Original Article

Effective diagnosis of sepsis in critically ill children using probabilistic graphical model

Tuong Minh Nguyen[1]^, Kim Leng Poh[1], Shu-Ling Chong[2,3], Jan Hau Lee[3,4]

[1]Department of Industrial Engineering and Management, National University of Singapore, Singapore, Singapore; [2]Children's Emergency, KK Women's and Children's Hospital, Singapore, Singapore; [3]Singhealth-Duke NUS Paediatrics Academic Clinical Programme, Duke-NUS Medical School, Singapore, Singapore; [4]Children's Intensive Care Unit, KK Women's and Children's Hospital, Singapore, Singapore

Contributions: (I) Conception and design: All authors; (II) Administrative support: None; (III) Provision of study materials or patients: None; (IV) Collection and assembly of data: All authors; (V) Data analysis and interpretation: All authors; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Tuong Minh Nguyen, MS. NUS Faculty of Engineering, Department of Industrial Systems Engineering and Management, 1 Engineering Drive 2, Blk E1A #06-25, Singapore 117576, Singapore. Email: minh.t.nguyen@u.nus.edu.

Background: Probabilistic graphical model, a rich graphical framework in modelling associations between variables in complex domains, can be utilized to aid clinical diagnosis. However, its application in pediatric sepsis remains limited. This study aims to explore the utility of probabilistic graphical models in pediatric sepsis in the pediatric intensive care unit.

Methods: We conducted a retrospective study on children using the first 24-hour clinical data of the intensive care unit admission from the Pediatric Intensive Care Dataset, 2010–2019. A probabilistic graphical model method, Tree Augmented Naive Bayes, was used to build diagnosis models using combinations of four categories: vital signs, clinical symptoms, laboratory, and microbiological tests. Variables were reviewed and selected by clinicians. Sepsis cases were identified with the discharged diagnosis of sepsis or suspected infection with the systemic inflammatory response syndrome. Performance was measured by the average sensitivity, specificity, accuracy, and area under the curve of ten-fold cross-validations.

Results: We extracted 3,014 admissions [median age of 1.13 (interquartile range: 0.15–4.30) years old]. There were 134 (4.4%) and 2,880 (95.6%) sepsis and non-sepsis patients, respectively. All diagnosis models had high accuracy (0.92–0.96), specificity (0.95–0.99), and area under the curve (0.77–0.87). Sensitivity varied with different combinations of variables. The model that combined all four categories yielded the best performance [accuracy: 0.93 (95% confidence interval (CI): 0.916–0.936); sensitivity: 0.46 (95% CI: 0.376–0.550), specificity: 0.95 (95% CI: 0.940–0.956), area under the curve: 0.87 (95% CI: 0.826–0.906)]. Microbiological tests had low sensitivity (<0.10) with high incidence of negative results (67.2%).

Conclusions: We demonstrated that the probabilistic graphical model is a feasible diagnostic tool for pediatric sepsis. Future studies using different datasets should be conducted to assess its utility to aid clinicians in the diagnosis of sepsis.

Keywords: Pediatric sepsis; probabilistic graphical model; tree augmented Naïve Bayes

Submitted Oct 11, 2022. Accepted for publication Feb 26, 2023. Published online Apr 04, 2023.
doi: 10.21037/tp-22-510
View this article at: https://dx.doi.org/10.21037/tp-22-510

^ ORCID: 0000-0002-8809-4800.

Introduction

Sepsis in critically ill children places them at risk of death and long-term morbidity (1,2). Early detection of sepsis allows for prompt treatment, while the early distinction of a child who is not in sepsis enables the option of timely de-escalation of antibiotic administration. However, these tasks remain challenging for clinicians in the pediatric intensive care unit (PICU) due to the complex nature and heterogeneity of pediatric sepsis. Probabilistic graphical models (PGM) provide an inferencing framework that can aid clinicians at the bedside. It uses graphs (nodes and edges) to model conditional dependencies between variables in complex domains and produce robust predictions (3). With graphical representation, PGM is more comprehensive and interpretable than other black-box machine learning methodologies (e.g., Neural Network, Support Vector Machine) (4). Bayesian Network (BN), a subclass of PGM, has been proven effective in disease diagnosis; adult sepsis is one of them (5-8). However, studies applying PGM in pediatric sepsis remain limited.

Our hypothesis for this preliminary study was that PGM is a robust diagnostic tool for pediatric sepsis. We employed Tree Augmented Naive Bayes (TAN), a PGM method, to develop diagnosis models to test our hypothesis and investigate the effectiveness of PGM in pediatric sepsis diagnosis. We present the following article in accordance with the TRIPOD reporting checklist (available at https://tp.amegroups.com/article/view/10.21037/tp-22-510/rc).

Highlight box

Key findings

- Probabilistic graphical model (PGM) is a feasible diagnostic tool for pediatric sepsis in critically ill children.

What is known and what is new?

- Most children receive antibiotics when there is a suspicion of infection. An early distinction between those who do and do not need antibiotics will help rationalize drug use and reduce drug resistance.
- PGM is an explainable machine learning methodology with graphical interfaces that are capable of both inference and prediction. It has been utilized for disease detection, image processing, and pattern recognizing. The use of PGM with methods such as Tree Augmented Naïve Bayesian Network (TAN) in pediatric sepsis remains limited.
- This study provides TAN models with high specificity and negative predictive value, which helps to rule out sepsis in the first 24 hours of pediatric intensive care unit admission.

What is the implication, and what should change now?

- PGM is a potential machine learning method that can be investigated further in pediatric sepsis. With future studies, clinicians have the flexibility to choose models based on the availability of the variables to predict sepsis outcomes and rationalize antibiotics in critically ill children.

Methods

Study design

We performed a retrospective study using the Pediatric Intensive Care Dataset (PICD), a publicly available dataset of patients aged 0–18 years admitted to the intensive care units (ICU) at Children's Hospital of Zhejiang University School of Medicine, Zhejiang, China, in 2010–2019 (9). This dataset was provided by PhysioNet and has been used for pediatric research recently (10-12). PhysioNet is a platform for freely accessible clinical data established by members of the Computational Physiology Laboratory at the Massachusetts Institute of Technology. PICD was hosted on PostgreSQL, an open-source relational database management system, for data extraction and processing.

Sepsis definition

We utilized two categories of patients to define sepsis. The first category was selected using the International Disease Classification-10 (ICD-10), who had discharge diagnosis of sepsis (ICD-10: A02.x, A22.x, A26.x, A32.x, A40.x, A41.x, A42.x, B37.x, O85.x, P36.x) (13). The second category included patients with systemic inflammatory response syndrome (SIRS, ICD-10: R65.x) and suspected infection (14). We considered patients with suspected infections as those who had microbiological cultures sampled followed by antibiotic administration within 72 hours, or antibiotic administration followed by cultures taken within 24 hours (15). The infection onset was taken either at the time of cultures or antibiotic administration, whichever occurred first. We reviewed patients with ICD-10 diagnosis of sepsis to ensure that timing of sepsis is close to PICU admission. If the patient did not meet the sepsis onset definition, we examined the dataset to ensure that these patients had an admission diagnosis related to sepsis (e.g., bacterial sepsis, unspecified sepsis, pneumonia), microbiological cultures taken, and/or antibiotics taken within 24 hours of PICU admission. Patient was considered to have septic shock when vasoactive agents

Table 1 List of clinical variables for sepsis diagnosis

Variable groups	Variables
Clinical symptom	
Overall symptoms	Grunt, cool extremities, cry, ill, rigor, moan, scream, quiet, feeble, sick, depression, twitching
Gastrointestinal symptoms	Vomiting, diarrhea, nausea, regurgitation, anorexia, hypoglycemia, hyperglycemia, abdominal distension, constipation
Central nervous symptoms	Unconsciousness, drowsiness, dizziness, lethargy, convulsion, headache, irritability
Skin symptoms	Jaundice, impetigo, conjunctivitis, pale, cellulitis, cyanosis
Respiratory symptoms	Apnea, cyanosis, respiratory distress, cough, phlegm, sputum, wheezing, dyspnea, expectoration, anhelation, asphyxia
Cardiovascular symptoms	Heart failure, heart murmur, pericarditis, endocarditis, myocarditis, ventricular, chest tightness, chest pain, tachycardia, brachy-cardia
Urinary tract symptoms	Urinary tract infection, oliguria
Infective symptoms	Infection, inflammation
Prematurity	Low birthweight, prematurity
Temperature symptoms	Fever, cold, hyperthermia, hypothermia
Vital signs	Temperature, heart rate, respiratory rate, vital oxygen saturation, systolic blood pressure, diastolic blood pressure
Laboratory tests	Acid bacilli, neutrophil percentage, neutrophil absolute count, lymphocyte percentage, lymphocyte absolute count, pct, platelet count, WBC count, PTT, PT, monocyte percentage, monocyte count, hemoglobin, sedimentation rate, glucose, lactate, blood oxygen saturation, SBE, creatinine, direct bilirubin, indirect bilirubin, total bilirubin, procalcitonin, ASO, CRP, PCO_2, PO_2
	Urine bacterial, urine WBC, urine epithelial, urine nitrite, urine bilirubin
	CSF RBC, CSF WBC
Microbiological culture tests (Blood culture, CSF culture, etc.)	Culture result

The variables were collected from different sepsis literature (16-24). ASO, anti-streptolysin O; CSF, cerebrospinal fluid; RBC, red blood cell; WBC, white blood cell; CRP, C-reactive protein; Pct, procalcitonin; PCO_2, partial pressure of carbon-oxygen; PO_2, partial pressure of oxygen; PT, prothrombin time; PTT, partial thromboplastin time; SBE, standard base excess.

were administered (Figure S1, Table S1). The non-sepsis cohort were patients that did not satisfy our study's sepsis definition.

Data extraction and processing

Within this database, an individual patient may have multiple hospitalizations, and each hospitalization may have multiple PICU admissions. As such, we considered each hospitalization as an independent event and only included clinical information from the first PICU admission within each hospitalization. Extracted clinical information included demographic data, clinical symptoms, vital signs, laboratory, and microbiological tests. PICU admission without clinical data or vital signs for the first 24 hours was excluded.

Clinicians reviewed and selected variables as covariates for the models (*Table 1*). Selected variables were consistent with published literature on sepsis diagnosis (16-18). Demographic data included gender, age, and prematurity status. As Han ethnicity was predominant (98%), we excluded ethnicity from the study. Age was categorized (i.e., newborn, neonate, infant, toddler and preschool, school-aged child, adolescent, and young adult) as recommended by Goldstein *et al.* (14). Four categories of covariates included in our model were clinical symptoms, vital signs, laboratory, and microbiological tests. Clinical symptoms were extracted from the physician's notes, considered as a binary variable and grouped into the following categories:

Table 2 List of models built with TAN methods

No.	Model Name	Input variable groups
Single variable group		
1	S	Clinical symptoms
2	V	Vital signs
3	M	Microbiological cultures
4	L	Laboratory tests
Combination of two variable groups		
5	SV	Clinical symptoms, vital signs
6	SM	Clinical symptoms, microbiological cultures
7	VM	Vital signs, microbiological cultures
8	SL	Clinical symptoms, laboratory tests
9	VL	Vital signs, laboratory tests
10	LM	Laboratory tests, microbiological cultures
Combination of three variable groups		
11	SVL	Clinical symptoms, vital signs, laboratory tests
12	SLM	Clinical symptoms, laboratory tests, microbiological cultures
13	VLM	Vital signs, laboratory tests, microbiological cultures
14	SVM	Clinical symptoms, vital signs, microbiological cultures
Combination of all four variable groups		
15	SVLM	Clinical symptoms, vital signs, laboratory tests, microbiological cultures

Details of each variable group are listed in *Table 1*. TAN, tree augmented Naïve Bayes; S, clinical symptoms; V, vital signs; L, laboratory tests; M, microbiological tests.

overall, gastrointestinal, central nervous, skin, respiratory, cardiovascular, urinary tract, infection, prematurity, and temperature symptoms. Vital signs were categorized into "high", "low", and "normal" based on age group (14). Laboratory tests included 34 variables from blood, urine, and cerebrospinal fluid tests. We utilized laboratory results as defined by PICD without further processing (i.e., "high", "normal", and "low" using the dataset's pre-defined thresholds) (9). Microbiological tests were positive when there was organism growth and negative when there was culture. If clinical symptoms were absent, or laboratory and microbiology tests were not conducted, these were considered as "no record".

Tree augmented Naïve Bayes

TAN relies on probability and graph theory to perform classification, which labels an outcome based on the provided evidence. Along with other Bayesian classifiers such as Naïve Bayes (NB), TAN is commonly used in medical diagnosis, pattern recognition, and natural language processing (7,25,26). Appendix 1 and Figure S2 show the simplified example of TAN and NB for sepsis diagnosis and the algorithm to construct TAN (27). TAN can also capture the correlations between variables while maintaining a simple graph structure (27,28). We built TAN using GeNie Academic software (Bayesfusion, USA) (29) and learned parameters with the Expectation-Maximization algorithm. A total of 15 models were built from the combinations of four categories of variables (*Table 2*). All models were trained and tested with ten-fold cross-validations (k=10). Data were randomly split into k parts, trained in k-1 parts, and tested in the left-over. Performance was measured by the average sensitivity (SEN), specificity (SPE), accuracy (ACC), and area under the curve (AUC) of the ten-fold cross-validations with 95% confidence interval (CI). We compared our model against logistic regression (LR), a supervised model for predicting the probability of binary outcomes, as the performance benchmark, using the same study settings and variables (30). We also validated the models with different data cut-off points (24 *vs.* 48 hours), decision thresholds (0.4 *vs.* 0.5 *vs.* 0.6), and cohort subgroups (premature *vs.* term infants, children under 30 days old *vs.* one month-one year *vs.* > one-year age, admission from general wards *vs.* emergency department).

ACC was calculated by: (true positive cases + true negative cases)/cohort samples. AUC was measured by comparing the true positive rate against the false positive rate. The primary outcome of the TAN classifier was sepsis, and we chose the default decision threshold of 0.5. Patients with outcome probability ≥0.5 were labeled as sepsis. The SEN and SPE determine the model accuracy in sepsis and non-sepsis cases, respectively. We also reported the negative predictive value (NPV) and positive predictive value (PPV).

Statistical analysis

We calculated medians [interquartile ranges (IQRs)] for continuous variables and percentages for categorical variables for data analysis. Significance between groups were analyzed by Mann-Whitney U and Chi-square test. All analyses were performed in Microsoft Excel (version 16.55, Microsoft, USA) with a statistical significance taken

as P<0.05.

Ethical statement

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the National University of Singapore's Institutional Review Board (ID: NUS-IRB-2021-673). Written informed consent was not required for our study because of the retrospective nature of the study and the public availability of the dataset.

Results

The PCID dataset had 13,449 hospitalizations with 13,941 PICU admissions with 357 sepsis cases (2.7%) with a PICU mortality rate of 42/357 (11.8%). We excluded 10,927 ineligible admissions, including 492 cases that were not the first PICU admission, 609 cases that did not have PICU clinical data, and 9,826 cases that did not have vital signs within 24 hours of the PICU stay. Thus, a total of 3,014 admissions with an overall median age of 1.13 (0.15–4.30) years old and 1,698 (56.3%) male patients were included in our study (*Figure 1*). Of these, 52 patients were admitted more than once (with 108 unique admissions). There were 134 (4.4%) patients identified with sepsis, including 55 cases of septic shock (41%). The majority of them were diagnosed as unspecified sepsis (ICD-10: A41.9, n=69, 51.5%) and bacterial sepsis (ICD-10: P36.9, n=60, 44.7%). Of these unspecified sepsis cases, there were 55 cases of viral sepsis and no case of fungal sepsis found (ICD-10: B37.7, Candida sepsis). The most common source of viral sepsis was pneumonia (43, 32%, ICD-10: J18.0, J18.9, P23.5, P23.9). The dominant organism was Gram-positive cocci (16, 11.9%), followed by Klebsiella pneumonia (13, 9.7%). 90 (67.2%) patients with sepsis had negative cultures. *Tables 3,4* show the baseline statistics of our cohort shows the distribution of the selected diagnosis variables.

Overall, all models yielded high ACC (0.92–0.96), SPE (0.95–0.99), and AUC (0.77–0.87). SEN varied with different combinations of categories (range: 0.10–0.46) (*Figure 2*). Individual categories reported low SEN (<0.30), especially vital signs and microbiological tests (<0.10). Models performed better when more variables were incorporated. Laboratory tests combination reported the highest SEN (range: 0.34–0.46). The combination of all four categories (clinical symptoms, vital signs, laboratory, and microbiological tests) yielded the best performance [ACC: 0.93 (95% CI: 0.916-0.936); SEN: 0.46 (95% CI: 0.376–0.550), SPE: 0.95 (95% CI: 0.940–0.956), AUC: 0.87 (95% CI: 0.826–0.906)]. The TAN network for all variables is shown in *Figure 3*. All models reported high NPV (range: 0.96–0.97) and low PPV (range <0.50).

A comparison of our model against the LR model and sensitivity analysis is shown in *Table 5*. The ACC, SPE, and PPV were higher in LR [ACC: 0.95 (95% CI: 0.941–0.967), SPE: 0.99 (95% CI: 0.982–0.998), PPV: 0.45 (95% CI: 0.250–0.661)]. The NPV was high in both models [TAN:0.97 (95% CI: 0.968–0.980), LR:0.96 (95% CI: 0.949–0.977)]. However, SEN and AUC in TAN [SEN: 0.46 (95% CI:0.376–0.550), AUC: 0.87 (95% CI: 0.826–0.906)] outperformed LR [SEN: 0.13 (95% CI: 0.05–0.321), AUC: 0.56 (95% CI: 0.509–0.611)]. As a result, TAN achieved a better performance than LR. The performance of our models considering data over 48 hours was better than the 24 hours, indicating that the additional data provided more information and improved the performance. In addition, the models yielded higher SEN when the decision threshold was lowered (*Figure 4*). They also performed better in the premature, and children less than 30 days old. There was no difference in model performance between patients from general wards and emergency departments (*Table 5*).

Discussion

In this study, we demonstrated that PGM is a potential diagnosis model-building tool for pediatric sepsis with a high ability to rule out sepsis within 24 hours of PICU admission. We also evaluated the diagnostic performance of commonly used variables in PICU, and it showed, unsurprisingly, better diagnostic performance when more variables were used.

PGM has been used in several applications from medical diagnosis, object recognition to virus evolution modeling (31-33). Despite limited use in pediatric sepsis, PGM has been investigated in various disease diagnoses, such as cancer, heart disease, and adult sepsis (8,34,35). In 2012, a study applied the Dynamic Bayesian Network (DBN) to detect early sepsis in 3,100 adults within 24 hours of admission and reported an AUC of 0.94 (36). In 2016, Jiang *et al.* proposed a BN-based sepsis monitoring framework for the elderly that can report patients' conditions periodically without human intervention. It was able to detect sepsis approximately 0.5 hours earlier than the traditional BN-based diagnosis model (6). Recently, BERG (Massachusetts) released bAIcis, a BN-based software for large-scale

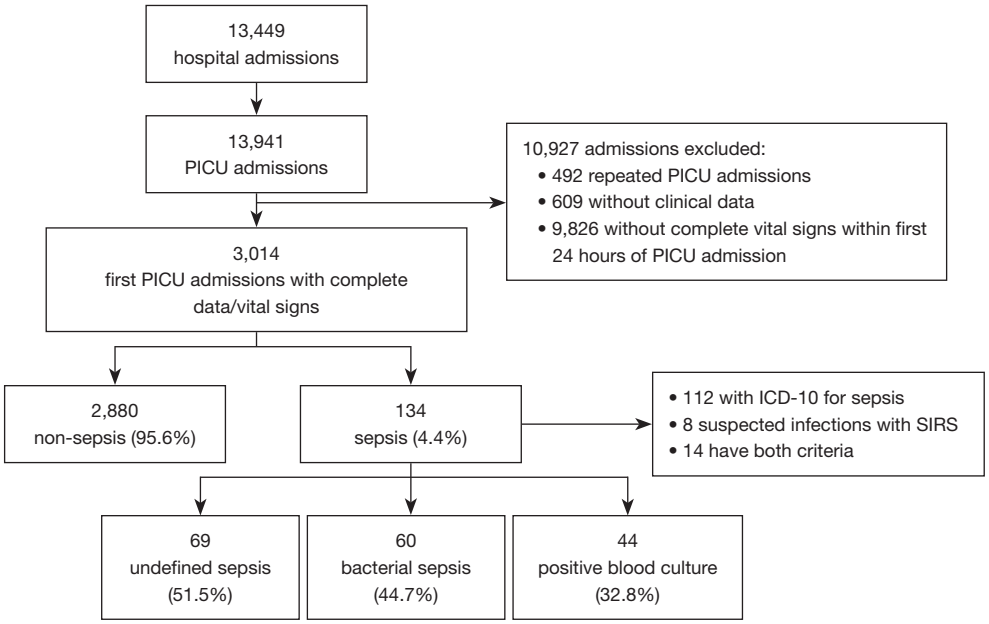

Figure 1 Data extraction flow chart to derive the study cohort. Sepsis cohort (n=134) consisted of 112 cases identified with ICD-10, 8 suspected infections with SIRS, and 14 cases satisfying for both criteria. Of these, 69 and 60 cases were undefined and bacterial sepsis, respectively. 44 out of 134 patients had positive blood culture. ICD-10, international classification of diseases-10; PICU, pediatric intensive care unit; SIRS, systemic inflammatory response syndrome.

Table 3 Baseline clinical demographics and clinical outcomes of the study cohort

Parameters	Overall (N=3,014)	Sepsis cohort (N=134)	Non-sepsis cohort (N=2,880)	P value
Median age, years (interquartile range)	1.13 (0.15–4.30)	0.07 (0.01–1.02)	1.21 (0.18–4.40)	<0.001
Male gender, n (%)	1,698 (56.3)	75 (56.0)	1,623 (56.4)	0.93
History of prematurity, n (%)	108 (3.6)	37 (27.6)	71 (2.5)	<0.001
Use of vasoactive agents, n (%)	1,406 (46.6)	55 (41.0)	1,351 (46.9)	0.183
Median ICU length of stay, days (interquartile range)	0.95 (0.8–3.77)	9.64 (3.14–21.83)	0.94 (0.79–2.91)	<0.001
Median hospital length of stay, days (interquartile range)	11.08 (7.02–18.12)	18.92 (10.33–32.74)	11.02 (7–17.61)	<0.001
In-hospital mortality, n (%)	223 (7.4)	5 (3.7)	218 (7.6)	0.097

Mann-Whitney U test was performed on continuous variables. χ^2 test was performed on categorical variables. Significant level P<0.05. ICU, intensive care unit.

health applications. It was capable of handling hundreds of thousands of features and reported true positive rates of 0.9 and a precision of 0.8 on the synthetics dataset (37). To the best of our knowledge, there were two studies using PGM in pediatric sepsis in 2014 and 2021. The first study applied Auto-Regressive Hidden Markov Model (HMM) to 24 critically ill and low birth-weight neonates in the neonatal ICU and could detect sepsis accurately with an AUC of 0.8 (38). The second study used Markov Chain to study sepsis transition in 140 patients with suspected sepsis after blood cultures (39). Consistent with these prior studies, our findings suggest that PGM is a potential framework for pediatric sepsis that can aid clinical decision-making.

The TAN model, and PGM in general, offer both

Table 4 Distribution of clinical variables within 24 hours of PICU admission

Variables	Overall (N=3,014)	Sepsis cohort (N=134)	Non-sepsis cohort (N=2,880)	P value
Presence of overall symptoms, n (%)	286 (9.5)	68 (50.8)	218 (7.6)	<0.001
Presence of gastrointestinal symptoms, n (%)	487 (16.2)	73 (54.5)	414 (14.4)	<0.001
Presence of central nervous symptoms, n (%)	434 (14.4)	42 (31.3)	392 (13.6)	<0.001
Presence of skin symptoms, n (%)	458 (15.2)	49 (36.6)	409 (14.2)	<0.001
Presence of respiratory symptoms, n (%)	577 (19.1)	82 (61.2)	495 (17.2)	<0.001
Presence of cardiovascular symptoms, n (%)	426 (14.1)	49 (36.6)	377 (13.1)	<0.001
Presence of infective symptoms, n (%)	473 (15.7)	59 (44.0)	414 (14.4)	<0.001
Presence of abnormal temperature symptoms, n (%)	472 (15.7)	61 (45.5)	411 (14.3)	<0.001
Median heart rate, bpm (interquartile range)	140 (121–156)	158 (144–169.5)	138 (120–155)	<0.001
	High: 807	High: 49	High: 758	
	Low: 23	Low: 1	Low: 22	
	Normal: 2,184	Normal: 84	Normal: 2,100	
Median respiratory rate, /min (interquartile range)	34 (28–56)	52 (42–56)	34 (28–43)	<0.001
	High: 576	High: 40	High: 536	
	Low: 352	Low: 1	Low: 351	
	Normal: 2,086	Normal: 93	Normal: 1,993	
Median temperature, ℃ (interquartile range)	37.3 (37–37.7)	37.1 (36.9–37.7)	37.3 (37–37.7)	0.068
	High: 209	High: 21	High: 188	
	Low: 33	Low: 2	Low: 31	
	Normal: 2,772	Normal: 111	Normal: 2,661	
Median oxygen saturation, % (interquartile range)	99 (98–100)	96 (92–100)	99 (98–100)	<0.001
	Normal: 2,413	Normal: 94	Normal: 2,319	
	Low: 601	Low: 40	Low: 561	
Median SBP, mmHg (interquartile range)	98 (85–110)	76 (58.5–95.5)	99 (86–95.5)	<0.001
	High: 797	High: 18	High: 779	
	Low: 557	Low: 62	Low: 495	
	Normal: 1,660	Normal: 54	Normal: 1,606	
Median DBP, mmHg (interquartile range)	56 (46–66)	43 (32–58)	56 (46–66)	<0.001
	High: 1,049	High: 40	High: 1,009	
	Low: 539	Low: 53	Low: 486	
	Normal: 1,426	Normal: 41	Normal: 1,385	
Median WBC count, $\times 10^9$/L (interquartile range)	11.34 (7.66–14.18)	10.51 (6.15–14.5)	11.38 (7.78–14.18)	0.386
Median neutrophil absolute count, $\times 10^9$/L (interquartile range)	6.47 (3.56–11.5)	4.97 (2.32–10.69)	6.6 (3.62–11.53)	0.127
Median lymphocyte absolute count, $\times 10^9$/L (interquartile range)	1.87 (1.25–2.77)	2.15 (1.04–3.11)	1.87 (1.25–2.76)	<0.001

Table 4 (*continued*)

Table 4 (*continued*)

Variables	Overall (N=3,014)	Sepsis cohort (N=134)	Non-sepsis cohort (N=2,880)	P value
Median platelet count, $\times 10^9$/L (interquartile range)	279.5 (211.25–356)	200 (123.25–324)	282 (215–357)	<0.001
Median PTT, second (interquartile range)	30 (26.6–35.7)	40.75 (30.9–57.6)	29.8 (26.5–35.3)	<0.001
Median PT, second (interquartile range)	21.3 (11.6–13.4)	14.3 (12.3–20.3)	20.8 (11.6–20.3)	<0.001
Median erythrocyte sedimentation rate, mm/h (interquartile range)	12 (4.25–27.5)	15 (6–26)	11 (4.25–32.5)	<0.99
Median glucose, mmol/L (interquartile range)	4.02 (2.55–4.83)	3.52 (2.55–4.39)	4.07 (2.93–5.19)	0.048
Median lactate, mmol/L (interquartile range)	1 (0.8–1.6)	1.6 (1.03–2.4)	1 (0.7–1.6)	<0.004
Median creatinine, μmol/L (interquartile range)	40 (33–51)	58 (43–77.5)	39 (33–50)	0.0049
Median procalcitonin, ng/ml (interquartile range)	0.337 (0.091–1.17)	2.66 (0.21–12.57)	0.318 (0.088–1.085)	1
Median CRP, mg/L (interquartile range)	1 (0.5–10)	6.28 (0.5–45.89)	1 (0.5–8)	<0.001
Positive microbiological test, n (%)	301 (10)	24 (17.9)	277 (9.6)	0.002

Vital signs were categorized into "high", "low", and "normal" based on age group following cut-offs from (14). The Mann-Whitney U test was performed on continuous variables. χ^2 test was performed on categorical variables. Significant level P<0.05. bpm, beat per minute; DBP, diastolic blood pressure; SBP, systolic blood pressure; CRP, c-reactive protein; PCT, procalcitonin; PICU, pediatric intensive care unit; PT, prothrombin time; PTT, partial thromboplastin time; WBC, white blood cell.

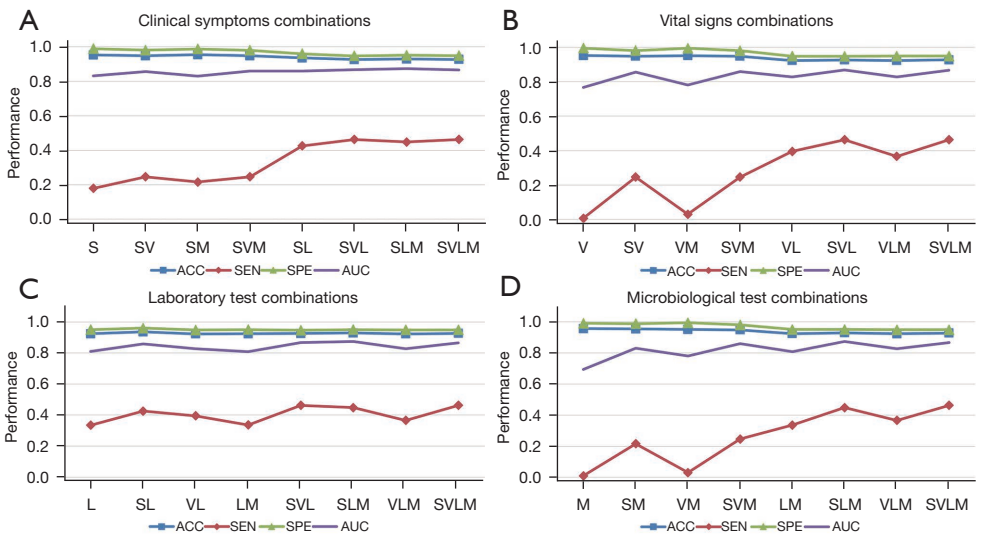

Figure 2 Performance of TAN diagnosis models: (A) clinical symptoms combinations, (B) vital signs combinations, (C) laboratory test combinations, (D) microbiological test combinations. Accuracy (ACC): line with square, Sensitivity (SEN): line with diamond, Specificity (SPE): line with triangle, area under the curve (AUC): simple line. Model names are listed in *Table 2*. Model names comprise of S, V, L, or M. For example, SL stands for model built with variable from clinical symptoms and laboratory tests. S, clinical symptoms; V, vital signs; L, laboratory tests; M, microbiological tests; TAN, tree augmented Naïve Bayes.

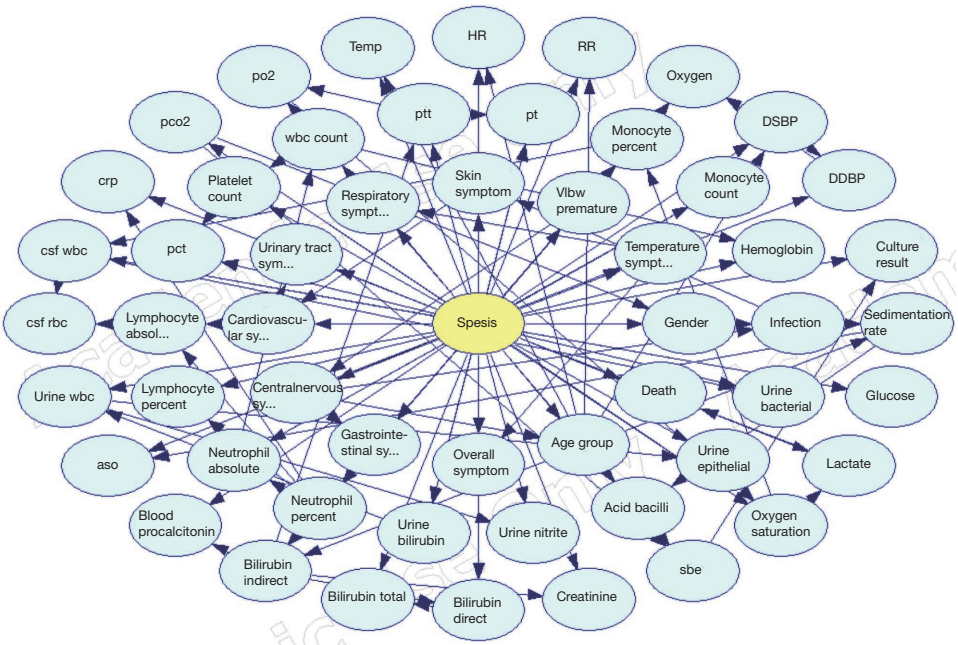

Figure 3 TAN model with full variables from four categories. Variables are listed in Table 1. ASO, anti-streptolysin O; CSF, cerebrospinal fluid; CSF RBC, cerebrospinal fluid red blood cell; CSF WBC, cerebrospinal fluid white blood cell; CRP, c-reactive protein; DDBP, diastolic blood pressure; DSBP, systolic blood pressure; HR, heart rate; TAN, tree augmented Naïve Bayes; Pct, procalcitonin; PCO_2, partial pressure of carbon-oxygen; PO_2, partial pressure of oxygen; PT, prothrombin time; PTT, partial thromboplastin time; RR, respiratory rate; SBE, standard base excess; Temp, temperature.

prediction and inference capabilities (i.e., it can produce predictions and allow variable investigation at the same time). Clinicians can investigate the association between sepsis and biomarkers and how one variable affects another. It is also possible to determine which biomarker contributes the most to sepsis recognition. In addition, PGM is a better tool to model uncertainty and offers a graphical representation that most machine learning methods do not have (40). Clinicians can get involved in the model building and validation process using this graphical representation. With these advantages, in our opinion, PGM is a robust tool for disease diagnosis modeling and for future studies examining the utility of machine learning in sepsis.

We observed that individual biomarker categories were not sufficient to diagnose sepsis. The SEN of the diagnosis models only improved when more data were incorporated, and inclusion of laboratory tests increased SEN. When combined with two initial important variable groups (i.e., clinical symptoms and vital signs), laboratory tests improved SEN by approximately 20% (*Figure 2*). This was not surprising as certain laboratory tests provided more reliable information about sepsis (e.g., C-reactive protein, procalcitonin), whilst vital signs and clinical symptoms were non-specific to sepsis alone. Therefore, even in resource-limited settings, laboratory tests should be considered in the identification of sepsis, compared to vital signs, clinical symptoms, and microbiological tests. However, as each laboratory order involves a certain cost and the variables in our study may not be optimal, future studies should investigate the contributing factor of each laboratory variable in sepsis recognition to help determine which tests should be prioritized in the recognition of sepsis in resource-limited settings. We also observed that there was a high incidence of negative microbiology culture results (67.2%) and microbiological tests contributed little to the SEN performance in our results. Considering the fact that the microbiology results are often unavailable at the time of diagnosis, they should only be used to confirm the presence of sepsis at a later stage instead of being an early diagnosis variable. The clinical symptoms, vital signs, and laboratory tests model (SVL), without microbiological tests, can be used as the diagnosis model instead of the full-feature model (SVLM) at the time of sepsis evaluation. Based on our results, vital signs and clinical symptoms were

Table 5 Performance of the full-feature model (SLVM) in different sub-groups of the cohort

Group/ Performance	ACC (95% CI)	SEN (95% CI)	SPE (95% CI)	AUC (95%CI)	PPV (95% CI)	NPV (95% CI)
Premature infants	0.766 (0.665–0.844)	0.483 (0.299–0.671)	0.892 (0.785–0.952)	0.736 (0.622–0.520)	0.667 (0.431–0.845)	0.795 (0.681–0.877)
Term infants	0.873 (0.840–0.899)	0.412 (0.293–0.550)	0.932 (0.904–0.952)	0.817 (0.750–0.884)	0.446 (0.316–0.584)	0.924 (0.895–0.946)
Age <30 days	0.889 (0.860–0.913)	0.338 (0.231–0.464)	0.961 (0.940–0.976)	0.783 (0.716–0.850)	0.534 (0.378–0.685)	0.917 (0.890–0.938)
Age between 1 m to 1 yr	0.960 (0.944–0.971)	0.281 (0.144–0.470)	0.986 (0.974–0.992)	0.889 (0.814–0.964)	0.429 (0.226–0.656)	0.973 (0.959–0.982)
Age >1 yr	0.967 (0.957–0.975)	0.235 (0.114–0.416)	0.983 (0.976–0.989)	0.883 (0.809–0.957)	0.242 (0.117–0.426)	0.983 (0.975–0.989)
General wards	0.973 (0.965–0.979)	0.213 (0.112–0.361)	0.989 (0.983–0.993)	0.853 (0.785–0.921)	0.286 (0.152–0.465)	0.984 (0.977–0.988)
Emergency units (ICU, PICU, NICU, SICU)	0.851 (0.822–0.876)	0.241 (0.159–0.347)	0.936 (0.913–0.953)	0.789 (0.731–0.847)	0.344 (0.230–0.478)	0.899 (0.872–0.920)
24 hours data cut-off	0.930 (0.916–0.936)	0.463 (0.376–0.550)	0.949 (0.940–0.956)	0.866 (0.826–0.906)	0.295 (0.235–0.362)	0.974 (0.968–0.980)
48 hours data cut-off	0.946 (0.918–0.936)	0.469 (0.377–0.550)	0.953 (0.941–0.957)	0.867 (0.828–0.906)	0.301 (0.239–0.368)	0.974 (0.968–0.980)
TAN	0.930 (0.916–0.936)	0.463 (0.376–0.550)	0.949 (0.940–0.956)	0.866 (0.826–0.906)	0.295 (0.235–0.362)	0.974 (0.968–0.980)
LR	0.950 (0.941–0.967)	0.130 (0.05–0.321)	0.990 (0.982–0.998)	0.560 (0.509–0.611)	0.450 (0.250–0.661)	0.960 (0.949–0.977)

S, clinical symptoms; L, laboratory tests; V, vital signs; M, microbiological tests; ACC, accuracy; AUC, area under the curve; ICU, intensive care unit; NICU, neonatal intensive care unit; NPV, negative predictive value; PICU, pediatric intensive care unit; PPV, positive predictive value; SEN, sensitivity; SICU, surgical intensive care unit; SPE, specificity; TAN, tree augmented Naïve Bayes; LR, logistic regression.

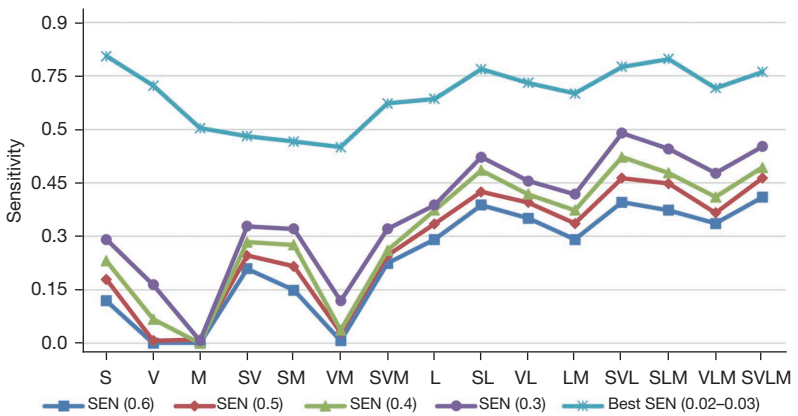

Figure 4 Model performance with different decision thresholds. Line with circles, triangles, diamonds, and squares represents SEN at a decision threshold of 0.3, 0.4, 0.5 and 0.6, respectively. The top line with stars represents the best SEN achieved by each model with the decision threshold ranging from 0.02–0.03. S, clinical symptoms; V, vital signs; M, microbiological tests; L, laboratory tests; SEN, sensitivity. Model names are listed in *Table 2*. Model names comprise of S, V, L, or M. For example, SL stands for model built with variable from clinical symptoms and laboratory tests.

good indicators to exclude sepsis in the first 24 hours of PICU admission. Models built on these categories should be used as a sepsis screening tool at the time of admission. Adding laboratory tests can be used to identify sepsis; whilst the combinations of vital signs and clinical symptoms can be used to rule out sepsis. At any time, clinicians have the flexibility to choose which model to use based on the availability of the variables (and resources).

Our models were not ideal for sepsis detection due to the low SEN. However, they were efficient at excluding sepsis because of the high SPE and NPV. When the laboratory tests and other variables are available, they can be quickly fed into the model to predict the patient's conditions. If the prediction is negative, clinicians can consider stopping antibiotics early. When the predicted probability is borderline, the clinician may consider keeping the antibiotics and stopping them only when indicated. Clinicians have the flexibility to choose which model to use based on the availability of the variables. We recommend using the SVL model because it yielded the same performance as the SVLM model without the need for microbiological tests. Once the microbiological result is available, the SVLM model can be used to confirm the patient's condition.

We included only the worst vital sign measure within 24 hours of PICU admission. This approach may not accurately reflect their dynamic nature. To improve the utility of vital signs in model building, investigators can consider taking hourly data and applying the temporal PGM (DBN and HMM) (41-43). This approach offers more data to the diagnosis models, provides insights into data trends, and improves prediction power. The presence of "no record" was another factor contributing to the low SEN of our models by adding biases and uncertainty. To minimize this issue, investigators can consider two options. First, increase the frequency of data to decrease the number of "no record" entries. This, however, would raise costs and create more challenges for resource-limited settings. Second, apply feature selection to filter the low-contributing variables so that the most valuable information will be retained for predictions. Our dataset suffered an imbalance between positive and negative classes, where the negative cases were predominant. Due to this imbalance, the model produced predictions in favor of the predominant class and underestimated the other. As a result, the model calibration performed poorly and required correction (44). This aspect should be addressed in future studies together with the solutions to the limitation of class imbalance.

There are several limitations to this preliminary study. First, by using the ICD-10 code, we may have missed cases of patients with bacteremia (ICD-10: R78.81) and other subgroups of sepsis (e.g., viral and fungal sepsis). This may have led to a small number of sepsis cases, causing the problem of imbalanced dataset. The ICD-10 code was assigned only at the hospital discharges, so it was also challenging to determine the onset of sepsis and assess the model performance in different subgroups (e.g., nosocomial infections *vs.* community-acquired sepsis). Second, the dataset lacked certain important clinical data (e.g., Glasgow Coma Scale, oxygen support, invasive mechanical ventilation settings), which precluded us from defining organ dysfunction based on the Surviving Sepsis Campaign 2020 (45). In addition, the dataset did not have granular data on fluid resuscitation and administration, which can be used to study sepsis progression. Third, the prematurity could only be identified using the clinician's notes, which may be prone to biases. The data dates were offset randomly; therefore, we could not investigate the models across different time periods. We also did not consider comorbidities (e.g., oncological disorders, immunodeficiencies) as a variable in this study. Therefore, we could not investigate the effect of this variable in our study. Furthermore, our study only extracted the children in the ICUs and the diagnostic models were trained solely on their clinical characteristics. Therefore, the models may not be applicable to other clinical settings outside of the ICU (e.g., acute care floors or intermediate care units). To apply the model in these settings, additional amendment and re-training would be required. Finally, this study was heavily dependent on the PICD data set, which may include several types of biases, such as changes in patient care or clinical programs. As of the current version, the dataset owners are facing several challenges in integrating and processing the data and are working on releasing better quality data in the next version. There is no information on the inter-rater reliability and third-party audits from them. The patients we removed from the cohorts due to missing data might have introduced additional bias to the study as well. Therefore, future studies should consider applying the PGM in different datasets with different sepsis population sizes to verify the robustness of this method. Alternatively, time-series data should be considered to enhance the predictive diagnosis in real-time. Future studies should consider comparing PGM with other probabilistic machine learning methods (e.g., Random Forests and Probabilistic Neural Networks) to evaluate its advantages and disadvantages.

Conclusions

Our study has shown that PGM is a reliable diagnostic tool for pediatric sepsis. Laboratory tests contributed the most to the predictions, while clinical symptoms and vital signs were highly capable of excluding sepsis. Microbiological tests were unreliable due to the high negative incidence. Further studies using different data sets should be conducted to verify the utility of PGM in pediatric sepsis.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at https://tp.amegroups.com/article/view/10.21037/tp-22-510/rc

Peer Review File: Available at https://tp.amegroups.com/article/view/10.21037/tp-22-510/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tp.amegroups.com/article/view/10.21037/tp-22-510/coif). JHL serves as an unpaid Deputy Editor-in-Chief of *Translational Pediatrics* from July 2022 to June 2024. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the National University of Singapore's Institutional Review Board (ID: NUS-IRB-2021-673). Written informed consent was not required for our study because of the retrospective nature of the study and the public availability of the dataset.

References

1. Shah S, Kaul A, Jadhav Y, et al. Clinical outcome of severe sepsis and septic shock in critically ill children. Trop Doct 2020;50:186-90.
2. Tan B, Wong JJ, Sultana R, et al. Global Case-Fatality Rates in Pediatric Severe Sepsis and Septic Shock: A Systematic Review and Meta-analysis. JAMA Pediatr 2019;173:352-62.
3. Koller D, Friedman N. Probabilistic graphical models: principles and techniques. Cambridge, MA: MIT Press; 2009. 1231 p.
4. Lee H, Kim S. Black-Box Classifier Interpretation Using Decision Tree and Fuzzy Logic-Based Classifier Implementation. Int J Fuzzy Log Intell Syst 2016;16:27-35.
5. Onisko A, Druzdzel MJ, Austin RM. Application of Bayesian network modeling to pathology informatics. Diagn Cytopathol 2019;47:41-7.
6. Jiang Y, Sha L, Rahmaniheris M, et al. Sepsis Patient Detection and Monitor Based on Auto-BN. J Med Syst 2016;40:111.
7. Gupta A, Liu T, Shepherd S. Clinical decision support system to assess the risk of sepsis using Tree Augmented Bayesian networks and electronic medical record data. Health Informatics J 2020;26:841-61.
8. Raghu VK, Zhao W, Pu J, et al. Feasibility of lung cancer prediction from low-dose CT scan and smoking factors using causal models. Thorax 2019;74:643-9.
9. Zeng X, Yu G, Lu Y, et al. PIC, a paediatric-specific intensive care database. Sci Data 2020;7:14.
10. Liu R, Greenstein JL, Fackler JC, et al. Prediction of Impending Septic Shock in Children With Sepsis. Crit Care Explor 2021;3:e0442.
11. Wang H, He Z, Li J, et al. Early Plasma Osmolality Levels and Clinical Outcomes in Children Admitted to the Pediatric Intensive Care Unit: A Single-Center Cohort Study. Front Pediatr 2021;9:745204.
12. Morooka H, Tanaka A, Kasugai D, et al. Abnormal magnesium levels and their impact on death and acute kidney injury in critically ill children. Pediatr Nephrol 2022;37:1157-65.
13. World Health Organization, editor. International statistical classification of diseases and related health problems. 10th revision, 2nd edition. Geneva: World Health Organization; 2004. 3 p.

14. 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.
15. 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.
16. Cantey JB, Lee JH. Biomarkers for the Diagnosis of Neonatal Sepsis. Clin Perinatol 2021;48:215-27.
17. Sharma D, Farahbakhsh N, Shastri S, et al. Biomarkers for diagnosis of neonatal sepsis: a literature review. J Matern Fetal Neonatal Med 2018;31:1646-59.
18. Z Oikonomakou M, Gkentzi D, Gogos C, et al. Biomarkers in pediatric sepsis: a review of recent literature. Biomark Med 2020;14:895-917.
19. Mahallei M, Rezaee MA, Mehramuz B, et al. Clinical symptoms, laboratory, and microbial patterns of suspected neonatal sepsis cases in a children's referral hospital in northwestern Iran. Medicine (Baltimore) 2018;97:e10630.
20. Launay E, Gras-Le Guen C, Martinot A, et al. Suboptimal care in the initial management of children who died from severe bacterial infection: a population-based confidential inquiry. Pediatr Crit Care Med 2010;11:469-74.
21. Santos RP, Tristram D. A practical guide to the diagnosis, treatment, and prevention of neonatal infections. Pediatr Clin North Am 2015;62:491-508.
22. Hammett E. Can you spot the signs and symptoms of sepsis? BDJ Team 2019;6:8-10.
23. Chan KH, Sanatani S, Potts JE, et al. The relative incidence of cardiogenic and septic shock in neonates. Paediatr Child Health 2020;25:372-7.
24. Riley C, Basu RK, Kissoon N, et al. Pediatric sepsis: preparing for the future against a global scourge. Curr Infect Dis Rep 2012;14:503-11.
25. Ren H, Guo Q. Flexible learning tree augmented naïve classifier and its application. Knowl-Based Syst 2022;260:110140.
26. Khonsa Izzaty AM, Mubarok MS, Huda NS, Adiwijaya. A Multi-Label Classification on Topics of Quranic Verses in English Translation Using Tree Augmented Naïve Bayes. In: 2018 6th International Conference on Information and Communication Technology (ICoICT) (Internet). Bandung: IEEE; 2018 (cited 2022 Nov 29). p. 103–6. Available online: https://ieeexplore.ieee.org/document/8528802/
27. Chow C, Liu C. Approximating discrete probability distributions with dependence trees. IEEE Trans Inf Theory 1968;14:462-7.
28. Padmanaban H. Comparative Analysis of Naive Bayes and Tree Augmented Naive Bayes Models (Internet) (Master of Science). (San Jose, CA, USA): San Jose State University; 2014 (cited 2022 Mar 9). Available online: https://scholarworks.sjsu.edu/etd_projects/356
29. Genie academic version (Internet). Bayesfusion; Available online: https://www.bayesfusion.com
30. Peng CYJ, Lee KL, Ingersoll GM. An Introduction to Logistic Regression Analysis and Reporting. J Educ Res 2002;96:3-14.
31. Sucar LE. Relational Probabilistic Graphical Models. In: Probabilistic Graphical Models (Internet). London: Springer London; 2015 (cited 2021 Apr 15). p. 219–35. (Advances in Computer Vision and Pattern Recognition). Available online: http://link.springer.com/10.1007/978-1-4471-6699-3_12
32. Gupta A, Slater JJ, Boyne D, et al. Probabilistic Graphical Modeling for Estimating Risk of Coronary Artery Disease: Applications of a Flexible Machine-Learning Method. Med Decis Making 2019;39:1032-44.
33. El-Awady A, Ponnambalam K. Integration of simulation and Markov Chains to support Bayesian Networks for probabilistic failure analysis of complex systems. Reliab Eng Syst Saf 2021;211:107511.
34. Leon C, Carrault G, Pladys P, et al. Early Detection of Late Onset Sepsis in Premature Infants Using Visibility Graph Analysis of Heart Rate Variability. IEEE J Biomed Health Inform 2021;25:1006-17.
35. Zhang Y, Zhu L, Wang X. NEM-Tar: A Probabilistic Graphical Model for Cancer Regulatory Network Inference and Prioritization of Potential Therapeutic Targets From Multi-Omics Data. Front Genet 2021;12:608042.
36. Nachimuthu SK, Haug PJ. Early detection of sepsis in the emergency department using Dynamic Bayesian Networks. AMIA Annu Symp Proc 2012;2012:653-62.
37. Zhang L, Rodrigues LO, Narain NR, et al. bAIcis: A Novel Bayesian Network Structural Learning Algorithm and Its Comprehensive Performance Evaluation Against Open-Source Software. J Comput Biol 2020;27:698-708.
38. Stanculescu I, Williams CK, Freer Y. Autoregressive hidden Markov models for the early detection of neonatal sepsis. IEEE J Biomed Health Inform 2014;18:1560-70.
39. Kausch SL, Lobo JM, Spaeder MC, et al. Dynamic Transitions of Pediatric Sepsis: A Markov Chain Analysis. Front Pediatr 2021;9:743544.
40. Sucar LE. Probabilistic Graphical Models (Internet).

London: Springer London; 2015 (cited 2022 Jan 26). (Advances in Computer Vision and Pattern Recognition). Available online: http://link.springer.com/10.1007/978-1-4471-6699-3
41. Dagum P, Galper A, Horvitz E, Seiver A. Uncertain reasoning and forecasting. Int J Forecast 1995;11:73-87.
42. Baum LE, Petrie T, Soules G, Weiss N. A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains. Ann Math Stat 1970;41:164-71.
43. Baum LE, Petrie T. Statistical Inference for Probabilistic Functions of Finite State Markov Chains. Ann Math Stat 1966;37:1554-63.
44. Van Calster B, McLernon DJ, van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. BMC Med 2019;17:230.
45. 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. Pediatr Crit Care Med 2020;21:e52-e106.

Cite this article as: Nguyen TM, Poh KL, Chong SL, Lee JH. Effective diagnosis of sepsis in critically ill children using probabilistic graphical model. Transl Pediatr 2023;12(4):538-551. doi: 10.21037/tp-22-510

Appendix 1 Tree Augmented Naïve Bayesian Network (TAN)

Classification refers to a task that assigns corresponding labels to outcomes. Supervised classification finds the suitable functions or rules to achieve the known outcomes using the evidence. Unsupervised classification clusters object into groups or categories when the outcomes are unknown. Let Y = {0,1} or {−1,1} be the binary indicating the outcome of the diagnosis, and n random variables **X**= {X1, X2, ..., Xn} represent the information needed to diagnose the patient's condition, known as variables or features. Applying the Bayes rules, we can derive the conditional probability of an outcome given attributes and the supervised classification problem is then described as:

$$argmax_Y\left(P\left(Y\left|\boldsymbol{X}\right.\right)\right)=argmax_Y\left(\frac{P\left(Y\right)P\left(\boldsymbol{X}\left|Y\right.\right)}{P\left(\boldsymbol{X}\right)}\right) \quad [1]$$

There is no independence assumption among variables in the general Bayesian classifier, leading to a complete network with complex and costly calculations. This type of network is only suitable for problems with few variables. Depending on the processing power, this number of variables can be increased. The more variables, the more it will tax the training time and resources. Here, we review some alternatives of Bayesian classifiers, including Naive Bayes, Semi-Naive Bayesian, and Tree Augmented Naive Bayes. The main difference between them is the assumption of conditional independence among the variables.

If the general Bayesian classifier considers no independence and the Naive Bayesian considers all independence, TAN is the model in between. It allows some dependence between variables, allowing information to flow better into the graph while maintaining a simple network structure and low computational costs. This dependence is measured using mutual information, a pairwise score based on information theory:

$$MI\left(X_i,X_j\left|Y\right.\right)=\sum_{k=1}^{N_{Xi}}\sum_{h=1}^{N_{Xj}}\sum_{l=1}^{N_Y}P\left(k,h,l\right)log\frac{P\left(k,h\left|l\right.\right)}{P\left(k\left|l\right.\right)P\left(h\left|l\right.\right)} \quad [2]$$

Tree Augmented Naïve Bayes construction algorithm

TAN, compared to other BN classifiers, is superior in capturing the associations of variables by utilizing the mutual information (MI) between them (21). This MI is calculated by the following equation:

$$MI\left(X_i,X_j\left|Y\right.\right)=\sum_{k=1}^{N_{Xi}}\sum_{h=1}^{N_{Xj}}\sum_{l=1}^{N_Y}P\left(k,h,l\right)log\frac{P\left(k,h\left|l\right.\right)}{P\left(k\left|l\right.\right)P\left(h\left|l\right.\right)} \quad [3]$$

where X_i, X_j are the pair of variables (e.g., gender, age), Y is sepsis outcome, N_{Xi}, N_{Xj}, N_Y are the number of values in X_i, X_j, and Y.

TAN network for sepsis diagnosis is built using the following steps (21):

Step 1: Build the complete undirected graphs with the variables. Assign the MI of each pair of variables as the weight of the respective edge between them.

Step 2: Find the spanning tree that has the maximum weight.

Step 3: Transform the maximum weighted spanning tree in step 2 to a directed graph with the arbitrary or chosen root node.

Step 4: Attach the sepsis node to the network with the arrows coming from sepsis node to other nodes.

https://dx.doi.org/10.21037/tp-22-510

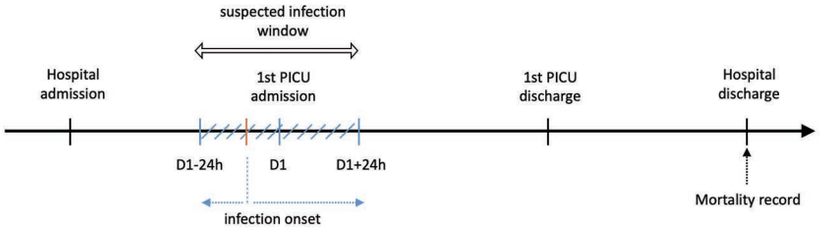

Figure S1 Definition of patients with suspected infection. D1 is the first day of the first PICU admission of hospital admission. Suspected infection was defined as the combination of culture drawn and antibiotic administration within a specific time interval (15). The suspected infection window is between 24 hours prior to and 24 hours after PICU admission. The onset of infection was taken to be either the time of cultures drawn or the time of antibiotic administration, whichever occurred first. Mortality was recorded at hospital discharge within the given hospitalization. PICU, pediatric intensive care unit.

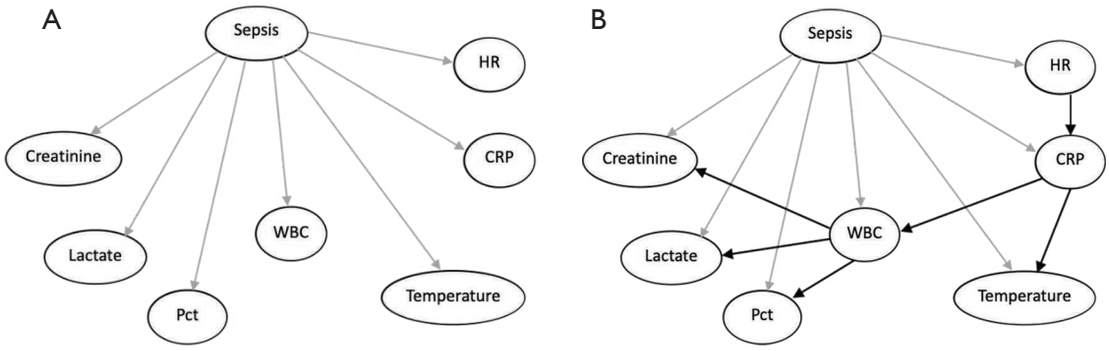

Figure S2 Example of Naïve Bayesian Network (NB) and Tree Augmented Naïve Bayes Network (TAN) to diagnose sepsis. The outcome of interest is the sepsis node. The prediction performed on the sepsis node is calculated based on the given state of seven other variables in the graph. NB classifier assumes that all the variables are independent, which offers a simple graphical representation and calculation (A). However, it may not be practical in some real-world scenarios. TAN is an extension of NB that allows additional dependency of one variable to one another, which enables the information to flow more naturally in the graph and improves performance (B). CRP, C-reactive protein; HR, heart rate; Pct, procalcitonin; WBC, white blood cell count.

Table S1 List of vasoactive agents for shock

Vasoactive agents
Adrenaline
Amrinone
Dobutamine
Dopamine
Dopexamine
Ephedrine
Epinephrine
Erlipressin
Isoprenaline
Levosimendan
Metaraminol
Milrinone
Noradrenaline
Norepinephrine
Phenylephrine
Vasopressin

https://dx.doi.org/10.21037/tp-22-510