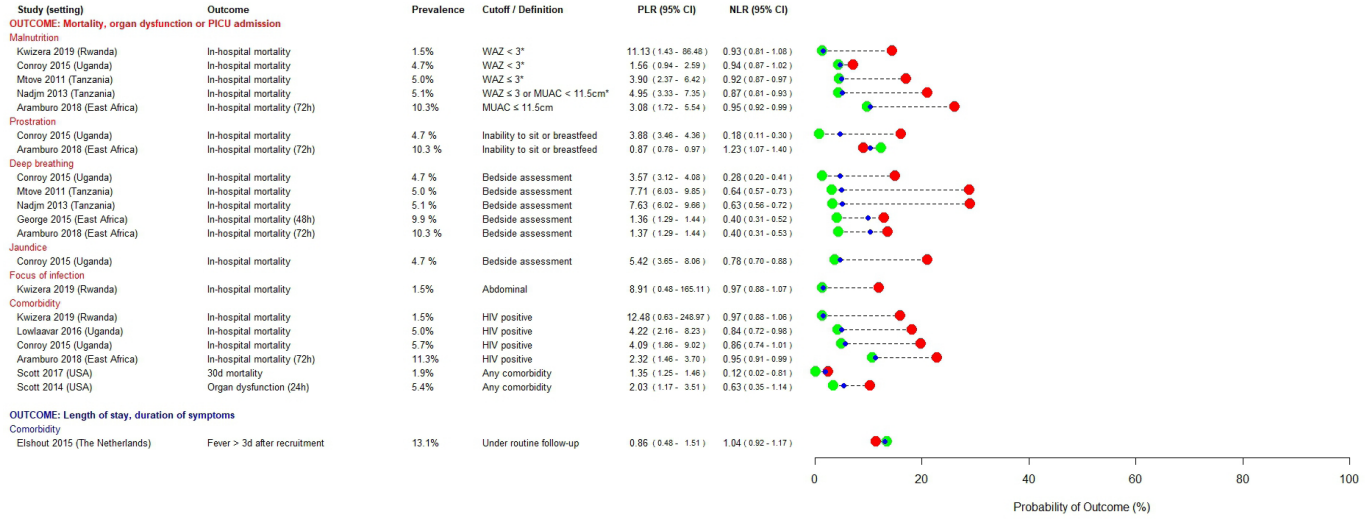

Figure 4 Prognostic factors identified as having rule-in (PLR ≥5.0) or rule-out (NLR ≤0.2) value for severe disease in at least one study—historical, anthropometric and metabolic variables. *Children with visible wasting or nutritional oedema were also classified as having severe malnutrition. In the study by Elshout *et al*, 'comorbidity' was defined as being under routine care of a paediatrician or ENT specialist. ENT, ear, nose and throat; MUAC, mid-upper arm circumference; NLR, negative likelihood ratio; PICU, paediatric intensive care unit; PLR, positive likelihood ratio; WAZ, weight-for-age z-score.

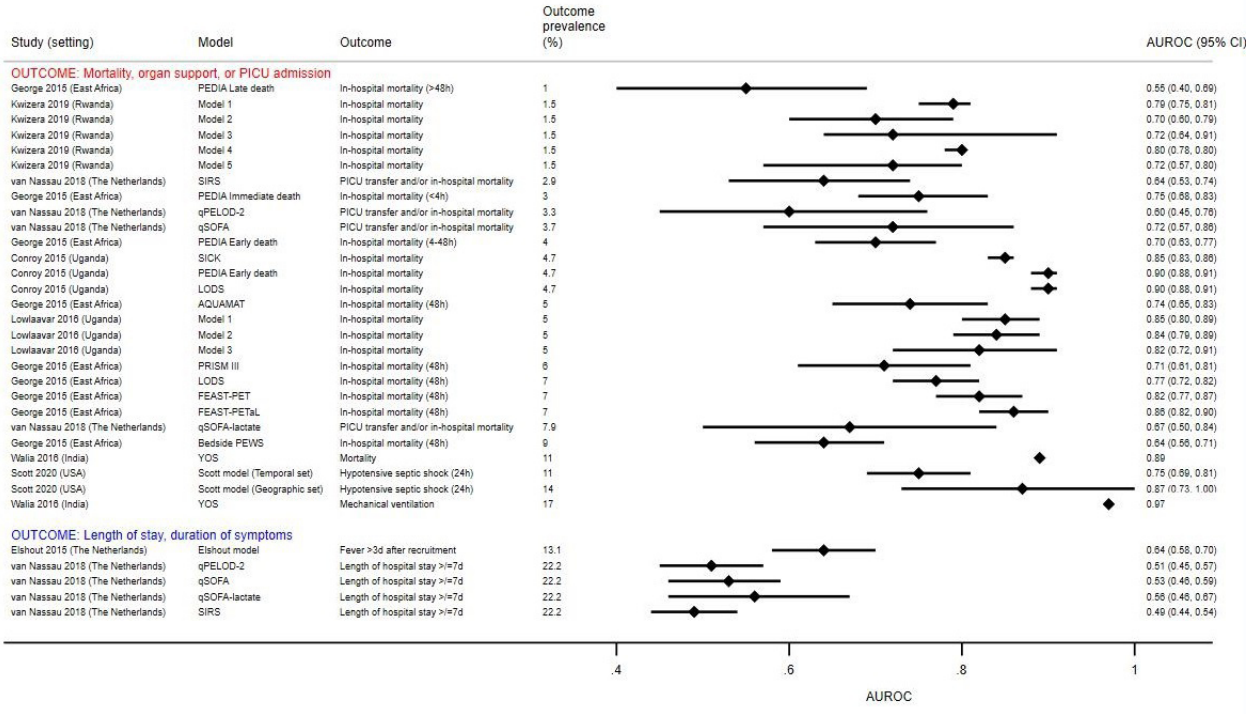

Figure 5 Discrimination of clinical prediction models to identify children at risk of severe disease. Individual studies evaluated different clinical prediction models using datasets with varying numbers of children with severe disease, depending on the data available. The outcome prevalence reflects the proportion of children with severe disease in the dataset used to evaluate that particular prediction model/outcome pair. This may be different from the overall prevalence of children with severe disease in the study, which is listed in table 1 and used to classify studies into low, moderate or high prevalence settings. No CIs were provided for the AUROC estimates in the study by Walia *et al*. AQUAMAT, African Quinine Artesunate Malaria Trial; AUROC, area under the receiver operating characteristic curve; FEAST-PET, FEAST-Paediatric Emergency Triage; FEAST-PETaL, FEAST-Paediatric Emergency Triage and Laboratory; LODS, Lambaréné Organ Dysfunction Score; PEDIA, Paediatric Early Death Index for Africa; PEWS, Paediatric Early Warning Score; PICU, paediatric intensive care unit; PRISM III, Paediatric Risk of Mortality; qPELOD-2, quick Paediatric Logistic Organ Dysfunction; qSOFA, quick Sequential Organ Failure Assessment; SICK, Signs of Inflammation in Children that Kill; SIRS, Systemic Inflammatory Response Syndrome; YOS, Yale Observation Score.

Table 2 Components of clinical prediction models evaluated in the included studies

Clinical prediction model	Variables used in the clinical prediction model in the included studies	Included study evaluating the model	Original study developing the model
AQUAMAT	Base deficit, impaired consciousness, convulsions, elevated blood urea, underlying chronic illness	George[31]	von Seidlein[60]
ELSHOUT model	Sore throat, palpable lymph nodes, duration of fever before consultation, C-reactive protein	Elshout[29]	Elshout[29]
FEAST-PET	Axillary temperature, heart rate, capillary refill time, conscious level, respiratory distress, lung crepitations, severe pallor, weak pulse	George[31]	George[31]
FEAST-PETaL	FEAST-PET with the addition of lactate, pH, blood urea nitrogen	George[31]	George[31]
KWIZERA model 1	Age, respiratory rate, heart rate, temperature, capillary refill time, altered mental state	Kwizera[41]	Kwizera[41]
KWIZERA model 2	Age, respiratory rate, heart rate, capillary refill time, altered mental state	Kwizera[41]	Kwizera[41]
KWIZERA model 3	Age, respiratory rate, temperature, capillary refill time, altered mental state	Kwizera[41]	Kwizera[41]
KWIZERA model 4	Age, respiratory rate, capillary refill time, altered mental state	Kwizera[41]	Kwizera[41]
KWIZERA model 5	Age, respiratory rate, altered mental state	Kwizera[41]	Kwizera[41]
LODS	Deep breathing, coma, and prostration	George[31]; Conroy[27]	Helbok[50]
LOWLAAVAR model 1	Conscious level, HIV, weight-for-age z-score	Lowlaavar[42]	Lowlaavar[42]
LOWLAAVAR model 2	Conscious level, HIV, mid-upper arm circumference	Lowlaavar[42]	Lowlaavar[42]
LOWLAAVAR model 3	Conscious level, mid-upper arm circumference	Lowlaavar[42]	Lowlaavar[42]
PEDIA-i	Anaemia, jaundice, indrawing, deep breathing, conscious level, prostration, convulsions/seizures, temperature	George[31]	Berkley[51]
PEDIA-e	Jaundice, indrawing, conscious level, prostration, convulsions/seizures, wasting, kwashiorkor*	George;[31] Conroy[27]	Berkley[51]
PEDIA-l	History >7 days, conscious level, prostration, convulsions/seizures, temperature, wasting, kwashiorkor	George[31]	Berkley[51]
PEWS†	Heart rate, capillary refill time, respiratory rate, oxygen saturation, systolic blood pressure	George[31]	Parshuram[61]
PRISM III‡	Heart rate, temperature, conscious level, systolic blood pressure, glucose, potassium, PCO_2, pH, acidosis, pupillary reflexes	George[31]	Pollack[62]
qPELOD-2	Systolic or mean arterial pressure, heart rate, altered mentation	van Nassau[38]	Leclerc[53]
qSOFA	Respiratory rate, altered mentation, systolic blood pressure	van Nassau[38]	Seymour[54]
qSOFA-L	qSOFA with the addition of lactate	van Nassau[38]	van Nassau[38]

Continued

Table 2 Continued

Clinical prediction model	Variables used in the clinical prediction model in the included studies	Included study evaluating the model	Original study developing the model
SCOTT model	Systolic blood pressure, diastolic blood pressure, temperature, age, respiratory rate, heart rate, arrival via emergency medical services, oncological comorbidity, indwelling central line on arrival, hospitalised in the last year	Scott[40]	Scott[40]
SICK	Level of consciousness, temperature, heart rate, respiratory rate, systolic blood pressure, SpO_2, capillary refill time, age	Conroy[27]	Kumar[52]
SIRS	Heart rate, respiratory rate, leucocyte count, temperature	van Nassau[38]	Goldstein[63]
YOS	Quality of cry, reaction to parent stimulation, state variation, colour, hydration, response to social overtures	Walia[39]	McCarthy[64]

*Kwashiorkor was not included in the PEDIA-e score in the Conroy *et al* study.
†Receipt of oxygen therapy and respiratory effort included in the original PEWS but not measured in the George *et al* study.
‡Pupillary reflexes, pH, total CO_2, arterial PaO_2, creatinine, urea, white blood cells, prothrombin time and platelets included in the original PRISM III score but not measured in the George *et al* study.
AQUAMAT, African Quinine Artesunate Malaria Trial; FEAST-PET, FEAST-Paediatric Emergency Triage; FEAST-PETaL, FEAST-Paediatric Emergency Triage and Laboratory; LODS, Lambaréné Organ Dysfunction Score; PEDIA-e, Paediatric Early Death Index for Africa (early death score); PEDIA-i, Paediatric Early Death Index for Africa (immediate death score); PEDIA-l, Paediatric Early Death Index for Africa (late death score); PEWS, Paediatric Early Warning Score; PRISM-III, Paediatric Risk of Mortality; qPELOD-2, quick Paediatric Logistic Organ Dysfunction; qSOFA, quick Sequential Organ Failure Assessment; qSOFA-L, quick Sequential Organ Failure Assessment-Lacate; SICK, Signs of Inflammation in Children that Kill; SIRS, Systemic Inflammatory Response Syndrome; YOS, Yale Observation Score.

of bias (figure 6). Most often this was due to inadequate reporting of model performance (studies reporting discrimination but not calibration), circularity between predictors and outcomes or having fewer than 100 participants with severe outcomes for model validation. It is noteworthy that one study which externally validated three models included 99 children who died.[27] Another study which derived and/or validated nine models undertook an additional external validation in a population of acutely unwell but non-febrile children (and thus not eligible for consideration in this review), which included more than 100 children who died.[31]

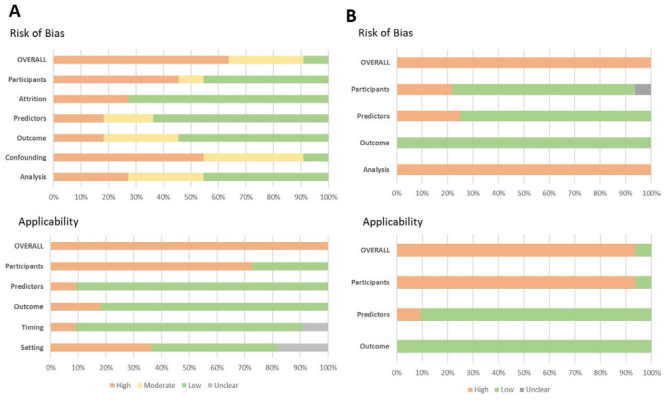

Figure 6 Risk of bias and applicability assessments for included studies using (A) the QUIPS tool (n=11 studies) and (B) PROBAST (n=33 clinical prediction model/outcome pairs from seven studies). All studies evaluating clinical prediction models were assessed using PROBAST, except for the study by Elshout *et al*, which was primarily a prognostic factor study and was therefore assessed using QUIPS. PROBAST, Prediction Model Risk of Bias Assessment Tool; QUIPS, Quality In Prognosis Studies.

In all but one study there was high concern regarding applicability to the review question.[40] This was largely due to the majority of studies including only children requiring hospitalisation, with recruitment occurring after the decision to admit had been made by the treating physician. Full details on risk of bias and applicability assessments are provided in online supplemental appendix S6.

DISCUSSION

This systematic review of prognostic factors and clinical prediction models assessing severity of disease in febrile children highlights that few well-conducted studies address this important public health question, particularly in unselected children presenting from the community. One of its main strengths is the inclusion of studies from a wide geographic context, aiding understanding of how predictive performance can vary across settings. By focusing on prognosis, we identified features that predict the likelihood that a child's illness might progress, rather than features associated with illness severity at the moment of assessment.

Most prognostic factors identified as valuable for predicting severe childhood febrile illness (PLR ≥5.0) overlapped with individual components of the most

promising clinical prediction models (AUROC ≥0.80): nutritional and HIV status, hypoxia, altered consciousness, and markers of acidosis (raised venous lactate or deep breathing) and poor peripheral perfusion (weak pulse, limb-core temperature gradient or prolonged capillary refill time).[27 31 32 34 36 38 42] Hypoglycaemia was a useful prognostic factor identified in our review, but omitted in most clinical prediction models. Many of these features, however, indicate a child that is already very unwell, reflecting the fact that most studies included only hospitalised children and focused on predicting mortality. Few prognostic factors adequately ruled-out (NLR ≤0.2) the possibility of progression to severe disease, a finding consistent with a previous systematic review evaluating the diagnostic utility of clinical features for serious bacterial infections.[10]

The major limitation of our work arises from the heterogeneity of studies, which precludes comparison of effect estimates. Second, it is difficult to determine if studies included children presenting to first-line health workers. We did not exclude studies solely based on the designated 'level' of a health facility: concerned parents in all settings use primary, secondary and tertiary care facilities as their first point-of-access. Third, most studies included only hospitalised children. This is a major barrier to understanding the potential for prognostic factors and prediction models to guide referral or admission decisions. Follow-up of children assessed as 'low-risk' (ie, those managed in the community) must be a priority for future studies seeking to determine the validity of prognostic factors and prediction models in outpatient settings.[43] Fourth, in line with other reviews we found most studies to be of low quality.[44] Recent guidance may help address this.[17] Finally, we framed the review around 'febrile illness', rather than, for example, 'clinically-suspected infection'. Our rationale was to ensure the findings were as relevant as possible for lesser-trained community health workers in resource-constrained settings, for whom a presumptive diagnosis of suspected infection can be challenging. Febrile illness is an accepted 'pragmatic point-of-entry' in these settings,[45] however, we acknowledge that some children (particularly younger infants) may not mount a fever in response to serious infection. Therefore, despite our deliberately broad definition of febrile illness (documented abnormal temperature and history of fever), and the inclusion of studies of children with 'suspected sepsis', relevant studies may have been missed. Of note, in view of a suggestion arising during the peer-review process we also performed a second MEDLINE search, using alternate search strings, which did not yield any additional eligible articles (online supplemental appendix S7).

Thirty out of 200 (15%) prognostic factors met our prespecified threshold for clinical relevance (PLR ≥5.0 or NLR ≤0.2). This may reflect the difficulty of identifying parsimonious predictors for all febrile children. While common pathophysiological pathways for severe disease have been identified across a spectrum of microbial aetiologies,[46 47] certain predictors may perform better for specific syndromes or pathogens, compared with all-cause febrile illness. Five studies in our review reported a high proportion of children as being either slide-positive or rapid diagnostic test-positive for malaria. Notwithstanding the issues of co-infection and/or concomitant incidental parasitaemia in settings of high malaria endemicity, it is possible that the findings of these studies are more pertinent to children with malaria. However, four of these studies compared the prognostic performances of hyperlactaemia, hypoglycaemia and the prediction models SICK, LODS and PEDIA, and found them to be broadly consistent between children with malaria, non-malarial fever and invasive bacterial disease.[26 27 32 34] Furthermore, as can be seen in figures 2–4, a number of predictors identified in malaria endemic regions also demonstrated prognostic utility in contexts where malaria is not endemic (eg, venous lactate, impaired peripheral perfusion, hypotension and altered consciousness). This, in conjunction with the subgroup analyses performed in the original studies, gives us a degree of confidence that the prognostic factors that we have identified are generalisable across different infecting pathogens. Nonetheless, future reviews using search strategies developed to retrieve syndrome-specific or pathogen-specific studies should explore this.

Another potential explanation for the relatively few valuable prognostic factors identified is work-up bias. In most studies, predictors were available to the treating clinicians: abnormal values are likely to have been acted on and predictive performance underestimated. For most predictors, particularly clinical signs, this is unavoidable as blinding is often neither possible nor ethical. When feasible, randomisation is required to definitively assess their potential impact.[48] This is particularly important for new tests proposed in resource-limited settings. For example, both lactate and hypoxia were identified as potentially of value in this review but introducing tests for these parameters at all first-line health facilities across the tropics would incur substantial cost, and as their predictive value may vary in different settings, could result in unnecessary or missed referrals. Randomisation can help determine whether new tests such as these add value to simple clinical assessment.[49]

Clinical prediction models performed better when derived and validated in similar populations[27]: in East Africa LODS and PEDIA-e (both derived in sub-Saharan Africa)[50 51] were superior to SICK (originally derived in India).[52] Model performance also improved when predicting the same outcome as the derivation study: quick Sequential Organ Failure Assessment and quick Paediatric Logistic Organ Dysfunction, derived to predict mortality, performed poorly when predicting prolonged length of stay.[38 53 54] These findings highlight the importance of deriving prediction models using populations and outcomes appropriate to the clinical question. While mortality is a 'hard' outcome, it seldom occurs in primary care. Furthermore, its reflection of disease

severity is influenced (mediated) by the level of care. It is striking that in Tanzania a raised lactate conveyed a post-test probability of in-hospital mortality comparable to that of 'organ dysfunction within 24 hours of arrival' in a similar prevalence setting in the USA.[32 34 37] Rather than relying on models derived in secondary care to generalise to outpatient settings across different epidemiological landscapes, alternative ways to quantify disease severity, which consider local context yet avoid circularity between predictor variables and outcome definitions, will be important to facilitate comparisons across settings and explore generalisability of risk prediction tools. Finally, the fact that most studies summarised model performance using only the AUROC means that is difficult to appreciate what the impact might be on clinical decision making.[55]

In LMIC primary care contexts, many variables are not feasible to collect,[56] and as noted above, some may incur substantial cost. Interestingly, HIV and nutritional status were both identified in our review and represent the only prognostic factors meeting our threshold for clinical relevance that may not necessarily reflect a child that is overtly very unwell. While biological plausibility for the prognostic utility of these two variables is high, it should be noted that the study which identified them was small and correspondingly the CI for the PLR is wide.[41] The WHO's Integrated Management of Childhood Illnesses 'Danger Signs' are recommended to guide referrals from community healthcare providers in resource-constrained settings.[57] Of these, only altered consciousness was widely represented among included studies, and most found it to be a good predictor of severe disease.[26 27 31 32 34 36 38 41 42] History of convulsions was examined in two studies while other 'Danger Signs' were not evaluated.[26 27]

CONCLUSION

Our findings emphasise the limitations of individual prognostic factors. Performance varies widely across settings and clinicians must be cognisant not to over interpret individual predictors. While prediction models can support clinical decision making, they must be derived and validated using appropriate methodology, and populations and outcomes relevant to the clinical problem. For the identification of children at risk of severe febrile illness, this will require multiple, large, collaborative, research initiatives across different settings, which collect harmonised data on predictors and outcomes,[58 59] and include unselected children presenting from the community.

Author affiliations
[1]Cambodia-Oxford Medical Research Unit, Angkor Hospital for Children, Siem Reap, Cambodia
[2]Centre for Tropical Medicine and Global Health, University of Oxford, Oxford, UK
[3]Unisanté Centre for Primary Care and Public Health, University of Lausanne, Lausanne, Switzerland
[4]University of Basel, Basel, Switzerland
[5]Swiss Tropical and Public Health Institute, Basel, Basel-Stadt, Switzerland
[6]Department of Women and Children's Health, King's College London, London, UK
[7]Academic Centre of General Practice, University of Leuven, Leuven, Flanders, Belgium
[8]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
[9]Ifakara Health Institute, Dar-es-Salaam, Tanzania
[10]Department of Pediatrics and Child Health, Muhimbili University Health and Allied Sciences, Dar-es-Salaam, Tanzania
[11]Mahidol-Oxford Tropical Medicine Research Unit, Bangkok, Thailand
[12]Division of Emergency Medicine, Department of Pediatrics, University Children's Hospital, Inselpital, University of Bern, Bern, Switzerland

Twitter Arjun Chandna @arji_barji, Rainer Tan @DrRainerTan and Paul Turner @PaulTurnerMicro

Acknowledgements We acknowledge the support of Nia Roberts and Jolanda Elmers in helping to translate MEDLINE search terms and access certain articles. We are grateful to Hanne Boon for assisting with the R code to create the dumbbell plots.

Contributors AC conceived the study; AC, RT, YL, PT and KK defined the review strategy; AC and RT conducted the search, screened retrieved articles and extracted the data; AC, RT, MC, AVDB and JV assessed quality of included articles; AC and RT analysed the data and drafted the report; CK provided statistical oversight; PT and KK commented on the drafted report; AC, RT, MC, AVDB, JV, CK, NS, YL, PT and KK commented on and approved the final manuscript.

Funding AC is supported by a Wellcome Trust Doctoral Training Fellowship. RT is supported by the Botnar Foundation. The Cambodia Oxford Medical Research Unit and Mahidol-Oxford Tropical Medicine Research Unit are part of the Wellcome Trust Thailand Africa and Asia Programme, which receives core-funding from the UK Wellcome Trust (106 698Z/14/Z). The project also received seed funding from the Tropical Health Education Trust in the UK.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon request. All data relevant to the study are included in the article or uploaded as supplementary information. The protocol for the study is available from: https://www.crd.york.ac.uk/PROSPERO/display_record.php?RecordID=140542.

Author note Technical Advisory Panel members: Jalemba Aluvaala, Quique Bassat, David Bell, John Crump, W. Conrad Liles, Rianne Oostenbrink and Shunmay Yeung. Affiliations are listed in online supplemental appendix S2.

ORCID iDs
Arjun Chandna http://orcid.org/0000-0003-1313-7922
Paul Turner http://orcid.org/0000-0002-1013-7815

REFERENCES

1 Prasad N, Sharples KJ, Murdoch DR, *et al*. Community prevalence of fever and relationship with malaria among infants and children in low-resource areas. *Am J Trop Med Hyg* 2015;93:178–80.
2 World Health Organization. *WHO informal consultation on fever management in peripheral health care settings: a global review of evidence and practice*. Geneva: World Health Organization, 2013.

3 Buntinx F, Mant D, Van den Bruel A, *et al*. Dealing with low-incidence serious diseases in general practice. *Br J Gen Pract* 2011;61:43–6.
4 Prasad N, Murdoch DR, Reyburn H, *et al*. Etiology of severe febrile illness in low- and middle-income countries: a systematic review. *PLoS One* 2015;10:e0127962.
5 Kruk ME, Gage AD, Joseph NT, *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* 2018;392:2203–12.
6 McDonald CR, Weckman A, Richard-Greenblatt M, *et al*. Integrated fever management: disease severity markers to triage children with malaria and non-malarial febrile illness. *Malar J* 2018;17:353.
7 Molyneux E, Ahmad S, Robertson A. Improved triage and emergency care for children reduces inpatient mortality in a resource-constrained setting. *Bull World Health Organ* 2006;84:314–9.
8 Keitel K, D'Acremont V. Electronic clinical decision algorithms for the integrated primary care management of febrile children in low-resource settings: review of existing tools. *Clin Microbiol Infect* 2018;24:845–55.
9 Moons KGM, Altman DG, Reitsma JB, *et al*. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.
10 Van den Bruel A, Haj-Hassan T, Thompson M, *et al*. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet* 2010;375:834–45.
11 Van den Bruel A, Thompson MJ, Haj-Hassan T, *et al*. Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review. *BMJ* 2011;342:d3082.
12 Southeast Asia Infectious Disease Clinical Research Network. Causes and outcomes of sepsis in Southeast Asia: a multinational multicentre cross-sectional study. *Lancet Glob Health* 2017;5:e157–67.
13 Wahl B, O'Brien KL, Greenbaum A, *et al*. Burden of Streptococcus pneumoniae and Haemophilus influenzae type B disease in children in the era of conjugate vaccines: global, regional, and national estimates for 2000-15. *Lancet Glob Health* 2018;6:e744–57.
14 Moons KGM, de Groot JAH, Bouwmeester W, *et al*. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the charms checklist. *PLoS Med* 2014;11:e1001744.
15 Riley RD, Moons KGM, Snell KIE, *et al*. A guide to systematic review and meta-analysis of prognostic factor studies. *BMJ* 2019;364:k4597.
16 Hayden JA, van der Windt DA, Cartwright JL, *et al*. Assessing bias in studies of prognostic factors. *Ann Intern Med* 2013;158:280–6.
17 Moons KGM, Wolff RF, Riley RD, *et al*. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1–33.
18 Moher D, Liberati A, Tetzlaff J. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* 2009;151:264. **151**(4): 264-9, w64.
19 Schlapbach LJ, Kissoon N. Defining pediatric sepsis. *JAMA Pediatr* 2018;172:313–4.
20 Geersing G-J, Bouwmeester W, Zuithoff P, *et al*. Search filters for finding prognostic and diagnostic prediction studies in MEDLINE to enhance systematic reviews. *PLoS One* 2012;7:e32844.
21 Haynes RB, McKibbon KA, Wilczynski NL, *et al*. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *BMJ* 2005;330:1179.
22 Ingui BJ, Rogers MA. Searching for clinical prediction rules in MEDLINE. *J Am Med Inform Assoc* 2001;8:391–7.
23 Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. what are the results and will they help me in caring for my patients? the evidence-based medicine Working group. *JAMA* 1994;271:703–7.
24 Fischer JE, Bachmann LM, Jaeschke R. A readers' guide to the interpretation of diagnostic test properties: clinical example of sepsis. *Intensive Care Med* 2003;29:1043–51.
25 Debray TPA, Damen JAAG, Snell KIE, *et al*. A guide to systematic review and meta-analysis of prediction model performance. *BMJ* 2017;356:i6460.
26 Aramburo A, Todd J, George EC, *et al*. Lactate clearance as a prognostic marker of mortality in severely ill febrile children in East Africa. *BMC Med* 2018;16:37.
27 Conroy AL, Hawkes M, Hayford K, *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* 2015;19:47.
28 Costa de Santana M. Duarte Mello Amoedo C, Nascimento-Carvalho CM. Clinical and epidemiological characteristics of children admitted with fever in emergency department with or without sepsis. *J Infect Dev Ctries* 2017;11:597–603.
29 Elshout G, Kool M, Bohnen AM, *et al*. Predicting prolonged duration of fever in children: a cohort study in primary care. *Br J Gen Pract* 2015;65:e578–84.
30 Freyne B, Divilley R, Kissoon-Harrison G, *et al*. Field testing the utility of procalcitonin and the acute infantile observation score in febrile infants 6 to 36 months old presenting to the pediatric emergency department with no obvious focus of infection. *Clin Pediatr* 2013;52:503–6.
31 George EC, Walker AS, Kiguli S, *et al*. Predicting mortality in sick African children: the feast paediatric emergency triage (PET) score. *BMC Med* 2015;13:174.
32 Mtove G, Nadjm B, Hendriksen ICE, *et al*. Point-Of-Care measurement of blood lactate in children admitted with febrile illness to an African district hospital. *Clin Infect Dis* 2011;53:548–54.
33 Mwandama D, Mwale C, Bauleni A, *et al*. Clinical outcomes among febrile children aged 2 to 59 months with negative malaria rapid diagnostic test results in Mchinji district, Malawi. *Malawi Med J* 2016;28:150–3.
34 Nadjm B, Mtove G, Amos B, *et al*. Blood glucose as a predictor of mortality in children admitted to the hospital with febrile illness in Tanzania. *Am J Trop Med Hyg* 2013;89:232–7.
35 Scott HF, Brou L, Deakyne SJ, *et al*. Association between early lactate levels and 30-day mortality in clinically suspected sepsis in children. *JAMA Pediatr* 2017;171:249–55.
36 Scott HF, Donoghue AJ, Gaieski DF, *et al*. Effectiveness of physical exam signs for early detection of critical illness in pediatric systemic inflammatory response syndrome. *BMC Emerg Med* 2014;14:24.
37 Scott HF, Donoghue AJ, Gaieski DF, *et al*. The utility of early lactate testing in undifferentiated pediatric systemic inflammatory response syndrome. *Acad Emerg Med* 2012;19:1276–80.
38 van Nassau SC, van Beek RH, Driessen GJ, *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* 2018;6:266.
39 Walia S, M HM, Kumble A, *et al*. Yale observation scale as a predictor of bacteremia and final outcome in 3-36 months old febrile children admitted in tertiary health centres: a hospital-based cross-sectional study. *Asian Journal of Pharmaceutical and Clinical Research* 2016;9:219.
40 Scott HF, Colborn KL, Sevick CJ, *et al*. Development and validation of a predictive model of the risk of pediatric septic shock using data known at the time of hospital arrival. *J Pediatr* 2020;217:145–51.
41 Kwizera A, Kissoon N, Musa N, *et al*. A machine Learning-Based triage tool for children with acute infection in a low resource setting. *Pediatr Crit Care Med* 2019;20:1–30.
42 Lowlaavar N, Larson CP, Kumbakumba E, *et al*. Pediatric in-hospital death from infectious disease in Uganda: derivation of clinical prediction models. *PLoS One* 2016;11:e0150683.
43 Hansoti B, Jenson A, Keefe D, *et al*. Reliability and validity of pediatric triage tools evaluated in low resource settings: a systematic review. *BMC Pediatr* 2017;17:37.
44 Bouwmeester W, Zuithoff NPA, Mallett S, *et al*. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med* 2012;9:e1001221.
45 World Health Organization. *IMCI chart booklet*. Geneva: World Health Organization, 2014.
46 Leligdowicz A, Richard-Greenblatt M, Wright J, *et al*. Endothelial activation: the Ang/Tie axis in sepsis. *Front Immunol* 2018;9:838.
47 Kinasewitz GT, Yan SB, Basson B, *et al*. Universal changes in biomarkers of coagulation and inflammation occur in patients with severe sepsis, regardless of causative micro-organism [ISRCTN74215569]. *Crit Care* 2004;8:R82–90.
48 Tan R, Kagoro F, Levine GA, *et al*. Clinical outcome of febrile Tanzanian children with severe malnutrition using anthropometry in comparison to clinical signs. *Am J Trop Med Hyg* 2020;102:427–35.
49 Keitel K, Samaka J, Masimba J, *et al*. Safety and efficacy of C-reactive Protein-guided antibiotic use to treat acute respiratory infections in Tanzanian children: a planned subgroup analysis of a randomized controlled Noninferiority trial evaluating a novel electronic clinical decision algorithm (ePOCT). *Clin Infect Dis* 2019;69:1926–34.
50 Helbok R, Kendjo E, Issifou S, *et al*. The Lambaréné organ dysfunction score (LODS) is a simple clinical predictor of fatal malaria in African children. *J Infect Dis* 2009;200:1834–41.
51 Berkley JA, Ross A, Mwangi I, *et al*. Prognostic indicators of early and late death in children admitted to district hospital in Kenya: cohort study. *BMJ* 2003;326:361.
52 Kumar N, Thomas N, Singhal D, *et al*. Triage score for severity of illness. *Indian Pediatr* 2003;40:204–10.

53 Leclerc F, Duhamel A, Deken V, *et al*. Can the pediatric logistic organ Dysfunction-2 score on day 1 be used in clinical criteria for sepsis in children? *Pediatr Crit Care Med* 2017;18:758–63.
54 Seymour CW, Liu VX, Iwashyna TJ, *et al*. Assessment of clinical criteria for sepsis: for the third International consensus definitions for sepsis and septic shock (Sepsis-3). *JAMA* 2016;315:762–74.
55 Fackler JC, Rehman M, Winslow RL. Please welcome the new team member: the algorithm. *Pediatr Crit Care Med* 2019;20:1200–1.
56 Fung JST, Akech S, Kissoon N, *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* 2019;14:e0211274.
57 World Health Organization. *Integrated management of childhood Ilness chart booklet*. Geneva: World Health Organization, 2014.
58 Li E. Guidelines for the standardized collection of predictor variables in studies for pediatric sepsis. In: Ansermino M, ed. *Scholars portal Dataverse*. V2 ed, 2020.
59 Wooldridge G, Murthy S, Kissoon N. Core outcome set in paediatric sepsis in low- and middle-income countries: a study protocol. *BMJ Open* 2020;10:e034960.
60 von Seidlein L, Olaosebikan R, Hendriksen ICE, *et al*. Predicting the clinical outcome of severe falciparum malaria in African children: findings from a large randomized trial. *Clin Infect Dis* 2012;54:1080–90.
61 Parshuram CS, Hutchison J, Middaugh K. Development and initial validation of the bedside paediatric early warning system score. *Crit Care* 2009;13:R135.
62 Pollack MM, Patel KM, Ruttimann UE. Prism III: an updated pediatric risk of mortality score. *Crit Care Med* 1996;24:743–52.
63 Goldstein B, Giroir B, Randolph A, *et al*. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr Crit Care Med* 2005;6:2–8.
64 McCarthy PL, Sharpe MR, Spiesel SZ, *et al*. Observation scales to identify serious illness in febrile children. *Pediatrics* 1982;70:802–9.

Predictors of disease severity in children presenting from the community with febrile illnesses: a systematic review of prognostic studies

Authors and affiliations

Arjun Chandna MD*,[1,2] Rainer Tan MD*,[3,4,5] Michael J Carter PhD,[6] Prof Ann Van den Bruel PhD,[7] Prof Jan Verbakel PhD,[7,8] Constantinos Koshiaris PhD,[8] Nahya Salim PhD,[9,10] Yoel Lubell PhD,[2,11] Paul Turner PhD**,[1,2] Kristina Keitel PhD**[4,12]

* Contributed equally
** Contributed equally

1. Cambodia Oxford Medical Research Unit, Angkor Hospital for Children, Cambodia
2. Centre for Tropical Medicine and Global Health, University of Oxford, UK
3. Unisanté, Centre for Primary Care and Public Health, University of Lausanne, Switzerland
4. Swiss Tropical and Public Health Institute (SwissTPH), Basel, Switzerland
5. University of Basel, Basel, Switzerland
6. Department of Women and Children's Health, King's College London, London, UK
7. Academic Centre of General Practice, University of Leuven, Leuven, Belgium
8. Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, UK
9. Ifakara Health Institute, Dar es Salaam, United Republic of Tanzania
10. Department of Pediatrics and Child Health, Muhimbili University Health and Allied Sciences (MUHAS), Dar es Salaam, United Republic of Tanzania
11. Mahidol-Oxford Tropical Medicine Research Unit, Wellcome Trust Thailand Africa and Asia Programme
12. Division of Emergency Medicine, Department of Pediatrics, Inselspital Bern, Switzerland

Corresponding author

Dr. Arjun Chandna
Cambodia Oxford Medical Research Unit
Angkor Hospital for Children
Siem Reap
Cambodia
Email: arjun@tropmedres.ac
Tel: +855-85-712-586

S1 Appendix

Table 1. Search strategy built following Cochrane Prognosis Methods Group recommendations[1] and adapting published search strings.[2-4]

	MEDLINE	Embase	Science Citation Index via Web of Science
1	Fever[MeSH Terms] OR Fever[Title/Abstract] OR Febrile[Title/Abstract] OR “suspected sepsis”[Title/Abstract]	fever/ or (fever* or febrile or suspected sepsis).ti,ab,kw.	TS=(fever* or febrile or "suspected sepsis")
2	pediatrics[MeSH Terms] OR pediatric*[Title/Abstract] OR paediatric*[Title/Abstract] OR child[MeSH Terms] OR child*[Title/Abstract] OR Infant[Mesh:NoExp] OR infant[Title/Abstract]	pediatrics/ or child/ or infant/ or preschool child/ or school child/ or toddler/ or boy/ or girl/ or (pediatric* or paediatric* or child* or infant*).mp.	TS=(pediatric* or paediatric* or child* or infant*)
3	(((((((Validat*[tw] OR Predict*[ti] OR Rule*[tw]) OR (Predict*[tw] AND (Outcome*[tw] OR Risk*[tw] OR Model*[tw])) OR ((History OR Variable*[tw] OR Criteria OR Scor*[tw] OR Characteristic*[tw] OR Finding*[tw] OR Factor*[tw]) AND (Predict*[tw] OR Model*[tw] OR Decision*[tw] OR Identif*[tw] OR Prognos*[tw])) OR (Decision*[tw] AND (Model*[tw] OR Clinical*[tw] OR “Logistic Models”[MeSH Terms])) OR (Prognostic AND (History OR Variable*[tw] OR Criteria OR Scor*[tw] OR Characteristic*[tw] OR Finding*[tw] OR Factor*[tw] OR Model*[tw]))))) OR ("Stratification" OR "ROC Curve"[MeSH Terms] OR "Discrimination" OR "Discriminate" OR "c-statistic" OR "c statistic" OR "Area under the curve" OR "AUC" OR "Calibration" OR "Indices" OR "Algorithm" OR "Multivariable")))))	predict*.ti. or (validat* or rule* or (predict and (outcome* or risk* or model*)) or ((history or variable or criteria or scor* or characteristic* or finding* or factor*) and (predict* or model* or decision* or identif* or prognos*)) or (decision* and (model* or clinical*)) or (prognostic and (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model*)) or stratification or discrimination or discriminate or c-statistic or "c statistic" or auc or calibration or indices or algorithm or multivariable).mp. or statistical model/ or "receiver operating characteristic"/ or "area under the curve"/	TI=(predict*) OR TS=(validat* or rule*) OR TS=((predict and (outcome* or risk* or model*))) OR TS=(((history or variable or criteria or scor* or characteristic* or finding* or factor*) and (predict* or model* or decision* or identif* or prognos*))) OR TS=((decision* and (model* or clinical*))) OR TS=((prognostic and (history or variable* or criteria or scor* or characteristic* or finding* or factor* or model*))) OR TS=(stratification or discrimination or discriminate or c-statistic or "c statistic" or auc or calibration or indices or algorithm or multivariable)
4	death[MeSH Terms] OR death[Title/Abstract] OR mortality[MeSH Terms] OR mortality[Title/Abstract] OR systemic inflammatory response syndrome[MeSH Terms] OR “systemic inflammatory response syndrome”[Title/Abstract] OR SIRS[Title/Abstract] OR sepsis[Title/Abstract] OR septic*[Title/Abstract] OR “severe disease*”[Title/Abstract] OR “severe infection*”[Title/Abstract] OR “severe bacterial infection*”[Title/Abstract] OR “severe illness”[Title/Abstract] OR “severe febrile illness”[Title/Abstract] OR “serious disease*”[Title/Abstract] OR “serious infection*”[Title/Abstract] OR “serious bacterial infection*”[Title/Abstract] OR “serious illness”[Title/Abstract] OR “serious febrile illness”[Title/Abstract]	mortality/ or childhood mortality/ or infant mortality/ or exp mortality rate/ or death/ or child death/ or fatality/ sepsis/ or systemic inflammatory response syndrome/ or exp septic shock/ or septicemia/ or (death or mortality or systemic inflammatory response or sirs or sepsis or septic* or ((severe or serious) adj2 (disease or illness* or infection*))).mp.	TS=(death or mortality or "systemic inflammatory response" or sirs or sepsis or septic*) OR TS=(((severe or serious) NEAR/2 (disease or illness* or infection*)))
5	1 AND 2 AND 3 AND 4	1 and 2 and 3 and 4	#4 AND #3 AND #2 AND #1
6	(“1999/05/31”[Date - Publication] : “2020/04/30”[Date – Publication])	conference*.pt.	#4 AND #3 AND #2 AND #1 Refined by: PUBLICATION YEARS: (2020 OR 2019 OR 2010 OR 2002 OR 2018 OR 2009 OR 2001 OR 2017 OR 2008 OR 2000 OR 2016 OR 2007 OR 1999 OR 2015 OR 2006 OR 2014 OR 2005 OR 2013 OR 2004 OR 2012 OR 2003 OR 2011)
7	5 AND 6	5 not 6	#4 AND #3 AND #2 AND #1 Refined by: PUBLICATION YEARS: (2020 OR 2019 OR 2010 OR 2002 OR 2018 OR 2009 OR 2001 OR 2017 OR 2008 OR 2000 OR 2016 OR 2007 OR 1999 OR 2015 OR 2006 OR 2014 OR 2005 OR 2013 OR 2004 OR 2012 OR 2003 OR 2011) AND [excluding] DOCUMENT TYPES: (MEETING ABSTRACT OR PROCEEDINGS PAPER)

S2 Appendix

Table 2. Membership of the Technical Advisory Panel (domain experts) responsible for peer-reviewing the search strategy, identifying omitted articles and suggesting key authors whose publication lists were reviewed.

Technical Advisory Panel member	Affiliation	Key authors proposed by Technical Advisory Panel
Dr. Jalemba Aluvaala	Paediatrics and Child Health, University of Nairobi, Nairobi, Kenya; KEMRI-Wellcome Trust Research Programme, Nairobi, Kenya	Ambrose Agweyu Andre Siqueira Anna Seale Anthony Scott Christopher C Moore Climent Casals-Pascual Elizabeth Molyneux Henriette Moll Kathryn Maitland Jay Berkeley Elizabeth Molyneux Quique Bassat Kristina E Rudd Martin Otyek Opio Michaëla A M Huson Mike English Mike Levin Ruud Nijman Samuel Akech Tim Baker Trevor Duke
Professor Quique Bassat	Centro de Investigação em Saúde de Manhiça, Maputo, Mozambique; ISGlobal, Hospital Clínic-Universitat de Barcelona, Barcelona, Spain; Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain.	
Dr. David Bell	Foundation for Innovative New Diagnostics (FIND), Campus Biotech, Building B, Level 0, Chemin des Mines 9, 1202, Geneva, Switzerland.	
Professor John Crump	Division of Infectious Diseases and International Health, Duke University Medical Center, Durham, North Carolina; Centre for International Health, University of Otago, Dunedin, New Zealand.	
Professor W. Conrad Liles	Department of Medicine, University of Washington, Seattle, WA.	
Dr. Rianne Oostenbrink	Department of General Paediatrics, Erasmus Medical Center Sophia Children's Hospital, Rotterdam, Netherlands.	
Dr. Shunmay Yeung	Clinical Research Department, London School of Hygiene and Tropical Medicine, London, UK; Department of Paediatrics, Imperial College Healthcare NHS Trust, London, UK.	

S3 Appendix

Table 3. Data extraction sheet based on the CHARMS and CHARMS-PF checklists

Domain	Item	General	Applicability	Risk of bias	Extraction
Study	Study label	YES	NO	NO	
	Year of publication	YES	NO	NO	
	Journal of publication	YES	NO	NO	
	DOI	YES	NO	NO	
Source of data	Study design	YES	YES	YES	
	Target population	NO	YES	NO	
Participants	Single center or multi-center	YES	YES	YES	
	Number of centers recruiting	YES	NO	NO	
	Type of centers recruiting	YES	YES	YES	
	Location of study	YES	NO	NO	
	Recruitment method	YES	YES	YES	
	Recruitment setting	YES	YES	YES	
	Age range	YES	YES	YES	
	Fever definition + duration	YES	YES	YES	
	Inclusion criteria	YES	YES	YES	
	Exclusion criteria	YES	YES	YES	
	Participant description	YES	NO	NO	
	Study dates	YES	YES	NO	
Outcomes to be predicted	Prognostic outcome and definition	YES	YES	YES	
	Method of measurement of outcome	NO	NO	YES	
	Same outcome definition for all participants	NO	NO	YES	
	Same measurement of outcome for all participants	NO	NO	YES	
	Type of outcome (single or combined endpoints?)	YES	YES	NO	
	Outcomes assessed without knowledge of the candidate predictor (blinded)?	NO	NO	YES	
	Were candidate prognostic factors part of the outcome (e.g. when using a panel or consensus outcome measurement)?	NO	NO	YES	
	Time of outcome occurrence	YES	YES	NO	
Prognostic factors	Demographic prognostic factors	YES	NO	NO	
	Anthropometric prognostic factors	YES	NO	NO	
	Socioeconomic prognostic factors	YES	NO	NO	
	Historical (PMH) prognostic factors	YES	NO	NO	
	Clinical symptoms (current and historical during the illness) prognostic factors	YES	NO	NO	
	Clinical signs prognostic factors	YES	NO	NO	
	Vital signs prognostic factors	YES	NO	NO	
	Laboratory measures prognostic factors	YES	NO	NO	
	Score prognostic factors with definition and weights	YES	NO	NO	
	Method for measurement of prognostic factors	NO	NO	YES	

	Method of measurement of PFs is the same for all study participants?	NO	NO	YES	
	Setting of measurement of PF	YES	YES	YES	
	Setting of measurement of PF is the same for all study participants?	NO	NO	YES	
	Timing of prognostic factor measurement	NO	YES	YES	
	Prognostic factor assessed blinded for outcome?	NO	NO	YES	
	Handling of prognostic factor in the analysis (continuous, linear, categorised, non-linear transformations)	NO	NO	YES	
Sample size	Number of participants	YES	NO	NO	
	Number of refusals	NO	NO	YES	
	Number of outcomes/events	YES	NO	NO	
	For model studies: Number of outcomes/events in relation to the number of candidate prognostic factors (events per variable)	NO	NO	YES	
Missing data	Proportion of data on PF available for analysis	NO	NO	YES	
	Number of participants with missing data for each outcome	NO	NO	YES	
	Method used for missing data	NO	NO	YES	
Analysis	Modelling method (linear, logistic, cox, parametric survival, competing risks, regression)	YES	NO	YES	
	How modelling assumptions were checked (in particular, for time-to-event outcomes and the analysis of hazard ratios, the method for assessing non-proportional hazards (non-constant hazard ratios over time)	NO	NO	YES	
	Method for selection of PF for INCLUSION in multivariable modelling (all considered, preselection of established PF, retain only those significant from univariable analysis)	NO	NO	YES	
	Method for selection of PF DURING multivariable modelling	NO	NO	YES	
	Inclusion of additional PF (not measured at admission or not included in above categories) for multivariable modelling?	NO	YES	YES	
	Criteria used for any selection or exclusion of PF DURING multivariable modelling (P value, Akaike info criterion)	NO	NO	YES	
	Method of handling each continuous PF (dichotomisation, categorisation, linear, non-linear), including values of any cut-points used and their justification for non-linear relationships (splines, fractional polynomials)	NO	NO	YES	
Results	Unadjusted effect estimates for each PF	YES	NO	NO	
	Adjusted effect estimates for each PF	YES	NO	NO	
Interpretation and discussion	Interpretation of presented results	YES	YES	YES	
	Comparison with other studies	YES	YES	NO	
	Discussion of generalisability	YES	YES	NO	
	Strengths	YES	YES	YES	
	Limitations	YES	YES	YES	

S4 Appendix

Table 4. Cut-points for clinical prediction models evaluated in the included studies associated with rule-in (positive likelihood ratio ≥ 5.0) or rule-out (negative likelihood ratio ≤ 0.2) value for progression to severe disease.

Model	Study	Outcome	Model score range	Cut-point to rule-in	PLR	NLR	Cut-point to rule-out	PLR	NLR
AQUAMAT	George 2015	In-hospital mortality (48h)	0-5	≥ 4	8.24	0.94			
FEAST-PET	George 2015	In-hospital mortality (48h)	0-10	≥ 6	7.95	0.54	≥ 3	1.36	0.10
LODS	George 2015	In-hospital mortality (48h)	0-3				≥ 1	1.13	0.00
LODS[a]	Conroy 2015	In-hospital mortality	0-3	> 1	6.49	0.21			
PEDIA-i	George 2015	In-hospital mortality (48h)	0-13	≥ 10	5.43	0.94	≥ 4	1.29	0.00
PEDIA-e	George 2015	In-hospital mortality (48h)	0-9	≥ 5	6.38	0.92	≥ 1	1.03	0.00
PEDIA-l	George 2015	In-hospital mortality (48h)	0-7	≥ 4	7.87	0.94	≥ 1	1.02	0.00
PEWS	George 2015	In-hospital mortality (48h)	0-19	≥ 15	5.66	0.97	≥ 1	1.00	0.00
PRISM III	George 2015	In-hospital mortality (48h)	0-24	≥ 6	5.70	0.58			
qPELOD-2[*]	van Nassau 2018	In-hospital mortality or PICU transfer	0-3	≥ 2	17.08	0.79			
qSOFA[*]	van Nassau 2018	In-hospital mortality or PICU transfer	0-3	≥ 2	7.46	0.54			
YOS	Walia 2016	Mortality	6-30	> 21	6.23	0.11	> 21[†]	6.23	0.11
YOS	Walia 2016	Mechanical ventilation	6-30	> 21	12.05	0.00	> 21[†]	12.05	0.00

NLR = negative likelihood ratio; PICU = pediatric intensive care unit; PLR = positive likelihood ratio

[*]Positive and negative likelihood ratios calculated from sensitivity and specificity provided in original manuscript; [†]For the Walia et al. study the same cut-point of > 21 was associated with a PLR ≥ 5.0 and NLR ≤ 0.2.

S5 Appendix

Table 5a. Unadjusted likelihood ratios for prognostic factors judged to be of limited value (neither positive likelihood ratio ≥ 5.0 nor negative likelihood ratio ≤ 0.2 found in any study) to identify children at risk of progressing to severe febrile illness ('hard' outcomes).

Study	Cohort	Outcome	Prev.	Prognostic factor	Definition / Cut-off	PLR	95% CI	NLR	95% CI
Demographic									
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Age	< 12m	0.97	(0.27 – 3.52)	1.01	(0.81 – 1.25)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Age	1y to < 5y	0.69	(0.34 – 1.40)	1.33	(0.90 – 1.98)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Age	5y to < 12y	1.41	(0.69 – 2.87)	0.86	(0.58 – 1.27)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Age	12y to < 18y	1.76	(0.48 – 6.46)	0.93	(0.75 – 1.16)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Age	13y to 17y	0.57	(0.20 - 1.67)	1.11	(0.96 - 1.29)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Age	6y to < 13y	1.21	(0.71 - 2.06)	0.91	(0.68 - 1.22)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Age	12m to < 6y	1.15	(0.76 - 1.73)	0.90	(0.61 - 1.31)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Age	< 12m	0.53	(0.08 - 3.66)	1.04	(0.96 - 1.13)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Age	< 12m	1.16	(0.88 - 1.52)	0.93	(0.80 - 1.08)
Mtove 2011	Hospitalised	In-hospital mortality	5.0	Age	< 12m	1.47	(1.22 - 1.78)	0.81	(0.71 - 0.92)
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Age	< 12m	1.42	(1.19 - 1.70)	0.81	(0.71 - 0.93)
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Age	< 24m	1.20	(1.10 - 1.30)	0.63	(0.47 - 0.83)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Age	< 12m	1.47	(1.20 - 1.79)	0.90	(0.83 - 0.96)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Sex	Female	0.87	(0.47 -1.60)	1.12	(0.88 – 1.06)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Sex	Female	1.05	(0.67 - 1.64)	0.97	(0.68 - 1.37)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Sex	Female	0.97	(0.77 - 1.22)	1.02	(0.86 - 1.23)
Lowlaavar 2016	Hospitalised	In-hospital mortality	5.0	Sex	Female	1.02	(0.78 – 1.34)	0.98	(0.78 – 1.23)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Sex	Female	1.01	(0.89 - 1.15)	0.99	(0.89 - 1.10)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Sex	Female	1.04	(0.91 - 1.17)	0.97	(0.87 - 1.08)
Anthropometric									
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Weight	< 6kg	1.57	(0.92 - 2.68)	0.98	(0.96 - 1.01)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Weight	< 8kg	1.29	(1.06 - 1.58)	0.92	(0.86 - 0.99)
Historical									
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Medical history	Non-oncological comorbidity	1.06	(0.68 - 1.66)	0.96	(0.67 - 1.36)
Scott 2017	Hospital OPD/ED	30-day mortality	1.9	Medical history	Oncological comorbidity	1.73	(1.17 - 2.54)	0.69	(0.46 - 1.03)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	Medical history	Immunosuppressed	4.64	(1.79 - 12.00)	0.74	(0.52 - 1.07)

Clinical symptoms									
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Convulsions	Caretaker history	1.90	(1.41 - 2.54)	0.81	(0.70 - 0.93)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Convulsions	Caretaker history	1.39	(1.05 - 1.84)	0.95	(0.90 - 1.00)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Fever	Caretaker history	1.00	(0.99 - 1.01)	1.42	(0.32 - 6.25)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Urine looks dark	Caretaker history	1.04	(0.76 - 1.41)	0.99	(0.95 - 1.04)
Clinical signs									
van Nassau 2018	Hospitalised	PICU transfer and/or in-hospital mortality	2.7	Abnormal temperature	> 38.5°C or < 36°C	1.43	(0.93 - 2.20)	0.77	(0.50 - 1.17)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Hyperthermia	Temperature > 38°C	0.59	(0.42 - 0.82)	1.34	(1.18 - 1.51)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Hyperthermia	Axillary temperature > 37°C	0.72	(0.66 - 0.80)	2.24	(1.92 - 2.62)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Hyperthermia	Axillary temperature > 39°C	0.60	(0.45 - 0.79)	1.13	(1.07 - 1.19)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Hypothermia	Temperature < 36°C	3.16	(1.73 - 5.77)	0.92	(0.86 - 0.99)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Hypothermia	Axillary temperature < 36°C	3.47	(2.58 - 4.67)	0.87	(0.83 - 0.92)
van Nassau 2018	Hospitalised	PICU transfer and/or in-hospital mortality	2.7	Heart rate	Age-adjusted	1.98	(1.33 - 2.94)	0.63	(0.40 - 0.99)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Heart rate	Age-adjusted	0.92	(0.77 - 1.09)	1.15	(0.90 - 1.46)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Heart rate	≥ 200bpm	4.61	(1.99 - 10.67)	0.98	(0.96 - 1.00)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Heart rate	Age-adjusted	0.74	(0.66 - 0.82)	1.69	(1.47 - 1.93)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Capillary refill time	≥ 3s	4.67	(3.00 - 7.28)	0.83	(0.75 - 0.92)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	Capillary refill time	Flash or > 2s	0.50	(0.07 - 3.35)	1.09	(0.92 - 1.29)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Capillary refill time	> 2s	1.28	(1.21 - 1.35)	0.48	(0.37 - 0.62)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Capillary refill time	< 2s	0.50	(0.39 - 0.65)	1.26	(1.19 - 1.33)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Capillary refill time	≥ 3s	1.98	(1.73 - 2.27)	0.69	(0.62 - 0.77)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Capillary refill time	2-3s	0.84	(0.72 - 0.99)	1.11	(1.02 - 1.22)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	Poor peripheral perfusion	Cold extremity	4.35	(0.52 - 36.17)	0.94	(0.80 - 1.10)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Poor peripheral perfusion	Limb-core temp. gradient	1.34	(1.25 - 1.44)	0.54	(0.44 - 0.67)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Poor peripheral perfusion	Limb-core temp. gradient	1.32	(1.23 - 1.42)	0.57	(0.47 - 0.70)
van Nassau 2018	Hospitalised	PICU transfer and/or in-hospital mortality	2.7	Respiratory rate	Age-adjusted	1.11	(0.93 - 1.33)	0.62	(0.22 - 1.78)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Respiratory rate	Age-adjusted	2.05	(1.66 - 2.53)	0.66	(0.54 - 0.80)
Conroy 2015	Hospitalised	In-hospital mortality	4.7	Respiratory distress	Subcostal recession	3.76	(3.18 - 4.45)	0.41	(0.31 - 0.54)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Respiratory distress	Chest wall retraction	1.15	(1.07 - 1.23)	0.70	(0.56 - 0.86)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Respiratory distress	Increased work of breathing or deep breathing	1.13	(1.09 - 1.17)	0.43	(0.29 - 0.63)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Respiratory distress	Increased work of breathing or deep breathing	1.12	(1.08 - 1.16)	0.48	(0.34 - 0.69)

George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Respiratory crackles	Physician assessment	1.82	(1.54 - 2.14)	0.79	(0.72 - 0.86)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Respiratory crackles	Physician assessment	1.81	(1.53 - 2.13)	0.80	(0.73 - 0.87)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Meningeal	2.71	(0.17 – 43.95)	0.98	(0.89 – 1.08)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Respiratory	0.50	(0.14 – 1.81)	1.20	(0.97 – 1.49)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Gastrointestinal	0.56	(0.08 – 3.71)	1.07	(0.92 – 1.23)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Urinary	1.69	(0.11-26.71)	0.99	(0.90 – 1.08)
Kwizera 2019	Hospitalised	In-hospital mortality	1.5	Focus of infection	Skin and/or soft-tissue	3.28	(0.20 -53.88)	0.98	(0.89 – 1.07)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Cough	Physician assessment	1.00	(0.93 - 1.08)	1.00	(0.83 - 1.21)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	Agitation	Physician assessment	4.17	(2.08 - 8.35)	0.61	(0.37 - 1.00)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Seizures	Physician assessment	1.24	(0.96 - 1.61)	0.96	(0.91 - 1.01)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Normal consciousness	AVPU = Alert (A)	0.29	(0.19 - 0.44)	1.23	(1.18 - 1.28)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Vomiting	Physician assessment	1.14	(1.03 - 1.27)	0.85	(0.74 - 0.98)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Dehydration	Decreased skin turgor	2.83	(2.07 - 3.89)	0.90	(0.86 - 0.95)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Dehydration	Sunken eyes or reduced skin turgor	2.52	(1.89 - 3.36)	0.89	(0.85 - 0.94)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Jaundice	Physician assessment	1.39	(1.21 - 1.60)	0.82	(0.75 - 0.91)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Severe pallor	Physician assessment	1.47	(1.35 - 1.59)	0.56	(0.47 - 0.67)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Severe pallor	Physician assessment	1.53	(1.37 - 1.71)	0.70	(0.62 - 0.80)
Laboratory									
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Glucose	> 5mmol/L	0.56	(0.47 - 0.67)	2.33	(2.02 - 2.68)
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Glucose	2.5-5mmol/L	1.39	(1.11 - 1.75)	0.88	(0.80 - 0.98)
Mtove 2011	Hospitalised	In-hospital mortality	5.0	Haemoglobin	< 4g/dL	1.98	(1.53 - 2.56)	0.83	(0.76 - 0.92)
Nadjm 2013	Hospitalised	In-hospital mortality	5.1	Haemoglobin	< 5g/dL	1.93	(1.51 - 2.46)	0.83	(0.75 - 0.92)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Haemoglobin	< 5g/dL	1.43	(1.24 - 1.64)	0.81	(0.73 - 0.89)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Haemoglobin	5-7g/dL	0.90	(0.68 - 1.19)	1.02	(0.97 - 1.07)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Haemoglobin	7-10g/dL	0.97	(0.79 - 1.18)	1.01	(0.94 - 1.09)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Haemoglobin	≥ 10g/dL	0.56	(0.42 - 0.75)	1.14	(1.09 - 1.20)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Lactate	< 2.5mmol/L	0.26	(0.17 - 0.37)	1.35	(1.29 - 1.40)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Lactate	2.5-5mmol/L	0.45	(0.34 - 0.58)	1.28	(1.21 - 1.35)
van Nassau 2018	Hospitalised	PICU transfer and/or in-hospital mortality	2.7	Leukocyte count	High or low (age-adjusted)	0.97	(0.64 - 1.48)	1.03	(0.67 - 1.58)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	pH	< 7.2	4.85	(3.79 - 6.21)	0.70	(0.63 - 0.77)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	pH	< 7.2	4.43	(3.45 - 5.68)	0.72	(0.65 - 0.79)
George 2015	Hospitalised	In-hospital mortality (48h)	9.9	Urea	> 20mg/dL	2.50	(2.08 - 3.00)	0.67	(0.58 - 0.76)

Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Urea	> 20mg/dL	2.37	(1.97 - 2.84)	0.69	(0.61 - 0.78)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Base deficit	> -8mmol/L	1.74	(1.61 - 1.88)	0.31	(0.23 - 0.43)
Aramburo 2018	Hospitalised	In-hospital mortality (72h)	10.3	Bicarbonate	< 15mmol/L	2.25	(2.03 - 2.51)	0.40	(0.31 - 0.50)
Composite scores									
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	CRS*	≥ 1	2.02	(1.26 - 3.23)	0.55	(0.28 - 1.11)
Scott 2014	Hospital OPD/ED	24-hour organ dysfunction	5.4	CRS*	≥ 2	4.35	(1.40 - 13.52)	0.81	(0.60 - 1.10)
SEAIDCRN 2017	Hospitalised	28-day mortality	1.9	Severe sepsis†	Goldstein criteria	3.08	(2.28 - 4.16)	0.29	(0.11 - 0.79)

AVPU = alert, voice, pain or unresponsive; CI = confidence interval; CRS = Clinical Recognition Signs; ED = emergency department; MUAC = mid-upper arm circumference; NLR = negative likelihood ratio; OPD = outpatient department; PLR = positive likelihood ratio; Prev. = outcome prevalence (%); WAZ = weight-for-age z-score

*CRS scored out of four variables including mental status, capillary refill time, peripheral pulse character, and presence of cold or mottled extremities);[5] †Children with sepsis were enrolled based on modified Goldstein criteria (see Table 1 in main manuscript). Severe sepsis was defined based on Goldstein criteria for severe sepsis.[6]

Table 5b. Unadjusted likelihood ratios for prognostic factors judged to be of limited value (neither positive likelihood ratio ≥ 5.0 nor negative likelihood ratio ≤ 0.2 found in any study) to identify children at risk of progressing to severe febrile illness ('soft' outcomes).

Study	Cohort	Outcome	Prev.	Prognostic factor	Definition / Cut-off	PLR	95% CI	NLR	95% CI
Demographic									
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Sex	Female	0.87	(0.53 - 1.42)	1.16	(0.73 - 1.83)
Socioeconomic									
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Household socioeconomic status	Highest wealth quintile	1.24	(0.89 - 1.73)	0.71	(0.36 - 1.40)
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Household socioeconomic status	Slept under ITN night prior to enrolment	0.73	(0.49 - 1.07)	2.08	(1.13 - 3.82)
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Parental education	None	0.71	(0.10 - 5.20)	1.02	(0.91 - 1.15)
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Parental education	Primary	1.05	(0.82 - 1.34)	0.86	(0.34 - 2.13)
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Parental education	Secondary	0.92	(0.31 - 2.74)	1.02	(0.83 - 1.25)
Clinical symptoms									
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	URTI/cold presentation	Caretaker history	1.27	(0.81 - 2.00)	0.79	(0.46 - 1.35)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Sore throat	Caretaker history	2.21	(1.39 - 3.51)	0.82	(0.70 - 0.96)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Otalgia	Caretaker history	1.51	(0.92 - 2.47)	0.90	(0.78 - 1.05)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Otalgia	Earache resulting in altered reaction or sleeping pattern	1.77	(0.85 - 3.68)	0.94	(0.85 - 1.04)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Rhinorrhea	Caretaker history	1.19	(0.84 - 1.67)	0.91	(0.74 - 1.12)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Cough	Caretaker history	1.20	(0.86 - 1.67)	0.90	(0.73 - 1.11)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Moaning respiration	Caretaker history	1.27	(1.01 - 1.60)	0.77	(0.56 - 1.05)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Abdominal pain	Caretaker history	1.45	(0.87 - 2.42)	0.92	(0.80 - 1.05)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Diarrhea > 2/day	Caretaker history	1.45	(0.98 - 2.15)	0.87	(0.72 - 1.04)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Vomiting	Caretaker history	1.31	(0.95 - 1.81)	0.85	(0.68 - 1.07)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Febrile convulsions	Caretaker history	2.36	(1.04 - 5.35)	0.93	(0.85 - 1.02)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Drowsy or difficult to wake	Caretaker history	0.85	(0.63 - 1.15)	1.15	(0.91 - 1.46)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Restlessness or confusion	Caretaker history	1.09	(0.74 - 1.62)	0.96	(0.80 - 1.15)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Feeling irritable	Caretaker history	1.31	(0.97 - 1.78)	0.84	(0.66 - 1.06)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Inconsolable crying	Caretaker history	0.99	(0.75 - 1.32)	1.00	(0.79 - 1.28)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Crying during diaper change	Caretaker history	0.95	(0.63 - 1.45)	1.02	(0.86 - 1.21)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Crying when picked up	Caretaker history	0.86	(0.55 - 1.35)	1.06	(0.90 - 1.24)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Different illness than usual	Caretaker history	1.16	(0.94 - 1.44)	0.81	(0.59 - 1.13)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Parental concern	Caretaker history	1.51	(0.92 - 2.47)	0.90	(0.78 - 1.05)

Elshout 2015	Primary care	Persistent fever at D3	13.1	Drinking less than half usual	Caretaker history	1.07	(0.76 - 1.51)	0.96	(0.78 - 1.18)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Skin rash	Caretaker history	0.92	(0.54 - 1.59)	1.02	(0.90 - 1.16)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Pale, grey or spotted skin	Caretaker history	0.91	(0.69 - 1.21)	1.09	(0.85 - 1.39)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Normal play behaviour	Caretaker history	1.09	(0.89 - 1.34)	0.87	(0.62 - 1.24)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Normal reaction to parents	Caretaker history	0.70	(0.26 - 1.89)	1.03	(0.96 - 1.11)
Clinical signs									
Mwandama 2016	Primary care	Persistent symptoms at D7	10.4	Hyperthermia	Axillary temperature ≥ 37.5°C	1.64	(1.01 - 2.66)	0.67	(0.39 - 1.15)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Hyperthermia	Rectal temperature ≥ 38°C	1.47	(1.08 - 2.01)	0.80	(0.63 - 1.00)
van Nassau 2018	Hospitalised	Length of stay >/= 7d	22.2	Abnormal temperature	> 38.5°C or < 36°C	0.81	(0.63 - 1.04)	1.11	(0.99 - 1.24)
van Nassau 2018	Hospitalised	Length of stay >/= 7d	22.2	Heart rate	Age-adjusted	1.67	(1.18 - 2.37)	0.88	(0.80 - 0.97)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Capillary refill time	> 2s	0.98	(0.35 - 2.71)	1.00	(0.93 - 1.07)
van Nassau 2018	Hospitalised	Length of stay >/= 7d	22.2	Respiratory rate	Age-adjusted	0.99	(0.89 - 1.10)	1.04	(0.73 - 1.47)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Respiratory distress (dyspnoea)	Physician assessment	1.06	(0.71 - 1.58)	0.98	(0.82 - 1.16)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Pharyngitis	Sign of throat infection	1.64	(1.25 - 2.16)	0.70	(0.54 - 0.91)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Rhinorrhoea	Physician assessment	0.91	(0.70 - 1.18)	1.11	(0.85 - 1.44)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Cough	Physician assessment	1.28	(0.95 - 1.72)	0.84	(0.66 - 1.07)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Palpable lymph nodes	Physician assessment	1.39	(1.09 - 1.77)	0.73	(0.54 - 0.98)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Meningism	Able to put chin to chest	0.34	(0.05 - 2.47)	1.03	(0.99 - 1.07)
Elshout 2015	Primary care	Persistent fever at D3	13.1	Ill appearance	Physician assessment	1.32	(0.61 - 2.85)	0.97	(0.89 - 1.06)
Laboratory									
van Nassau 2018	Hospitalised	Length of stay >/= 7d	22.2	Leukocyte count	High or low (age-adjusted)	1.15	(0.96 - 1.36)	0.86	(0.70 - 1.06)
Freyne 2013	Hospitalised	Length of stay > 96h	26.1	Leukocyte count	> 15,000cells/mm	0.97	(0.44 - 2.15)	1.02	(0.57 - 1.82)
Freyne 2013	Hospitalised	Length of stay > 96h	26.1	Procalcitonin	> 1.0ng/L	1.00	(0.31 - 3.23)	1.00	(0.68 - 1.48)
Freyne 2013	Hospitalised	Length of stay > 96h	26.1	C-reactive protein	> 20mg/dL	1.27	(0.61 - 2.64)	0.82	(0.43 - 1.56)
Composite scores									
Freyne 2013	Hospitalised	Length of stay > 96h	26.1	AIOS*	> 10	1.00	(0.51 - 1.97)	1.00	(0.51 - 1.97)

AIOS = Acute Infantile Observation Score; CI = confidence interval; ITN = insecticide-treated bednet; NLR = negative likelihood ratio; PLR = positive likelihood ratio; Prev. = outcome prevalence (%)

*Acute Infantile Observation Score calculated as described for Yale Observation Score (YOS) in webappendix3 Table 3a.

S6 Appendix

Table 6a: Risk of bias and applicability assessments for the included clinical prediction model studies (n=7) using PROBAST (Prediction model Risk Of Bias ASsessment Tool). Each prediction model/outcome pair (n=32) is assessed independently.

Study	Clinical prediction model	Outcome	Risk of Bias					Applicability			
			Overall	Analysis	Outcome	Predictors	Participants	Overall	Outcome	Predictors	Participants
George 2015	FEAST-PET (D)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	FEAST-PETaL (D)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	LODS (D)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	PEDIA-i (V)	In-hospital mortality (<4h)	H	H	L	L	L	H	L	L	H
George 2015	PEDIA-e (V)	In-hospital mortality (4-48h)	H	H	L	L	L	H	L	L	H
George 2015	PEDIA-l (V)	In-hospital mortality (>48h)	H	H	L	L	L	H	L	L	H
George 2015	PRISM (V)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	PEWS (V)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
George 2015	AQUAMAT (V)	In-hospital mortality (48h)	H	H	L	L	L	H	L	L	H
Conroy 2015	LODS (V)	In-hospital mortality	H	H	L	L	L	H	L	L	H
Conroy 2015	SICK (V)	In-hospital mortality	H	H	L	L	L	H	L	L	H
Conroy 2015	PEDIA-i (V)	In-hospital mortality	H	H	L	L	L	H	L	L	H
Lowlaavar 2016	Model 1 (D)	In-hospital mortality	H	H	L	H	L	H	L	H	H
Lowlaavar 2016	Model 2 (D)	In-hospital mortality	H	H	L	H	L	H	L	H	H
Lowlaavar 2016	Model 3 (D)	In-hospital mortality	H	H	L	H	L	H	L	H	H
Walia 2016	YOS (V)	Mortality	H	H	L	L	U	H	L	L	H
Walia 2016	YOS (V)	Mechanical ventilation	H	H	L	L	U	H	L	L	H
van Nassau 2018	qSOFA (V)	PICU transfer and/or in-hospital mortality	H	H	L	L	L	H	L	L	H
van Nassau 2018	qPELOD-2 (V)	PICU transfer and/or in-hospital mortality	H	H	L	L	L	H	L	L	H
van Nassau 2018	SIRS (V)	PICU transfer and/or in-hospital mortality	H	H	L	L	L	H	L	L	H
van Nassau 2018	qSOFA-L (V)	PICU transfer and/or in-hospital mortality	H	H	L	L	L	H	L	L	H
van Nassau 2018	qSOFA (V)	Length of stay ≥ 7 days	H	H	L	L	L	H	L	L	H
van Nassau 2018	qPELOD-2 (V)	Length of stay ≥ 7 days	H	H	L	L	L	H	L	L	H
van Nassau 2018	SIRS (V)	Length of stay ≥ 7 days	H	H	L	L	L	H	L	L	H

van Nassau 2018	qSOFA-L (V)	Length of stay ≥ 7 days	H	H	L	L	L	H	L	L	H
Kwizera 2019	Model 1 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Kwizera 2019	Model 2 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Kwizera 2019	Model 3 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Kwizera 2019	Model 4 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Kwizera 2019	Model 5 (D)	In-hospital mortality	H	H	L	H	H	H	L	L	H
Scott 2020	Temporal (V)	Hypotensive shock ≤ 24h	H	H	H	L	H	L	L	L	L
Scott 2020	Geographic (V)	Hypotensive shock ≤ 24h	H	H	H	L	H	L	L	L	L

D = derivation; H = high risk/concern; L = low risk/concern; V = validation

Table 6b. Risk of bias and applicability assessments for included prognostic factor studies (n=11) using the QUIPS (Quality in Prognosis Studies) tool.

Study ID	Risk of Bias							Applicability					
	Overall	Analysis	Confounding	Outcome	Predictors	Attrition	Participants	Overall	Setting	Timing	Outcome	Predictors	Participants
Elshout 2015	H	M	M	H	M	H	H	H	L	L	H	L	H
Scott 2012	H	L	H	M	L	L	H	H	L	H	L	L	H
Scott 2014	H	L	H	H	L	L	L	H	L	L	L	L	H
Scott 2017	L	L	L	L	L	L	L	H	L	L	L	L	H
Freyne 2013	H	H	H	M	L	L	H	H	U	L	L	L	H
Mtove 2011	M	L	M	L	L	L	L	H	H	L	L	L	L
Nadjm 2013	M	M	M	L	L	L	L	H	H	L	L	L	L
Aramburo 2018	M	L	M	L	L	L	L	H	H	L	L	L	H
Costa 2017	H	H	H	L	H	L	H	H	U	U	L	L	H
Mwandama 2016	H	M	H	M	M	H	H	H	L	L	H	L	H
SEAIDCRN 2017	H	H	H	L	H	H	M	H	H	L	L	H	L

H = high risk/concern; L = low risk/concern; U = unclear risk/concern

S7 Appendix

Table 7. Alternate search strategy

	ORIGINAL MEDLINE SEARCH	ALTERNATE MEDLINE SEARCH
1	Fever[MeSH Terms] OR Fever[Title/Abstract] OR Febrile[Title/Abstract] OR "suspected sepsis"[Title/Abstract]	Fever[MeSH Terms] OR Fever[Title/Abstract] OR Febrile[Title/Abstract] OR "suspected sepsis"[Title/Abstract] OR Hypothermia[MeSH Terms] OR Hypothermia[Title/Abstract] OR "history of fever"[Title/Abstract]
2	pediatrics[MeSH Terms] OR pediatric*[Title/Abstract] OR paediatric*[Title/Abstract] OR child[MeSH Terms] OR child*[Title/Abstract] OR Infant[Mesh:NoExp] OR infant[Title/Abstract]	pediatrics[MeSH Terms] OR pediatric*[Title/Abstract] OR paediatric*[Title/Abstract] OR child[MeSH Terms] OR child*[Title/Abstract] OR Infant[Mesh:NoExp] OR infant[Title/Abstract]
3	(((((((Validat*[tw] OR Predict*[ti] OR Rule*[tw]) OR (Predict*[tw] AND (Outcome*[tw] OR Risk*[tw] OR Model*[tw])) OR ((History OR Variable*[tw] OR Criteria OR Scor*[tw] OR Characteristic*[tw] OR Finding*[tw] OR Factor*[tw]) AND (Predict*[tw] OR Model*[tw] OR Decision*[tw] OR Identif*[tw] OR Prognos*[tw])) OR (Decision*[tw] AND (Model*[tw] OR Clinical*[tw] OR "Logistic Models"[MeSH Terms])) OR (Prognostic AND (History OR Variable*[tw] OR Criteria OR Scor*[tw] OR Characteristic*[tw] OR Finding*[tw] OR Factor*[tw] OR Model*[tw]))))) OR ("Stratification" OR "ROC Curve"[MeSH Terms] OR "Discrimination" OR "Discriminate" OR "c-statistic" OR "c statistic" OR "Area under the curve" OR "AUC" OR "Calibration" OR "Indices" OR "Algorithm" OR "Multivariable")))))	(((((((Validat*[tw] OR Predict*[ti] OR Rule*[tw]) OR (Predict*[tw] AND (Outcome*[tw] OR Risk*[tw] OR Model*[tw])) OR ((History OR Variable*[tw] OR Criteria OR Scor*[tw] OR Characteristic*[tw] OR Finding*[tw] OR Factor*[tw]) AND (Predict*[tw] OR Model*[tw] OR Decision*[tw] OR Identif*[tw] OR Prognos*[tw])) OR (Decision*[tw] AND (Model*[tw] OR Clinical*[tw] OR "Logistic Models"[MeSH Terms])) OR (Prognostic AND (History OR Variable*[tw] OR Criteria OR Scor*[tw] OR Characteristic*[tw] OR Finding*[tw] OR Factor*[tw] OR Model*[tw]))))) OR ("Stratification" OR "ROC Curve"[MeSH Terms] OR "Discrimination" OR "Discriminate" OR "c-statistic" OR "c statistic" OR "Area under the curve" OR "AUC" OR "Calibration" OR "Indices" OR "Algorithm" OR "Multivariable")))))
4	death[MeSH Terms] OR death[Title/Abstract] OR mortality[MeSH Terms] OR mortality[Title/Abstract] OR systemic inflammatory response syndrome[MeSH Terms] OR "systemic inflammatory response syndrome"[Title/Abstract] OR SIRS[Title/Abstract] OR sepsis[Title/Abstract] OR septic*[Title/Abstract] OR "severe disease*"[Title/Abstract] OR "severe infection*"[Title/Abstract] OR "severe bacterial infection*"[Title/Abstract] OR "severe illness"[Title/Abstract] OR "severe febrile illness"[Title/Abstract] OR "serious disease*"[Title/Abstract] OR "serious infection*"[Title/Abstract] OR "serious bacterial infection*"[Title/Abstract] OR "serious illness"[Title/Abstract] OR "serious febrile illness"[Title/Abstract]	death[MeSH Terms] OR death[Title/Abstract] OR mortality[MeSH Terms] OR mortality[Title/Abstract] OR "severe disease*"[Title/Abstract] OR "severe infection*"[Title/Abstract] OR "severe bacterial infection*"[Title/Abstract] OR "severe illness"[Title/Abstract] OR "severe febrile illness"[Title/Abstract] OR "serious disease*"[Title/Abstract] OR "serious infection*"[Title/Abstract] OR "serious bacterial infection*"[Title/Abstract] OR "serious illness"[Title/Abstract] OR "serious febrile illness"[Title/Abstract]
5	1 AND 2 AND 3 AND 4	1 AND 2 AND 3 AND 4
6	("1999/05/31"[Date - Publication] : "2020/04/30"[Date – Publication])	("1999/05/31"[Date - Publication] : "2020/04/30"[Date – Publication])
7	5 AND 6	5 AND 6

Following suggestions arising during the peer review process we constructed an alternate search strategy which explicitly included the concept of 'hypothermia' and 'history of fever' in the first search string, and excluded the components of the third search string which were closely related to the concept of 'suspected sepsis'. This search retrieved 2,470 articles on MEDLINE, 280 of which had not been retrieved by our original search. The Venn diagram below illustrates the overlap in studies retrieved by the two search strategies.

Two authors (AC and RT) independently screened the 280 additional articles against the eligibility criteria used for the systematic review: 279 were excluded by screening of title and abstract; one article proceeded to full text review but was subsequently excluded as 85% (306/360) of the cohort were neonates and data disaggregated by age were not presented. Hence, this alternate search strategy did not identify any additional eligible articles.

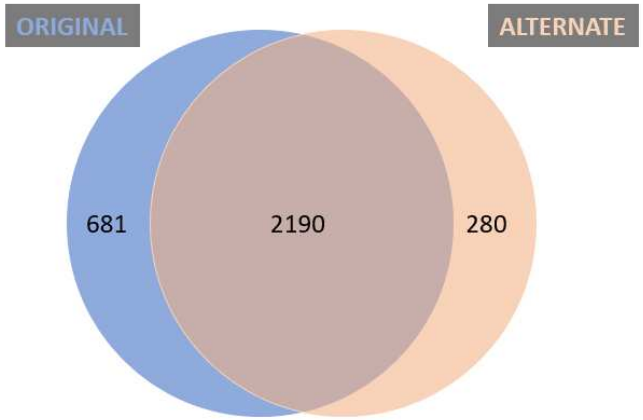

Figure 1. Venn diagram to illustrate the overlap in retrieved studies between the original and alternate search strategies.

References

1. Cochrane Prognosis Methods Group. Cochrane Methods Prognosis Tools. 2020. https://methods.cochrane.org/prognosis/tools (accessed 30 April 2020 2020).
2. Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. *PLoS One* 2012; **7**(2): e32844.
3. Haynes RB, McKibbon KA, Wilczynski NL, Walter SD, Werre SR. Optimal search strategies for retrieving scientifically strong studies of treatment from Medline: analytical survey. *Bmj* 2005; **330**(7501): 1179.
4. Ingui BJ, Rogers MA. Searching for clinical prediction rules in MEDLINE. *Journal of the American Medical Informatics Association : JAMIA* 2001; **8**(4): 391-7.
5. Scott HF, Donoghue AJ, Gaieski DF, Marchese RF, Mistry RD. Effectiveness of physical exam signs for early detection of critical illness in pediatric systemic inflammatory response syndrome. *BMC Emergency Medicine* 2014; **14**(1): 24.
6. Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatric critical care medicine : a journal of the Society of Critical Care Medicine and the World Federation of Pediatric Intensive and Critical Care Societies* 2005; **6**(1): 2-8.