BMJ Paediatrics Open

Identification of distinct clinical profiles of sepsis risk in paediatric emergency department patients using Bayesian profile regression

Patricia Gilholm ,[1] Sainath Raman,[1,2] Adam Irwin ,[2,3] Paula Lister,[4,5] Amanda Harley,[2,6] Luregn J Schlapbach,[1,7] Kristen S Gibbons[1]

To cite: Gilholm P, Raman S, Irwin A, *et al*. Identification of distinct clinical profiles of sepsis risk in paediatric emergency department patients using Bayesian profile regression. *BMJ Paediatrics Open* 2025;**9**:e003100. doi:10.1136/bmjpo-2024-003100

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjpo-2024-003100).

Received 4 October 2024
Accepted 24 February 2025

For numbered affiliations see end of article.

Correspondence to
Dr Patricia Gilholm; p.gilholm@uq.edu.au

ABSTRACT

Background Sepsis affects 25 million children and neonates annually, causing significant mortality and morbidity. Early identification and treatment are crucial for improving outcomes. Identifying children at risk is challenging due to clinical heterogeneity and overlap with other conditions. Current evaluations of sepsis criteria adopt a variable-centred approach, evaluating each criterion independently. The objective of this study was to explore associations between patterns of sepsis screening criteria and sepsis risk in children screened in the emergency department (ED) to identify distinct profiles that describe the clinical heterogeneity of suspected sepsis.

Methods This secondary analysis involved 3473 children screened for sepsis across 12 EDs in Queensland, Australia. Bayesian profile regression was used to construct data-driven clinical profiles derived from sepsis screening criteria and their association with suspected sepsis, defined as senior medical officer diagnosis and antibiotic administration in the ED. Posterior risk probabilities (Prs) with 95% credible intervals (CIs) were calculated for each profile. Profiles were internally validated by assessing their association with sepsis, septic shock, organ dysfunction and infection sources, in both adjusted and unadjusted models.

Results Seven distinct clinical profiles were identified. Two profiles were labelled as high risk of suspected sepsis (profile 1, n=22: Pr 0.73, 95% CI 0.55, 0.89; profile 2, n=150: Pr 0.69, 95% CI 0.59, 0.80), four as moderate risk and one as low risk. High-risk profiles were characterised by severe illness indicators and elevated lactate levels. Moderate-risk profiles included criteria such as altered behaviour, young age (<3 months) and respiratory distress. High-risk profiles had strong associations with all clinical outcomes.

Conclusions Seven clinical profiles were identified that varied in their risk of suspected sepsis and associated outcomes. Validation of these profiles in diverse populations and identification of which profiles are likely to benefit from certain interventions is needed.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Early recognition of sepsis remains a hurdle towards timely administration of sepsis treatment. The non-specific symptoms for sepsis in children can be difficult to recognise, especially in their early stages. Current evaluations of sepsis screening criteria for children are limited.

WHAT THIS STUDY ADDS

⇒ Distinct clinical profiles of children screened for sepsis in the emergency department were identified and internally validated. The profiles were characterised by different sepsis risk factors and moderate and severe illness features and varied in terms of their predicted risk of sepsis.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ This study highlights the potential for using data-driven clinical profiles to better stratify the risk of suspected sepsis in children, particularly in emergency settings. By identifying distinct risk profiles, it suggests the need for targeted monitoring and interventions, especially for children in moderate-risk groups who may require ongoing reassessment.

INTRODUCTION

Sepsis affects 25 million children and neonates per year and is a leading cause of childhood mortality and morbidity worldwide.[1] It is recognised that the early identification and treatment of children with sepsis or septic shock is critical for improving outcomes.[2] Protocol-based care has been found to decrease sepsis-related organ dysfunction, hospital and intensive care unit length of stay, and mortality.[3–6] To aid the timely recognition and treatment of sepsis, the 2020 Surviving Sepsis Campaign guidelines advocate for the use of systematic screening tools to assist in early recognition.[7 8] However, the non-specific symptoms for sepsis in children can be difficult to recognise, especially in their early stages. There is significant clinical heterogeneity in the physiological signs indicative of sepsis and considerable overlap in the presenting signs and symptoms with other

conditions, such as milder infections.[9] As such, identifying children who may benefit most from treatment remains challenging.[10 11]

Sepsis screening tools are increasingly being used in emergency departments (EDs) to improve the prompt recognition of paediatric sepsis and timely initiation of treatment. Paediatric sepsis screening tools comprise a range of vital signs or risk factors indicative of sepsis in children, which are either embedded as alerts into the electronic health record[11–13] or completed as a checklist by treating clinicians,[14 15] typically as part of a protocolised bundle of care.[5 16–18] Current evaluations of sepsis criteria adopt a variable-centred approach,[19] where the association between each criterion and sepsis is evaluated independently. Variable-centred approaches, such as multiple regression, assume that all individuals in a sample belong to a single population, where a set of averaged parameters is estimated.[20] In contrast, person-centred approaches assume that the sample may include multiple subpopulations characterised by different sets of parameters.[20] The identification of these subpopulations can help to understand different responses to treatment and support effective decision-making through stratification of care.[21 22]

Person-centred approaches, such as latent-class analysis and growth mixture modelling, have been used to identify clinical phenotypes of paediatric sepsis patients[23 24] and paediatric patients with acute respiratory distress syndrome.[25] However, these studies involved patients with confirmed diagnoses, and to our knowledge, this approach has not been employed to identify profiles of children screened for sepsis. Given the clinical heterogeneity in the signs and symptoms of paediatric sepsis, we used a person-centred approach to explore the association between patterns of sepsis screening criteria and sepsis risk in children screened for sepsis in the ED. We applied Bayesian profile regression[26] to construct clinical profiles based on covariate patterns of sepsis screening criteria and their association with clinician-diagnosed suspected sepsis.

METHODS

Cohort selection

This is a secondary analysis of a screening tool incorporated in the Queensland Paediatric Sepsis Pathway (PSP). The Queensland Paediatric Sepsis Collaborative was a state-wide quality improvement initiative evaluating the implementation of the PSP, designed to improve the recognition, escalation and management of paediatric sepsis.[16] Overall, 3473 children (<18 years old) were screened for sepsis at 12 participating EDs in Queensland during the evaluation period from 4 August 2018 to 31 December 2019.[15]

Exposure

The PSP screening tool contained 32 criteria associated with sepsis which were separated into 4 blocks of features corresponding to sepsis indicators (i.e., triggers to initiate screening), sepsis risk factors and features of severe and moderate illness.[14] Children were screened using the tool if the ED clinician suspected that the child may have sepsis or the child met one or more of the sepsis indicator criteria. Children were escalated to review by a senior medical officer (SMO) if any of the severe or moderate illness features were present.[14]

Outcomes

The primary outcome used to derive the clinical profiles was suspected sepsis, defined by two criteria: (1) the child was diagnosed with suspected sepsis by an SMO following screening on the PSP and (2) the child was treated with intravenous antibiotics in the ED.[16]

Secondary clinical outcomes of interest included sepsis, septic shock, organ dysfunction and infection source. Sepsis (children who met criteria for suspected sepsis and displayed signs of organ dysfunction) and septic shock (children who met criteria for suspected sepsis and displayed signs of cardiovascular organ dysfunction) were defined as per the 2005 International Pediatric Sepsis Definition Consensus Conference (online supplemental etable 1).[27] Organ dysfunction was recorded during the 24 hours following suspected sepsis diagnosis in the ED, including the presence of any organ dysfunction, as well as central nervous system (CNS) and cardiovascular system (CVS), renal, haematologic, respiratory and gastrointestinal organ dysfunction (online supplemental etable 1). The suspected source of infection was identified from the PSP or ED notes or retrieved from the medical record by research staff and categorised into sepsis with meningitis, sepsis (source unknown, excluding meningitis), pneumonia, intra-abdominal, urinary, cellulitis/skeletal/soft tissue, central venous access device, febrile neutropenia or other sources.

Patient and public involvement

This study is a secondary analysis of the screening tool incorporated in the Queensland PSP, and as such, there was no patient or public involvement in this specific work. However, the Queensland Sepsis Collaborative established a paediatric sepsis family group, which contributed to workshops on the screening tool design and implementation.

Statistical analysis

Descriptive statistics of the cohort are presented as counts and percentages for categorical variables and medians and IQR for continuous variables.

Covariates

The covariates used were the criteria assessed on the PSP screening tool. Previous work evaluating this screening tool identified 16 of the 32 criteria as being most strongly associated with suspected sepsis.[14] Therefore, this study uses only these 16 criteria as covariates in the derivation of the clinical profiles (table 1).

Table 1 Characteristics and distribution of sepsis screening criteria for each clinical profile

		Clinical profiles						
Characteristic	Total, **N=3473**	High risk 1 (severe features), **N=22**	High risk 2 (elevated lactate), **N=150**	Moderate risk 1 (altered behaviour), **N=111**	Moderate risk 2 (age <3 months), **N=299**	Moderate risk 3 (severe tachycardia), **N=420**	Moderate risk 4 (moderate features), **N=533**	Low risk, **N=1938**
Age (years)	2.1 (0.9, 5.6)	1.6 (0.6, 7.9)	1.0 (0.2, 5.1)	3.1 (1.4, 9.0)	0.1 (0.1, 0.2)	1.6 (1.0, 2.8)	2.9 (1.3, 6.8)	2.8 (1.3, 6.6)
Gender (male)	1903 (54.8%)	14 (63.6%)	86 (57.3%)	68 (61.2%)	166 (55.5%)	228 (54.3%)	276 (51.8%)	1065 (55.0%)
Sepsis screening criteria								
Sepsis indicators								
Healthcare worker concern	1198 (34.5%)	*16* (*72.7%*)	*111* (*74.0%*)	*71* (*64.0%*)	97 (32.4%)	228 (54.3%)	439 (82.4%)	236 (12.2%)
Looks sick	1596 (46.0%)	*20* (*90.9%*)	*114* (*76.0%*)	*81* (*73.0%*)	5 (1.7%)	*329* (*78.3%*)	*483* (*90.6%*)	564 (29.1%)
Altered behaviour	493 (14.2%)	*16* (*72.7%*)	24 (16.0%)	*93* (*83.8%*)	45 (15.1%)	77 (18.3%)	102 (19.1%)	136 (7.0%)
Sepsis risk factors								
Age <3 months	384 (11.1%)	2 (9.1%)	42 (28.0%)	10 (9.0%)	*299* (*100.0%*)	9 (2.1%)	6 (1.1%)	16 (0.8%)
Immunocompromised	100 (2.9%)	1 (4.5%)	9 (6.0%)	0	15 (5.0%)	15 (3.6%)	13 (2.4%)	47 (2.4%)
Chronic disease	141 (4.1%)	*6* (*27.2%*)	14 (9.3%)	17 (15.3%)	1 (0.3%)	31 (7.4%)	26 (4.9%)	46 (2.4%)
Severe illness features								
Need oxygen	195 (5.6%)	*22* (*100.0%*)	19 (12.7%)	11 (9.9%)	0	137 (32.6%)	0	6 (0.3%)
Severe respiratory distress	363 (10.4%)	*16* (*72.7%*)	27 (18.0%)	9 (8.1%)	36 (12.0%)	*234* (*55.7%*)	0	41 (2.1%)
Severe tachycardia	510 (14.7%)	*20* (*90.9%*)	57 (38.0%)	24 (21.6%)	37 (12.4%)	*244* (*58.1%*)	0	128 (6.6%)
Hypotension	30 (0.9%)	*5* (*22.7%*)	8 (5.3%)	6 (5.4%)	0	1 (0.2%)	1 (0.2%)	9 (0.5%)
Lactate ≥2 mmol/L	298 (8.6%)	*15* (*68.2%*)	*116* (*77.3%*)	28 (25.2%)	43 (14.4%)	31 (7.4%)	29 (5.4%)	36 (1.9%)
Altered AVPU	152 (4.4%)	*19* (*86.4%*)	2 (1.3%)	*107* (*96.4%*)	5 (1.7%)	18 (4.3%)	0	1 (0.1%)
Non-blanching rash	123 (3.5%)	3 (13.6%)	*30* (*20.0%*)	9 (8.1%)	9 (3.0%)	3 (0.7%)	13 (2%)	56 (2.9%)
Hypothermia	24 (0.7%)	3 (13.6%)	2 (1.3%)	5 (4.5%)	1 (0.3%)	2 (0.5%)	9 (2.4%)	2 (0.1%)
Moderate illness features								
Capillary refill ≥3 s	137 (3.9%)	0	0	0	*25* (*8.4%*)	0	*93* (*17.4%*)	19 (1.0%)
Cold extremities	93 (2.7%)	0	0	0	7 (2.3%)	0	*77* (*14.4%*)	9 (0.5%)

Sepsis screening criteria most prevalent in each profile are italicised.
AVPU, Alert, Verbal, Pain, Unresponsiveness Scale.

To assess the association between the clinical profiles and secondary outcomes, additional clinical control variables of gender, triage category (1: highest priority to 5: lowest priority), hospital category (quaternary hospital, dedicated paediatric hospital, mixed hospital) and admission time (midnight to 6:00, 6:oo to midday, midday to 18:00, 18:00 to midnight) were included.

Profile regression

Profile regression was used to cluster the sepsis screening criteria to identify clinical profiles (i.e., common covariate patterns of sepsis screening criteria). Profile regression is a Bayesian clustering algorithm which (1) uses a Bayesian mixture model framework to partition the observations based on similarity in covariate responses, (2) uses Markov chain Monte Carlo (MCMC) sampling methods to fit the model, which outputs a different partition (i.e., different number and composition of clinical profiles) at each iteration of the sampler, (3) links the profiles to the outcome via a regression model at each iteration of the sampler and (4) identifies the 'best' clustering solution (i.e., optimal number and composition of clinical profiles) and uses model averaging techniques to summarise the posterior distribution and quantify the uncertainty in the profiles.[26] Detailed descriptions of this method have been reported elsewhere.[26 28 29]

After obtaining the optimal profiles, we categorised them into three clinically meaningful categories of sepsis risk, based on their posterior mean probability of suspected sepsis. Profiles with a posterior mean probability of suspected sepsis between ≥10% and <50% were defined as moderate risk of sepsis, and groups with a posterior probability of suspected sepsis ≥50% were defined as high risk of sepsis. Profiles with a <10% posterior mean probability of suspected sepsis were categorised as low risk. These cut-offs were selected through consultation with clinical experts based on their experiences with identification and treatment of children with paediatric sepsis presenting to the ED.

To identify profiles that had higher or lower risk of suspected sepsis on average, as well as identify which specific sepsis criteria were most prevalent in each profile, we calculated centred parameters by subtracting the profile-specific probabilities from the overall mean probabilities across all profiles.[26] Furthermore, we examined the posterior distributions of the predicted risk for each profile by calculating the probability that the centred risk is greater than 0 (i.e., greater than the average risk of all profiles)[26] and by calculating the odds ratio (OR) and 95% credible intervals (CIs) between each profile and the reference profile, which we defined as the profile with the lowest predicted risk.[30] Further details of the model specification, prior distributions, model estimation and convergence, and model postprocessing are reported in online supplemental methods. All analyses were performed using the PReMiuM package[29] in R (V.4.1.1).[31]

Associations between clinical profiles and secondary clinical outcomes

We used Bayesian generalised linear models to assess the associations between each clinical profile and the secondary clinical outcomes of organ dysfunction and suspected source of infection. Models were only estimated for secondary outcomes with ≥5 cases in each profile. Log-binomial regression models were used to estimate the relative risk (RR) of each secondary outcome for each profile, alongside the 95% CIs. Clinical covariates were included as control variables. Models were fit in R using the Bayesian Applied Regression Modelling via Stan (rstanarm) package.[32] Default weakly informative priors, to provide moderate regularisation and help stabilise computation, were used.[32]

Missing data

As per previous publications using this cohort,[14 16] children who were missing data used to define the primary outcome of suspected sepsis (SMO review and/or antibiotic administration in the ED) were coded as 0 (i.e., no suspected sepsis). There were no missing data for the sepsis screening criteria or any clinical covariates except triage category, which was missing for <1% of cases. As a Bayesian modelling framework was used, imputation of these missing cases was not required.

Internal Validation

We internally validated the clinical profiles by executing the profile regression on two stratified random subsamples of the data (50/50 split, stratified by the original profile assignments). We fit the profile regression model to each subsample and selected the optimal number of clusters to be the same as the original analysis. We compared the resulting cluster compositions from the original analysis to each subsample by visual inspection of alluvial diagrams and through calculation of Cohen's kappa and the adjusted Rand index.[33–35]

To confirm that the clinical profiles were derived based on the underlying sepsis risk, we validated the profiles by evaluating the association between the clinical profiles and the sepsis and septic shock outcomes using Bayesian log-binomial regression models.

RESULTS

Cohort description

3473 children were screened for sepsis using the PSP. The median age of the children screened was 2.1 years (IQR 0.9–5.6) and 54.8% were male (table 1). 523 (15.1%) met the criteria for suspected sepsis (table 2). Of these, 210 (40.1%) children met criteria for sepsis, including 146 (27.9%) that met criteria for septic shock. The top three sources of infection for all children screened were pneumonia (n=268, 7.7%), sepsis with meningitis (n=243,

Table 2 Sepsis, organ dysfunction and focus of infection outcomes for each clinical symptom profile

		Clinical profiles						
Characteristic	Overall, **N=3473**	High risk 1 (severe features), **N=22**	High risk 2 (elevated lactate), **N=150**	Moderate risk 1 (altered behaviour), **N=111**	Moderate risk 2 (age <3 months), **N=299**	Moderate risk 3 (severe tachycardia), **N=420**	Moderate risk 4 (moderate features), **N=533**	Low risk, **N=1938**
Sepsis outcomes								
Suspected sepsis*	523 (15.1%)	19 (86.4%)	142 (94.7%)	44 (39.6%)	94 (31.4%)	93 (22.1%)	95 (17.8%)	36 (1.9%)
Sepsis	210 (40.2%)	18 (94.7%)	69 (48.6%)	33 (75.0%)	29 (30.9%)	29 (31.2%)	23 (24.2%)	9 (25.0%)
Septic shock	146 (27.9%)	12 (63.2%)	55 (38.7%)	15 (34.1%)	19 (20.2%)	15 (16.1%)	25 (26.3%)	5 (13.9%)
Organ dysfunction								
Any organ dysfunction†	518 (14.9%)	21 (95.5%)	71 (47.3%)	74 (66.7%)	62 (20.7%)	87 (20.7%)	69 (12.9%)	134 (6.9%)
CNS dysfunction	230 (44.4%)	16 (76.2%)	21 (29.6%)	61 (82.4%)	16 (25.8%)	44 (50.6%)	24 (34.8%)	48 (35.8%)
CVS dysfunction	260 (50.2%)	13 (61.9%)	54 (76.1%)	29 (39.2%)	30 (48.4%)	40 (46.0%)	32 (46.4%)	62 (46.3%)
Renal dysfunction	16 (3.1%)	1 (4.6%)	1 (1.4%)	3 (4.1%)	2 (3.2%)	2 (2.3%)	4 (5.8%)	3 (2.2%)
Haematological dysfunction	26 (5.0%)	0	5 (7.0%)	1 (1.4%)	5 (8.1%)	0	4 (5.8%)	11 (8.2%)
Respiratory dysfunction	48 (9.3%)	11 (52.4%)	7 9.9%)	6 (8.1%)	2 (3.2%)	15 (17.2%)	2 (2.9%)	5 (3.7%)
GIT dysfunction	55 (10.6%)	0	5 (7.0%)	2 (2.7%)	16 (25.8%)	5 (5.7%)	11 (15.9%)	16 (11.9%)
Focus of infection								
Sepsis with meningitis	243 (7.0%)	8 (36.4%)	43 (28.7%)	32 (28.8%)	68 (22.7%)	28 (6.7%)	25 (4.7%)	39 (2.0%)
Sepsis (source unknown, excluding meningitis)	189 (5.4%)	3 (13.6%)	35 (23.3%)	14 (12.6%)	32 (10.7%)	22 (5.2%)	43 (8.1%)	40 (2.1%)
Pneumonia	268 (7.7%)	7 (31.9%)	19 (12.7%)	11 (9.9%)	10 (3.3%)	101 (24.0%)	42 (7.9%)	78 (4.0%)
Intra-abdominal	57 (1.6%)	1 (4.5%)	4 (2.7%)	0	3 (1.0%)	5 (1.2%)	18 (3.4%)	26 (1.3%)
Urinary	138 (3.9%)	1 (4.5%)	16 (10.7%)	5 (4.5%)	20 (6.7%)	7 (1.7%)	28 (5.3%)	61 (3.1%)
Soft tissue/bone	103 (3.0%)	1 (4.5%)	7 (4.7%)	1 (0.9%)	1 (0.3%)	6 (1.4%)	30 (5.6%)	57 (2.9%)
CVAD	2 (0.1%)	0	0	0	0	0	0	2 (0.1%)
Febrile neutropenia	25 (0.7%)	0	1 (0.7%)	0	0	2 (0.5%)	7 (1.3%)	15 (0.7%)
Other	335 (9.6%)	2 (9.1%)	12 (8.0%)	17 (15.3%)	22 (7.4%)	43 (10.2%)	60 (11.3%)	179 (9.2%)

*Percentages of sepsis and septic shock are calculated out of the total with suspected sepsis.
†Percentages calculated out of the total with any organ dysfunction.
CNS, central nervous system; CVAD, Central Venous Access Device; CVS, cardiovascular system; GIT, gastrointestinal.

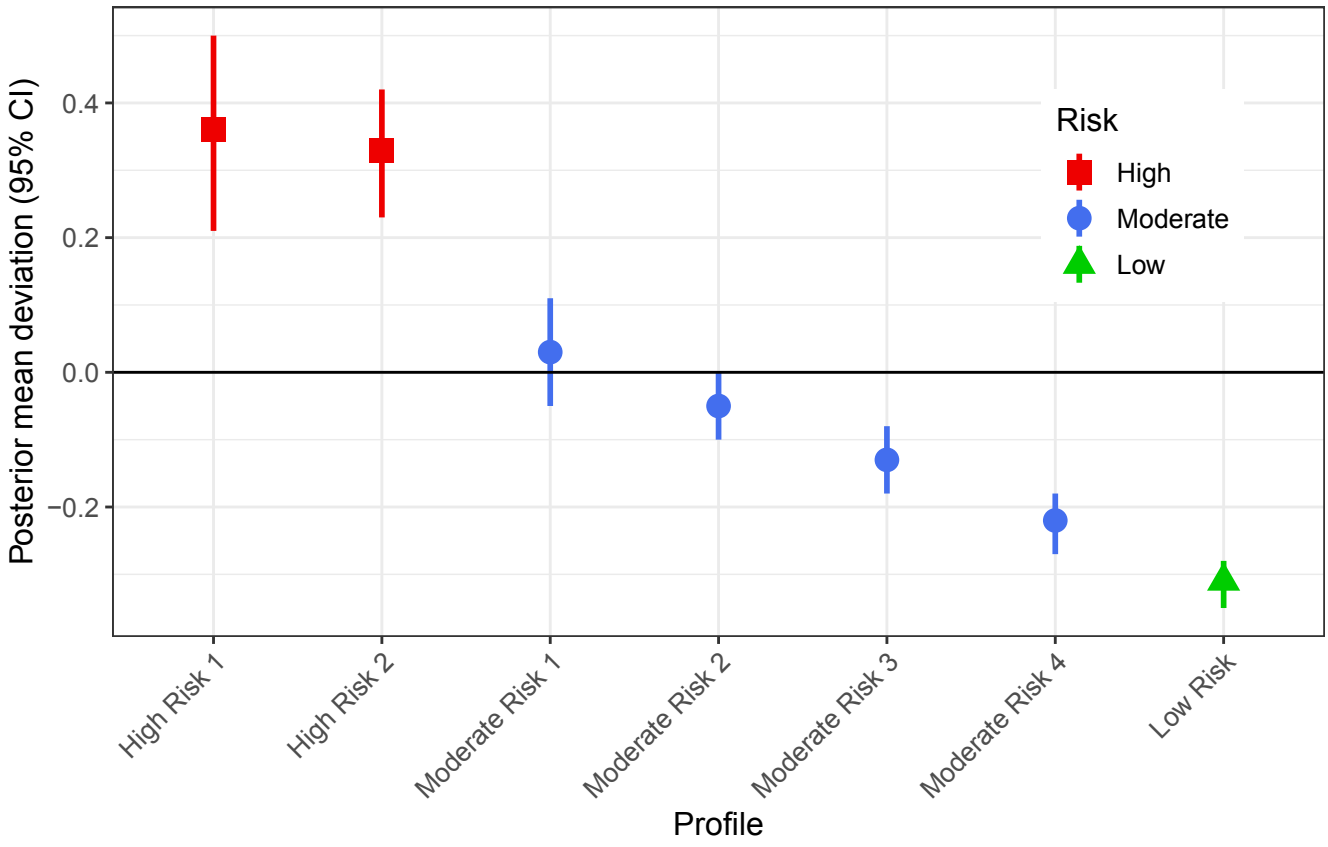

Figure 1 Posterior mean and 95% credible intervals of the centred risk for each profile. The line at zero indicates the average risk of sepsis across all profiles.

7.0%) and sepsis (source unknown, excluding meningitis) (n=189, 5.4%). Of the children screened, 518 (14.9%) had organ dysfunction, of which 308 (59.5%) did not meet criteria for suspected sepsis. The most common types of organ dysfunction were CVS (n=260, 50.1%) and CNS (n=230, 44.4%) dysfunction.

Description of profiles

The convergence statistics and procedures indicated that all parameters had sufficiently converged (online supplemental methods, efigure 1, etable 2). Postprocessing the MCMC chains revealed seven profiles, each with a differing predicted risk probability of suspected sepsis (table 2, figure 1) and characterised by a distinct subset of the sepsis screening criteria (table 1, figure 2, online supplemental efigure 2).

Two profiles had a high risk of sepsis (high risk 1 profile: risk probability (Pr) 0.73, 95% CI 0.55, 0.89; high risk 2 profile: Pr 0.69, 95% CI 0.59, 0.80) (table 3, figure 1). The high risk 1 profile was the smallest profile (n=22), representing less than 1% of the total cohort and included children with features of severe illness indicating likely organ dysfunction as well as multiple sepsis indicator and risk factor characteristics, including 'altered behaviour' and 'chronic disease' (table 1, figure 2). Four percent (n=150) of the cohort were assigned to high risk 2 profile, which was primarily characterised by elevated lactate and non-blanching rash (table 1, figure 2). Both these high-risk profiles were also characterised by two of the sepsis indicator criteria, 'healthcare worker concern' and 'looks sick'.

Four profiles had moderate risk of sepsis (table 3, figure 1) and were each characterised by a distinct profile of sepsis screening criteria (table 1, online supplemental efigure 2). Moderate risk 1 profile was characterised by altered behaviour and/or reduced level of consciousness. Moderate risk 2 profile was primarily composed of infants (age <3 months). Moderate risk 3 profile was characterised by children with severe respiratory distress and tachycardia, and moderate risk 4 profile had children with features of moderate illness (increased capillary refill time and cold extremities). All remaining children were allocated to the low-risk profile (online supplemental efigure 2), which was not characterised by any sepsis screening criteria and had the lowest predicted risk of sepsis.

Secondary outcomes

All the high-risk and moderate-risk profiles had higher risk for all organ dysfunction outcomes compared with the low-risk profile (table 4). For any organ dysfunction and cardiovascular dysfunction,

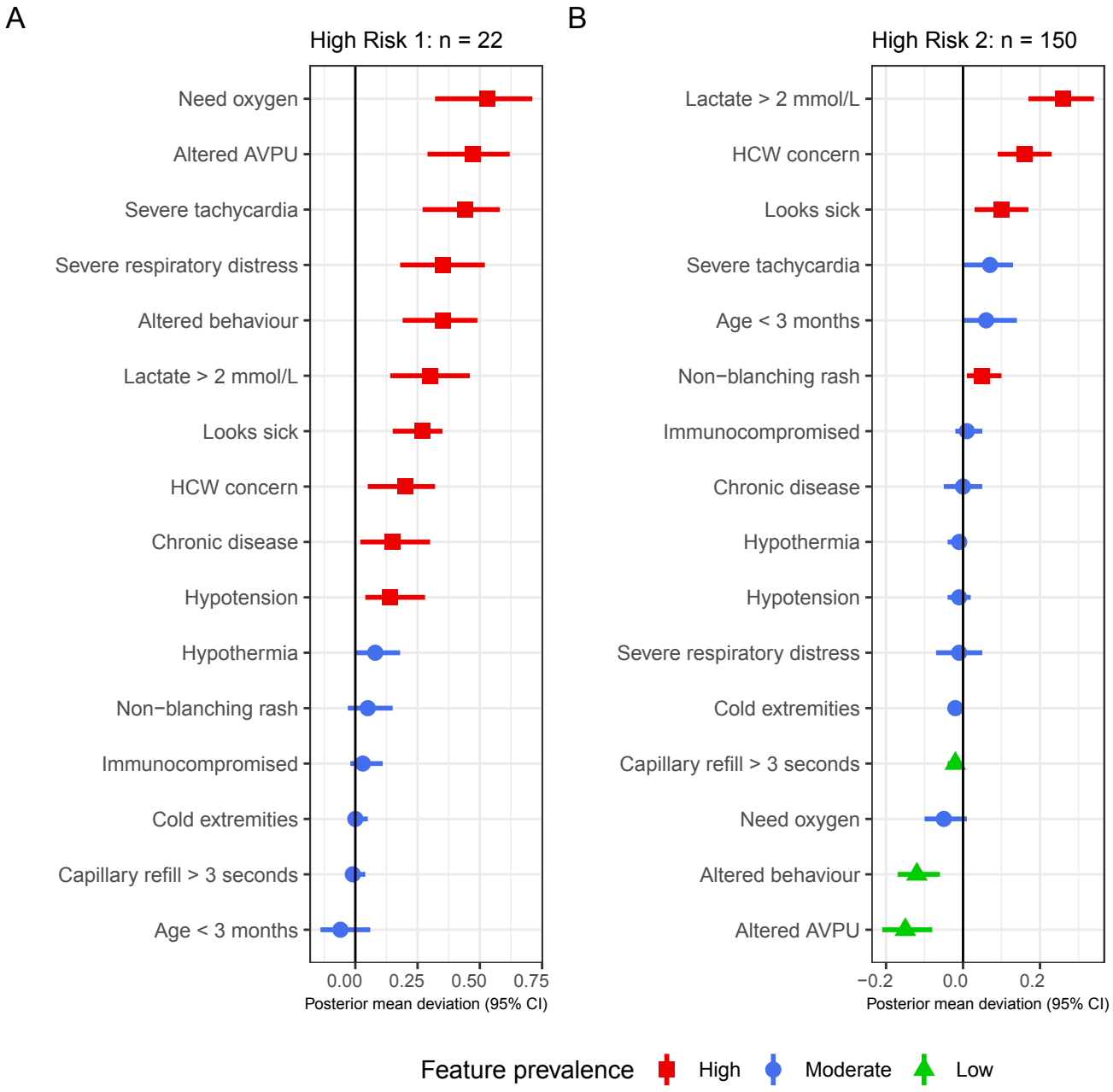

Figure 2 (A) Profile plot of the highest-risk profile (profile 1). (B) Profile plot of the high-risk profile (profile 2). For each feature, the posterior mean deviation and 95% credible interval (CI) from the average prevalence across all profiles (represented by the line at 0) is displayed. Features which are more prevalent within the subgroup compared with the average across all groups are presented in red, and features which are less prevalent within the subgroup compared with the average across all groups are presented in green. AVPU, Alert, Verbal, Pain, Unresponsive scale; HCW, healthcare worker.

the risk was greatest for the high-risk profiles and moderate risk 1 profile (characterised by altered behaviour). The highest risk of CNS dysfunction was exhibited by high risk 1 and moderate risk 1 profiles.

All high-risk and moderate-risk profiles had higher risk for sepsis with meningitis as the source of infection, compared with the low-risk profile (table 4), with both high-risk profiles and moderate risk 1 (altered behaviour) and 2 (age <3 months) profiles exhibiting the highest risk. For sepsis (source unknown, excluding meningitis) as the focus of infection, all high and moderate risk profiles had higher risk compared with the low-risk profile, with high-risk 2 profile, characterised by elevated lactate, displaying the highest risk. Finally, all high-risk and moderate-risk profiles had greater risk of pneumonia as the source of infection compared with the low-risk profile, except for moderate-risk 2 profile (age

Table 3 Predicted risk of suspected sepsis for each of the seven clinical profiles: profile size, posterior mean and 95% CIs of the uncentred risk, probability that the centred risk is greater than zero, and odds ratio of the predicted risk with the low-risk profile as the reference category

Profile	Profile size	Posterior mean (95% CI)	Pr ($p_c^* > 0$)	OR (95% CI)
High risk 1 (severe features)	22	0.73 (0.55, 0.89)	1.00	61.66 (15.15, 125.30)
High risk 2 (elevated lactate)	150	0.69 (0.59, 0.80)	1.00	46.60 (22.58, 75.14)
Moderate risk 1 (altered behaviour)	111	0.39 (0.31, 0.48)	0.78	12.96 (8.01, 18.37)
Moderate risk 2 (age <3 months)	299	0.31 (0.27, 0.36)	0.03	9.02 (6.42, 11.80)
Moderate risk 3 (severe tachycardia)	420	0.23 (0.19, 0.27)	0.00	5.97 (4.32, 7.75)
Moderate risk 4 (moderate features)	533	0.14 (0.11, 0.17)	0.00	3.17 (2.27, 4.12)
Low risk	1938	0.05 (0.04, 0.06)	0.00	1.00 (Reference)

Pr ($p_c^* > 0$): Probability that the centred risk is greater than zero.
CI, credible interval; OR, odds ratio.

<3 months). The high-risk 1 profile (severe illness features) and the moderate-risk 3 profile (severe tachycardia) displayed the highest risk.

Internal validation

The cluster memberships derived in the subsamples demonstrated moderate agreement to the original cluster memberships (Cohen's κ subsample 1=0.68, 95% CI (0.65, 0.70), adjusted Rand index 0.71; Cohen's κ subsample 2=0.71, 95% CI (0.68, 0.74), adjusted Rand index 0.75). Alluvial diagrams displaying the agreement between the cluster memberships show some consistency in the cluster assignments with some splitting and merging of clusters (online supplemental efigure 3 and 4). The original high-risk 1 profile, the smallest high-risk profile, was not represented in either subsample.

The evaluation of the association between the clinical profiles and the sepsis and septic shock outcomes confirmed that the clinical profiles were derived based on the underlying sepsis risk, whereby the highest risk profiles exhibited over 100-fold risk of sepsis and septic shock risk compared with the low-risk profile (online supplemental etable 3). For the moderate risk profiles, the risk of sepsis and septic shock aligned with the original risk probabilities of suspected sepsis, confirming the validity of the risk assignments.

DISCUSSION

Using Bayesian profile regression, we identified seven distinct clinical profiles of sepsis screening criteria in children screened for sepsis in the ED, each differing in their associated risk of suspected sepsis. Two profiles had a high risk of suspected sepsis: High risk 1 profile, with the smallest number of children assigned to it, was characterised by features of severe illness and linked to higher severity presentations of sepsis,[36–39] and high-risk 2, primarily marked by elevated lactate, which is known to predict sepsis in children.[40–42] Both high-risk profiles were strongly associated with organ dysfunction and infection outcomes. Four profiles had moderate risk for suspected sepsis, each with unique characteristics suggesting potential for deterioration. For example, the moderate risk 1 profile was associated with neurological dysfunction and linked to CNS organ dysfunction, while moderate risk 3, marked by severe respiratory distress and tachycardia, was strongly associated with pneumonia. The last profile had a low risk of suspected sepsis.

Bayesian profile regression has several strengths over other clustering methods and mixture model approaches. First, Bayesian profile regression is a semisupervised method which identifies representative profiles based not only on the similarity of the features (i.e., sepsis screening criteria), but also on their association with the primary outcome (i.e., suspected sepsis). Second, unlike other clustering algorithms and mixture models,[43 44] Bayesian profile regression is a non-parametric model that uses a Dirichlet process prior to model the number of clusters and, therefore, does not require the number of profiles to be prespecified.[26 45] Finally, by using a Bayesian framework, the uncertainty in the profile assignments can be explicitly quantified through Bayesian model averaging.

This cohort had low acuity and were all from the same well-resourced geographical area. As Bayesian profile regression is largely dependent on the data that is used to construct the profiles, external validation is needed to verify if the profiles observed using these data are represented in profiles derived from other cohorts, particularly from diverse settings. Previous research has found similar profiles in a very different population of children,[46] suggesting that some profiles may be stable across different populations. When internally validating the profiles through split-stratified sampling, we found moderate agreement between the clustering solutions. Although the highest-risk profile, with the smallest number of children assigned to it, was not represented in either subsample, most observations from this high-risk group merged with observations from the high-risk 2 profile and moderate-risk 1 profile in both subsamples

Table 4 Associations between clinical profiles and secondary outcomes of organ dysfunction type and sources of infection

Outcome	Clinical profile	Unadjusted RR (95% CI)	Adjusted RR (95% CI)*
Organ dysfunction outcomes			
Any organ dysfunction	High risk 1 (severe features)	13.26 (10.61, 16.20)	11.04 (8.81, 13.66)
	High risk 2 (elevated lactate)	6.79 (5.27, 8.61)	6.74 (5.34, 8.60)
	Moderate risk 1 (altered behaviour)	9.56 (7.66, 11.84)	9.44 (7.58, 11.67)
	Moderate risk 2 (age <3 months)	2.98 (2.23, 3.92)	2.96 (2.21, 3.88)
	Moderate risk 3 (severe tachycardia)	3.00 (2.31, 3.80)	2.98 (2.34, 3.83)
	Moderate risk 4 (moderate features)	1.85 (1.40, 2.44)	1.84 (1.39, 2.42)
	Low risk	1.00 (Reference)	1.00 (Reference)
Central nervous system dysfunction	High risk 1 (severe features)	28.34 (18.79, 41.00)	23.85 (15.85, 34.76)
	High risk 2 (elevated lactate)	5.62 (3.34, 8.96)	5.71 (3.46, 9.03)
	Moderate risk 1 (altered behaviour)	22.13 (16.28, 30.73)	22.22 (15.88, 31.19)
	Moderate risk 2 (age <3 months)	2.14 (1.16, 3.73)	2.13 (1.18, 3.59)
	Moderate risk 3 (severe tachycardia)	4.23 (2.84, 6.26)	4.30 (2.85, 6.38)
	Moderate risk 4 (moderate features)	1.80 (1.09, 2.88)	1.84 (1.09, 2.93)
	Low risk	1.00 (Reference)	1.00 (Reference)
Cardiovascular dysfunction	High risk 1 (severe features)	17.88 (10.68, 26.14)	16.41 (10.37, 23.85)
	High risk 2 (elevated lactate)	11.25 (8.14, 15.43)	10.80 (7.72, 15.00)
	Moderate risk 1 (altered behaviour)	8.09 (5.30, 12.14)	7.93 (5.15, 11.78)
	Moderate risk 2 (age <3 months)	3.12 (2.00, 4.70)	3.04 (1.93, 4.61)
	Moderate risk 3 (severe tachycardia)	2.98 (2.00, 4.32)	2.93 (2.00, 4.27)
	Moderate risk 4 (moderate features)	1.87 (1.21, 2.79)	1.81 (1.18, 2.72)
	Low risk	1.00 (Reference)	1.00 (Reference)
Focus of infection outcomes			
Sepsis with meningitis	High risk 1 (severe features)	17.26 (8.12, 30.11)	17.34 (8.25, 30.97)
	High risk 2 (elevated lactate)	14.13 (9.40, 20.94)	13.69 (9.17, 20.47)
	Moderate risk 1 (altered behaviour)	14.26 (9.19, 21.61)	13.68 (8.91, 21.14)
	Moderate risk 2 (age <3 months)	11.26 (7.85, 16.44)	11.22 (7.69, 16.38)
	Moderate risk 3 (severe tachycardia)	3.28 (2.04, 5.17)	3.28 (1.98, 5.16)
	Moderate risk 4 (moderate features)	2.31 (1.40, 3.77)	2.31 (1.36, 3.74)
	Low risk	1.00 (Reference)	1.00 (Reference)
Sepsis (source unknown, excluding meningitis)	High risk 1 (severe features)	5.94 (1.38, 15.98)	5.88 (1.34, 14.65)
	High risk 2 (elevated lactate)	11.32 (7.35, 17.59)	11.08 (7.20, 16.68)
	Moderate risk 1 (altered behaviour)	5.97 (3.13, 10.56)	5.64 (3.05, 9.87)

Continued

Table 4 Continued

Outcome	Clinical profile	Unadjusted RR (95% CI)	Adjusted RR (95% CI)*
	Moderate risk 2 (age <3 months)	5.17 (3.27, 8.10)	5.23 (3.34, 8.03)
	Moderate risk 3 (severe tachycardia)	2.51 (1,50, 4.25)	2.47 (1.45, 4.03)
	Moderate risk 4 (moderate features)	3.91 (2.58, 6.00)	3.76 (2.42, 5.80)
	Low risk	1.00 (Reference)	1.00 (Reference)
Pneumonia	High risk 1 (severe features)	7.50 (3.48, 13.51)	7.14 (3.44, 12.46)
	High risk 2 (elevated lactate)	3.08 (1.86, 4.78)	3.15 (1.93, 4.92)
	Moderate risk 1 (altered behaviour)	2.42 (1.24, 4.23)	2.33 (1.21, 4.07)
	Moderate risk 2 (age <3 months)	0.81 (0.39, 1.46)	0.84 (0.40, 1.51)
	Moderate risk 3 (severe tachycardia)	5.99 (4.54, 7.81)	6.00 (4.55, 7.94)
	Moderate risk 4 (moderate features)	1.94 (1.32, 2.81)	1.97 (1.37, 2.82)
	Low risk	1.00 (Reference)	1.00 (Reference)

*Adjusted for gender, triage category, site type (quaternary, specialised, mixed) and admission time (midnight–6:00, 6:00–midday, midday–18:00, 18:00–midnight).
CI, credible interval; RR, risk ratio.

(both high severity groups with high risk of suspected sepsis). This demonstrates that although the profiles may not be exactly reproducible, they still may be generalisable in terms of their sepsis risk.

This exploratory study used clustering methods to identify data-driven clinical symptom profiles, which may not be generalisable to other cohorts and are not intended for clinically diagnostic use. The primary outcome for deriving the profiles was clinician-diagnosed suspected sepsis, based on SMO judgement that 'sepsis is likely', which may introduce misclassification bias,[13] but reflects real clinical practice. Internal validation showed consistent risk probabilities between the clinical profiles and the sepsis and septic shock outcomes. Predefined cut-offs were used to classify profiles into low, moderate or high sepsis risk, with a predicted risk >50% for high-risk and 10%–50% for moderate-risk profiles. These cut-offs, developed in consultation with clinicians, may need further refinement. Previous research on sepsis subgroups has mostly focused on adult cohorts with a confirmed sepsis diagnosis,[47] limiting our ability to compare to existing derived subgroups. However, our two high-risk profiles share similarities with groups identified in the reanalysis of a landmark paediatric fluid trial examining fluid bolus in resuscitation among febrile children in Africa (FEAST trial),[46] which used Bayesian clustering to identify groups marked by severe respiratory and neurological dysfunction and elevated lactate, both associated with higher mortality.

Our study aimed to explain the clinical heterogeneity of children screened for sepsis in the ED, irrespective of diagnosis. This approach allowed us to identify the most common patterns of presenting signs and symptoms of children screened for sepsis in the ED and the associated risk of sepsis of each of those common patterns. The moderate-risk profiles that were uncovered reflect the clinical heterogeneity of paediatric sepsis, where although children at a high risk or low risk of sepsis can easily be identified through their presenting signs and symptoms, for the moderate-risk profiles, there is lower certainty for ruling out sepsis and, therefore, rescreening or ongoing monitoring of these children may be required.

This study contributes to the development of clinical profiles in children screened for sepsis in the ED. Using Bayesian profile analysis, we have identified seven distinct profiles which differ in terms of their predicted risk of suspected sepsis, as well as their association with organ dysfunction and source of infection outcomes. Future research should focus on the stability of these profiles, as well as identify which profiles are more or less likely to benefit from certain interventions. In addition, combining these clinical profiles with biological and genetic markers would provide a holistic perspective into the mechanisms associated with the development of sepsis in children.

Author affiliations
[1]Children's Intensive Care Research Program, Child Health Research Centre, The University of Queensland, South Brisbane, Queensland, Australia
[2]Queensland Children's Hospital, South Brisbane, Queensland, Australia
[3]University Of Queensland Centre for Clinical Research, Herston, Queensland, Australia

[4]Paediatric Critical Care Unit, Sunshine Coast University Hospital, Sunshine Coast, Queensland, Australia
[5]School of Medicine, Griffith University, Brisbane, Queensland, Australia
[6]School of Nursing, Midwifery, and Social Work, The University of Queensland, Brisbane, Queensland, Australia
[7]Department of Intensive Care and Neonatology and Children's Research Center, University Children's Hospital Zürich, Zurich, Switzerland

X Patricia Gilholm @TrishGilholm

Acknowledgements The authors would like to thank Queensland Health, the Queensland Paediatric Sepsis Program and Clinical Excellence Queensland for supporting this work, as well as the participating Emergency Department clinicians, and the families and children involved.

Contributors KSG conceived the study and supervised analyses; PG designed the study, performed all statistical analyses and drafted the manuscript; SR, AI, PL, AH and LJS acquired the data and critically interpreted the results. All authors critically reviewed the manuscript and approve the final version to be published. PG is responsible for the overall content as guarantor.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests All authors have completed the ICMJE uniform disclosure form at http://www.icmje.org/disclosure-of-interest/ and declare: no support from any organisation for the submitted work; AI and KSG have received research fellowships from the National Health and Medical Research Council, LJS has received support from the Thomas & Doris Ammann Foundation and the NOMIS Foundation. AI has received payment from Biomerieux and Gilead for lectures and presentations. All authors declare no other relationships or activities that could appear to have influenced the submitted work.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Ethical approval and waiver of individual consent was granted by the Children's Health Queensland Human Research Ethics Committee (HREC/18/QRCH/167).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. Data inquiries should be addressed to the corresponding author. The data are not publicly available due to privacy and ethical restrictions. Approval of data release will be governed by the Queensland Paediatric Sepsis Program, The University of Queensland and relevant regulatory bodies.

ORCID iDs
Patricia Gilholm http://orcid.org/0000-0001-8135-1520
Adam Irwin http://orcid.org/0000-0001-8974-6789

REFERENCES

1 Rudd KE, Johnson SC, Agesa KM, *et al*. Global, regional, and national sepsis incidence and mortality, 1990-2017: analysis for the Global Burden of Disease Study. *Lancet* 2020;395:200–11.
2 Idrees M, Macdonald SP, Kodali K. Sepsis Early Alert Tool: Early recognition and timely management in the emergency department. *Emerg Med Australas* 2016;28:399–403.
3 Paul R, Melendez E, Stack A, *et al*. Improving adherence to PALS septic shock guidelines. *Pediatrics* 2014;133:e1358–66.
4 Balamuth F, Weiss SL, Fitzgerald JC, *et al*. Protocolized Treatment Is Associated With Decreased Organ Dysfunction in Pediatric Severe Sepsis. *Pediatr Crit Care Med* 2016;17:817–22.
5 Evans IVR, Phillips GS, Alpern ER, *et al*. Association Between the New York Sepsis Care Mandate and In-Hospital Mortality for Pediatric Sepsis. *JAMA* 2018;320:358–67.
6 Paul R, Neuman MI, Monuteaux MC, *et al*. Adherence to PALS Sepsis Guidelines and Hospital Length of Stay. *Pediatrics* 2012;130:e273–80.
7 Hilarius KWE, Skippen PW, Kissoon N. Early Recognition and Emergency Treatment of Sepsis and Septic Shock in Children. *Pediatr Emerg Care* 2020;36:101–6.
8 Weiss SL, Peters MJ, Alhazzani W, *et al*. Surviving sepsis campaign international guidelines for the management of septic shock and sepsis-associated organ dysfunction in children. *Intensive Care Med* 2020;46:10–67.
9 Harley A, Schlapbach LJ, Johnston ANB, *et al*. Challenges in the recognition and management of paediatric sepsis - The journey. *Australas Emerg Care* 2022;25:23–9.
10 Balamuth F, Schlapbach LJ. Paediatric patient stratification in the emergency department. *Lancet Child Adolesc Health* 2020;4:557–8.
11 Balamuth F, Alpern ER, Abbadessa MK, *et al*. Improving Recognition of Pediatric Severe Sepsis in the Emergency Department: Contributions of a Vital Sign-Based Electronic Alert and Bedside Clinician Identification. *Ann Emerg Med* 2017;70:759–68.
12 Sepanski RJ, Godambe SA, Mangum CD, *et al*. Designing a pediatric severe sepsis screening tool. *Front Pediatr* 2014;2:56.
13 Balamuth F, Alpern ER, Grundmeier RW, *et al*. Comparison of Two Sepsis Recognition Methods in a Pediatric Emergency Department. *Acad Emerg Med* 2015;22:1298–306.
14 Gilholm P, Gibbons K, Lister P, *et al*. Validation of a paediatric sepsis screening tool to identify children with sepsis in the emergency department: a statewide prospective cohort study in Queensland, Australia. *BMJ Open* 2023;13:e061431.
15 Bradshaw C, Goodman I, Rosenberg R, *et al*. Implementation of an Inpatient Pediatric Sepsis Identification Pathway. *Pediatrics* 2016;137:e20144082.
16 Harley A, Lister P, Gilholm P, *et al*. Queensland Pediatric Sepsis Breakthrough Collaborative: Multicenter Observational Study to Evaluate the Implementation of a Pediatric Sepsis Pathway Within the Emergency Department. *Crit Care Explor* 2021;3:e0573.
17 Paul R, Melendez E, Wathen B, *et al*. A Quality Improvement Collaborative for Pediatric Sepsis: Lessons Learned. *Pediatr Qual Saf* 2018;3:e051.
18 Lane RD, Funai T, Reeder R, *et al*. High Reliability Pediatric Septic Shock Quality Improvement Initiative and Decreasing Mortality. *Pediatrics* 2016;138:2015–4153.
19 Menon K, Schlapbach LJ, Akech S, *et al*. Criteria for Pediatric Sepsis-A Systematic Review and Meta-Analysis by the Pediatric Sepsis Definition Taskforce. *Crit Care Med* 2022;50:21–36.
20 Morin AJS, Bujacz A, Gagné M. Person-Centered Methodologies in the Organizational Sciences. *Organ Res Methods* 2018;21:803–13.
21 Saunders R, Buckman JEJ, Pilling S. Latent variable mixture modelling and individual treatment prediction. *Behav Res Ther* 2020;124:103505.
22 Saunders R, Cape J, Fearon P, *et al*. Predicting treatment outcome in psychological treatment services by identifying latent profiles of patients. *J Affect Disord* 2016;197:107–15.
23 Wösten-van Asperen RM, la Roi-Teeuw HM, van Amstel RB, *et al*. Distinct clinical phenotypes in paediatric cancer patients with sepsis are associated with different outcomes-an international multicentre retrospective study. *EClinicalMedicine* 2023;65:102252.
24 Yehya N, Fitzgerald JC, Hayes K, *et al*. Temperature Trajectory Sub-phenotypes and the Immuno-Inflammatory Response in Pediatric Sepsis. *Shock* 2022;57:645–51.
25 Dahmer MK, Yang G, Zhang M, *et al*. Identification of phenotypes in paediatric patients with acute respiratory distress syndrome: a latent class analysis. *Lancet Respir Med* 2022;10:289–97.
26 Molitor J, Papathomas M, Jerrett M, *et al*. Bayesian profile regression with an application to the National Survey of Children's Health. *Biostatistics* 2010;11:484–98.
27 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.
28 Hastie DI, Liverani S, Azizi L, *et al*. A semi-parametric approach to estimate risk functions associated with multi-dimensional exposure profiles: application to smoking and lung cancer. *BMC Med Res Methodol* 2013;13:1–13.

29 Liverani S, Hastie DI, Azizi L, *et al*. PReMiuM: An R Package for Profile Regression Mixture Models Using Dirichlet Processes. *J Stat Softw* 2015;64:1–30.
30 Papathomas M, Molitor J, Richardson S, *et al*. Examining the joint effect of multiple risk factors using exposure risk profiles: lung cancer in nonsmokers. *Environ Health Perspect* 2011;119:84–91.
31 The r project for statistical computing. R Foundation for Statistical Computing, 2021. Available: https://www.R-project.org/
32 rstanarm: Bayesian applied regression modeling via Stan. Version R package version 2.32.1, 2024. Available: https://mc-stan.org/rstanarm/
33 Petersen JM, Parker SE, Dukes KA, *et al*. Machine learning-based placental clusters and their associations with adverse pregnancy outcomes. *Paediatr Perinat Epidemiol* 2023;37:350–61.
34 Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutor Quant Methods Psychol* 2012;8:23–34.
35 Hubert L, Arabie P. Comparing partitions. *J Classif* 1985;2:193–218.
36 Romaine ST, Sefton G, Lim E, *et al*. Performance of seven different paediatric early warning scores to predict critical care admission in febrile children presenting to the emergency department: a retrospective cohort study. *BMJ Open* 2021;11:e044091.
37 Chapman SM, Wray J, Oulton K, *et al*. "The Score Matters": wide variations in predictive performance of 18 paediatric track and trigger systems. *Arch Dis Child* 2017;102:487–95.
38 Romaine ST, Potter J, Khanijau A, *et al*. Accuracy of a Modified qSOFA Score for Predicting Critical Care Admission in Febrile Children. *Pediatrics* 2020;146:e20200782.
39 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.
40 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.
41 Schlapbach LJ, MacLaren G, Festa M, *et al*. Prediction of pediatric sepsis mortality within 1 h of intensive care admission. *Intensive Care Med* 2017;43:1085–96.
42 Schlapbach LJ, MacLaren G, Straney L. Venous vs Arterial Lactate and 30-Day Mortality in Pediatric Sepsis. *JAMA Pediatr* 2017;171:813.
43 McLachlan GJ, Rathnayake S. On the number of components in a Gaussian mixture model. *WIREs Data Min & Knowl* 2014;4:341–55.
44 Patel P, Sivaiah B, Patel R. Approaches for finding optimal number of clusters using k-means and agglomerative hierarchical clustering techniques. 2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSP); 2022
45 Li Y, Schofield E, Gönen M. A tutorial on Dirichlet Process mixture modeling. *J Math Psychol* 2019;91:128–44.
46 Levin M, Cunnington AJ, Wilson C, *et al*. Effects of saline or albumin fluid bolus in resuscitation: evidence from re-analysis of the FEAST trial. *Lancet Respir Med* 2019;7:581–93.
47 Seymour CW, Kennedy JN, Wang S, *et al*. Derivation, Validation, and Potential Treatment Implications of Novel Clinical Phenotypes for Sepsis. *JAMA* 2019;321:2003–17.