Predictive model for bacterial late-onset neonatal sepsis in a tertiary care hospital in Thailand

Dominicus Husada1*, Pornthep Chanthavanich2, Uraiwan Chotigeat3, Piyarat Sunttarattiwong3, Chukiat Sirivichayakul2, Krisana Pengsaa2, Watcharee Chokejindachai2 and Jaranit Kaewkungwal2

Abstract

Background: Early diagnosis of neonatal sepsis is essential to prevent severe complications and avoid unnecessary use of antibiotics. The mortality of neonatal sepsis is over 18% in many countries. This study aimed to develop a predictive model for the diagnosis of bacterial late-onset neonatal sepsis.

Methods: A case-control study was conducted at Queen Sirikit National Institute of Child Health, Bangkok, Thailand. Data were derived from the medical records of 52 sepsis cases and 156 non-sepsis controls. Only proven bacterial neonatal sepsis cases were included in the sepsis group. The non-sepsis group consisted of neonates without any infection. Potential predictors consisted of risk factors, clinical conditions, laboratory data, and treatment modalities. The model was developed based on multiple logistic regression analysis.

Results: The incidence of late proven neonatal sepsis was 1.46%. The model had 6 significant variables: poor feeding, abnormal heart rate (outside the range 100–180 x/min), abnormal temperature (outside the range 36–37.9 °C), abnormal oxygen saturation, abnormal leucocytes (according to Manroe’s criteria by age), and abnormal pH (outside the range 7.27–7.45). The area below the Receiver Operating Characteristics (ROC) curve was 95.5%. The score had a sensitivity of 88.5% and specificity of 90.4%.

Conclusion: A predictive model and a scoring system were developed for proven bacterial late-onset neonatal sepsis. This simpler tool is expected to somewhat replace microbiological culture, especially in resource-limited settings.

Keywords: Predictive model, Bacterial late-onset neonatal sepsis, Scoring system, Thailand

Background

Neonatal sepsis is a global challenge causing high morbidity and mortality among newborns [1–4]. The global infant mortality rate in 2014 was 29 per 1000 live births—the common cause being infection [5]. Neonatal sepsis accounted for 1.4 million neonatal deaths or around 40% of total lives lost, annually [6] About 99% of neonatal deaths occur in low and middle-income countries (LMIC) and approximately 62% occurred during the first 3 days of life [7].

The exact data of neonatal sepsis in the LMIC is limited [8–11]. Two studies from Nigeria showed a prevalence rate of 47.2 and 21.8% [12, 13]. A study from Indonesia found 46.6% prevalence [14]. In Thailand, two decades ago, the prevalence of late-onset neonatal sepsis at Siriraj Hospital, the largest hospital in the country, was 0.05/1000 live births [15]. Ramathibodi Hospital in Bangkokalso recorded almost similar prevalence [16]. Another study from 2012, involving 4 countries, including Thailand, found a prevalence of 21.22 per 1000 admissions [17].

Neonatal sepsis is defined as a clinical syndrome of bacteremia with systemic signs and symptoms of infection in the first 4 weeks of life [18]. Although various organisms can cause neonatal sepsis, the focus of this...
study was bacterial sepsis. Bacteria are the most common cause of neonatal sepsis in the world [2, 4, 5].

There are two types of neonatal sepsis, early- and late-onset. There is little consensus about applicable age limits in literature [19]. Usually, the age limit defined for early-onset sepsis varies from 3 to 7 days [1, 20]. Some clinicians and researchers use 7 days as the limit [17, 19, 21, 22]. Late-onset sepsis is usually caused by organisms acquired after delivery and considered as nosocomial community-acquired infection [17, 22].

Many factors contribute to newborns’ susceptibility to sepsis. The common risk factors are maternal, neonatal, and other conditions that predispose infants to infections, such as invasive procedures [19, 22–25]. Neonates born early or with very low birth weight are highly likely to contract sepsis [2, 26, 27].

Early diagnosis of sepsis improves survival and functional outcome [28, 29]. The other benefit of an early and correct diagnosis is related to antibiotic consumption. A five-year study in Poland revealed the reduction of antibiotic usage [30]. Overuse of antibiotics causes resistance problems worldwide [31].

The detection of neonatal sepsis is difficult due to the non-specific clinical signs and symptoms and the relative diagnostic inaccuracy of the available parameters or biomarkers [32]. Many of non-infectious syndromes have initial clinical presentations similar to severe infections [33]. The gold standard for diagnosis of systemic bacterial infection is the isolation of pathogens, usually from peripheral blood. Unfortunately, the sensitivity of this method is low. Thus, the diagnosis of sepsis cannot be excluded even when the results are negative [34, 35]. When the cultures are negative, but the infant manifests signs consistent with an infection, it may be assumed that they have clinical sepsis [3].

The clinical prediction rule (or predictive model, probability assessment, decision rule, risk score) [36] is a decision-making tool for clinicians with three or more variables obtained from the history, physical examination, and simple diagnostic tests. They are derived from the data collected directly from patients [36–38]. They provide powerful tools to improve clinical decision making [39].

Predictive models quantify the relative importance of individual clinical indicators for evaluating the risk of an adverse outcome for an individual patient [40]. These models attempt to formally test, simplify, and increase the accuracy of a clinician’s diagnostic and prognostic assessment and are most likely to be useful in situations where decision making is complex, the clinical stakes are high, or there are opportunities to achieve cost savings without compromising patient care [36, 41, 42]. This study aimed to develop a predictive model for the diagnosis of late-onset neonatal sepsis. The model, expectedly, helps clinicians determine the infection status of the neonates without waiting for the microbiology facility.

**Methods**

**Study design and site**

This case-control study was performed at Queen Sirikit National Institute of Child Health (QSNICH), Bangkok, Thailand. It has 3 neonatal sick wards, including 1 neonatal intensive care unit (NICU). The initial dataset was compiled from three year periods of the medical record in 2005–2007 and then recalculated in 2014. The need in specific areas was considered, especially in many parts of low- and middle-income countries, including South East Asia. Many areas are immensely burdened by neonatal sepsis patients and require simple tools to overcome the difficulty with the microbiology culture facilities.

**Samples**

Neonates diagnosed with sepsis were included in the case group. Late-onset neonatal sepsis was defined as sepsis at 7 days or more. The inclusion criteria were: age ≤ 28 days on admission, sepsis as the final diagnosis (either main diagnosis or additional), and at least one positive laboratory test for a bacterial pathogen (It could be positive bacterial culture result/polymerase chain reaction (PCR)/gram-staining/latex agglutination tests/antigen-antibody detection for bacteria). The hospital used BacTec (Becton Dickinson Microbiological System, Maryland) for bacterial culture. All the patients with severe congenital malformation that underwent surgery before the diagnosis of sepsis or were admitted for less than 6 h in the hospital were excluded. Inclusion criteria for the control group were: age ≤ 28 days on admission, a final diagnosis other than sepsis, admitted in range of 20 days before or after the comparing sepsis patient, except for the NICU where the time range was expanded to the same year, hospitalized in the same ward with the comparing case, and at least 7 days old on the day of data taken. Thus, each case had 3 controls.

**Definitions**

Neonates: an infant who is less than four weeks old.

Late-onset neonatal sepsis: sepsis diagnosed among neonates at the age of 7 days or more [24, 26, 28].

Clinical sepsis: sepsis in which blood cultures are not performed, not detected, or for which the physician institutes treatment for sepsis. Clinical sepsis patients were not used in this study.

**Data collection and management**

The dependent variable in this study was proven sepsis. The independent variables had 4 categories: risk factors (basic/demographic data, maternal history: antepartum, intrapartum, and postpartum), clinical manifestations,
laboratory findings, and treatment modalities. Initially, 144 variables were considered.

Data collection began by obtaining the list of neonatal patients from the medical record office. The three-year data were compiled and divided into three groups: (a) sepsis with positive bacterial culture result/PCR/gram-staining/latex agglutination tests/antigen-antibody detection for bacteria, (b) clinical sepsis, but without definite specific results as mentioned in the group (a), and (c) non-sepsis. Patients in the group (b) were not included in this study. Group (a) was identified using the International Classification of Diseases (ICD)-10 code of P360 to P368; meanwhile, the ICD-10 code for group (b) was P369.

While selecting the “sepsis group,” the data from culture result records in the neonatal ward was also searched to increase the number of subjects. All medical records of the sepsis group were checked to ensure the fulfillment of inclusion criteria. Subsequently, data from medical records were transferred to case record forms. For clinical and laboratory examinations, the data used were the worst result (could be highest or lowest) in the range of 24 h before or after the diagnosis of sepsis. If such data were not available, the most recent previous data were chosen. The name, address, and hospital number of the patients were not recorded as case records. The hospital numbers were only written in the master log record.

After obtaining all data for the sepsis group, the patients were divided according to the date of admission and the ward/site in the hospital. The control group was selected based on this division and the master medical record list. Controls were matched to the sepsis group based on: (a) date of admission (in the range of 20 days before or after the cases) and (b) hospitalized in the same ward (9, 10, or NICU) with the comparing sepsis patient, and (c) at least 7 days of age. The amount of control: sepsis patients were 3:1. The medical records of patients in the control group were checked to ensure the fulfillment of inclusion criteria. Data from medical records were then transferred to the case record forms. For the control group, the data used were the worst after 7 days of age. Therefore, the records of clinical conditions and laboratory results were observed daily. All the patients in the control group were not diagnosed with sepsis before the data were taken. All data from case record form were transferred to Statistical Package for the Social Sciences (SPSS) database, and data accuracy was rechecked after completing every single record form.

Data analysis
Once the data were available, descriptive, univariate (with t-test, Mann Whitney U, or Chi-Square tests)—as appropriate, and multivariable analysis with multiple logistic regression and calculation of diagnostic test aspects (sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratio (LR), and receiver operating characteristics (ROC) Curve) were performed. All univariate analysis used two-tailed p-value < 0.05. Multivariable analysis used p-value ≤ 0.1. The software used for data analysis was SPSS Version 11.5 (SPSS Inc., Chicago, IL).

The first step of the analytical process was evaluating missing data. Variables with too many incomplete data were not used. For the remaining variables, the missing data were replaced by the imputation method. For the control group, the mean of normal value (based on the literature) were considered. The second step was descriptive analysis. This was done by finding the frequency distributions, mean ± standard deviation (SD), and median (and range).

The third step was the univariate tests, which were done to compare 2 groups: sepsis and non-sepsis. The tests used for comparison were t-test, Mann Whitney U, and chi-square test, depending on the type of data. Variables with p > 0.1 were excluded. The variables with p ≤ 0.1 proceeded to the next step. The fourth step was the selection of remaining variables based on clinical consideration, collinearities, and similarities. The fifth step was the multivariate analysis by multiple logistic regression using the “enter” method. The considerations for the final decision were: number of variables, ease of usage, clinical judgment, performances, and results from several other studies as the comparison. This process resulted in the final equation. The sixth step calculated the sensitivity, specificity, PPV, NPV, LR, and the ROC curve on certain cutoff values of the final equation (or model). In the final, seventh step, the equation was transformed into a scoring system for practical purposes. This scoring system was developed based on the coefficients of each variable in the equation. Some proposed score systems (varying in the process of rounding coefficients) were tried and the best results were selected based on the ROC curve.

Ethical approval
Ethical approval for this study was obtained from two Ethical Committees—The Faculty of Tropical Medicine, Mahidol University, and Queen Sirikit National Institute of Child Health, Bangkok.

Results
Searching for medical records
The study explored 550 medical records from the Medical Record Unit Queen Sirikit National Institute of Child Health (QSNICH), Bangkok. Finally, there were 52 neonates with late-onset sepsis and 156 controls. Forty-five neonates with early-onset sepsis and the other 297
participants were not included because they did not fulfil the inclusion criteria or because of the exclusion criteria. Figure 1 illustrates the results of medical records searching.

**Patient characteristics**

For 3 years, there were 3557 neonatal patients admitted in QSNICH. This study used 11% of the total neonatal patients. Table 1 lists some baseline characteristics of the studied neonates. Most of the neonatal patients in QSNICH were male, weighted between 2500 and 4000 g, and were admitted in the first 24 h of their life. The overall incidence of proven neonatal sepsis at Queen Sirikit Institute of Child (QSNICH) Bangkok was 2.7% (denominator: all neonatal patients in QSNICH). The incidence of proven late-onset neonatal sepsis (LOS) was 1.46%.

The most common diagnosis among the control group was hyperbilirubinemia (79%). The other diagnosis was asphyxia, apnea of prematurity, and respiratory disease syndrome.

**Microbiology and antibiotic**

There were 52 neonatal patients who showed positive culture results from the blood. Three patients also had positive gram stain from cerebrospinal fluid (CSF), and 1 had a positive latex agglutination test from CSF. All those gram stain and latex agglutination test results were comparable with the hemoculture. Among the control group, 2 patients had positive hemoculture for *Coagulase-negative Staphylococcus* (CONS) and 1 had a positive enzyme-linked immunoabsorbent assay (ELISA) test for dengue infection. However, the data from these 2 patients with CONS were taken before the culture procedure. The most common bacteria were *Klebsiella pneumoniae*, CONS, and *Enterobacter* spp. Ampicillin was used as a single or combination drug for 78% of the septic neonates in this study.

**Comparison of outcome**

Most of the patients (53.3%) developed sepsis during the age of 15–28 days. These differ from the control group ($p < 0.001$). The patients who had sepsis had a significantly higher mortality rate and a longer hospitalization compared to the control group. Table 2 lists the comparison between the outcome, age, sex, and the duration of hospitalization.

**Comparison of risk factors**

The odds ratio (OR) regarding the risk factors for sepsis are listed in Table 3. Over 50% of neonatal sepsis patients were born from high-risk pregnancies, compared to only 35% in the control group. Most of their mothers were aged between 15 and 30 years, and were employed as laborers or were unemployed and lived in the slum area. Most of them were educated until elementary or high school. 51.9% mothers were given a steroid injection before birth which protected the neonates. Premature rupture of the membranes was not significantly different from the control group in the sepsis group. Only 6 mothers from this study had chorioamnionitis. Preeclampsia was the most common complication in the pregnancies, (9 cases). The majority of all neonates had good Apgar score either at the first or fifty minutes. The highest odds ratio for risk factors was found for the duration of the hospitalization (4.284), intracranial hemorrhage (3.419), high-risk pregnancies (2.727), and resuscitation of the neonates (2.060).

---

**Fig. 1 The Medical Records Searching Results**
The equation and probability of proven sepsis

Multiple logistic regression produced the following final result: $Y = (2.398 \times \text{poor feeding}) + (3.087 \times \text{abnormal heart rate}) + (3.995 \times \text{abnormal body temperature}) + (1.387 \times \text{abnormal leukocytes}) + (4.328 \times \text{abnormal pH}) - 4.328$. The formula for the probability ($P$) was $P = \frac{\text{EXP}(Y)}{1 + \text{EXP}(Y)}$.

The coding for this equation is listed in the Supplementary Material (Additional file 1). There were 6 variables in the final regression equation—4 from the clinical condition and 2 from laboratory data. The result of the regression equation was placed in the exponential equation to calculate the probability. Probability implies the probability to have proven sepsis in this equation and is expressed as a percentage. Table 5 lists the odds ratio and adjusted odds for all variables in the equations.

The score

To make the final equation easily applicable, a scoring system was derived. The score also included 6 variables. The performance (sensitivity, specificity, PPV, NPV, LR+, and LR (−)) of the equation and the scoring system are presented in two tables in the Supplementary Material (Additional file 1: 2 and 3). The areas under the ROC Curve for the equation and their score were 95.6 and 95.5%, respectively. The proposed cut off for the equation and probability of proven sepsis was 0.606 (3557 neonatal patients during the 3-year study period. Comparing the incidence of neonatal sepsis in countries was not easy since many reports used different criteria for early- and late-onset neonatal sepsis [42]. In Pakistan, Bosnia, and Malaysia, the incidences of LOS were 29, 71.3, and 90.2%, respectively [1, 23, 43].

Discussion

Ninety-seven sepsis patients were identified in this study from 3557 neonatal patients during the 3-year study period. Comparing the incidence of neonatal sepsis in countries was not easy since many reports used different criteria for early- and late-onset neonatal sepsis [42]. In Pakistan, Bosnia, and Malaysia, the incidences of LOS were 29, 71.3, and 90.2%, respectively [1, 23, 43].
Intrapartum history

- **Race (Thai: non-Thai)**: 1.5, p = 0.327
- **Smoking (Yes: No)**: 3.0, p = 0.438
- **Chronic disease of mother (Yes: No)**: 1.0, p = 0.218
- **High risk pregnancy (Yes: No)**: 2.7, p = 0.002
- **Long term drug user (Yes: No)**: 1.0, p = 0.105

Postpartum history

- **Icterus after birth (Yes: No)**: 1.3, p = 0.770
- **Seizure (Yes: No)**: 1.5, p = 1.00
- **Breastfed (Yes: No)**: 1.6, p = 0.205

Notes: $ = Chi-square test; * = Mann-Whitney U test

from four other countries, including Thailand, found an incidence of 5 per 1000 live births [17]. The prevalence was 21.8 or more in Nigeria [12, 13]. A report from the largest hospital in Indonesia found an incidence of 35% [44].

Among all the cases of neonatal sepsis, the percentage of neonates weighing less than 2500 g was 64.1%. Based on the gestational age, the percentage of preterm neonates was 48.9, 69.2, and 59.8% for early-onset sepsis (EOS), LOS, and total sepsis, respectively. These results were similar to other body weight-based reports. Another study reported that the incidence of LOS among very low birth weight (VLBW) neonates was 25–30% and 6–10% in late preterm neonates, with the mortality rate of 36–51% [22]. Data from Kenya and Gambia showed a CFR of 26 and 31% [45, 46]. The percentage of gram-negative organisms in this study was 67.3% (35/52). *Klebsiella pneumoniae* and CONS were the most common microorganisms. These data were comparable with other developing countries [42, 47]. A 10-year prospective surveillance in Brazil revealed 51.6% episodes of neonatal infection caused by gram-negative rods (mainly *Klebsiella* spp. and *E. coli*) [48].

Antibiotics are one of the most important treatments for neonatal sepsis, although some people may not receive this treatment because of the facility limitation in some rural areas [8]. The first line of antibiotics for neonatal sepsis in many countries, like in the studied hospital, are a combination of penicillin group and gentamicin. At least 78% of the LOS patients in this study were administered ampicillin. However, broad-spectrum antibiotics can create problems of resistance. Multi-resistant organisms, such as *A. baumanii* and *K. pneumoniae*, are consistently increasing in many countries, especially in LMIC [8, 44]. Our study focused on bacterial sepsis. All neonatal sepsis patients used antibiotics. This was not used as a decisive variable in our study.

All possible proven neonatal sepsis patients during the 3-year period were included in this study. Nevertheless, this study had a larger sample size than previous studies. The NOSEP Score by Mahieu et al. (2000) used 43 proven episodes and 104 suspected sepsis episodes but did not use non-suspected sepsis patients [49]. Okascharoen et al. (2005) used 1870 neonates, with only 17 proven sepsis patients [16]; Singh et al. (2003) used 30 episodes of definite, 17 most probable, and 58 non-sepsis patients in their study [50]. Recently, the system by Singh was modified using 497 infants in Bangladesh [51]. In 1982, Toller created the first neonatal sepsis score using basic clinical and laboratory data. He used 667 neonates in Ulm hospital [52]. The dependent variable for this study was proven neonatal sepsis. The proof was mostly based on the culture results, particularly hemoculture. All unproven sepsis patients were excluded. The clearly defined outcome variable is an essential requirement [53]. Confirmed sepsis guaranteed consistency and validity of the outcome [51]. The unproven neonatal sepsis patients were excluded from this study to avoid incorporation bias. This bias would appear if the possible predictive factors became part of the diagnostic criteria [3, 34]. Independent variables in the study originated from previous studies about the predictive model for neonatal sepsis and some scores for neonatal morbidity and mortality. In other clinical prediction rules, predictor variables were identified by the process of selecting, exploring, and modeling large amounts of data to discover unknown patterns or relations [36]. In this study,
the independent variables were added by some changes of continuous variables into qualitative forms. Others were made from the unification of some variables.

Initially, the original variables were classified as risk factors / history, clinical conditions, laboratory data, and treatment modalities, as suggested in some previous reports [54]. Some newer laboratory examinations such as procalcitonin [55], various interleukins [56, 57], and PCR methods [58] were not included in this study due to availability and financial reason.

The risk factors included demographic data and maternal history. In this study, the maternal history considered the mother’s habits (smoking, drug use), and the mother’s diseases (fever, amnionitis, history of antibiotics). Maternal diseases significantly contribute to neonatal sepsis—mostly for the early-onset sepsis. Puerperal infection was associated with 2:1 adjusted Risk Ratio for early neonatal mortality. Around 5% of all deaths in the first week of life were attributable to signs suggestive of puerperal infections [59].

To reduce the number of predictor variables and to make the statistical selection, some univariate tests were used as appropriate. In these tests $p < 0.1$ was used, although some other models used $p < 0.2$ [53]. Singh et al. did not use the univariate test for the study [50]. The selection of variables was based on the positive likelihood ratio. The results of the univariate tests were 68 (21 risk factors, 11 clinical condition, 34 laboratories, and 2 treatment modalities) variables.

Multivariate analysis used multiple logistic regression because the outcome variable was dichotomous, and this test was easy [53]. The reselection process was done based on clinical judgment, collinearities (more than 1 variables measured the same thing), similarities, and performances. If continuous and qualitative data were present, the qualitative would be chosen due to practicability. The use of dichotomized data was also accurate and more useful in clinical practice. The original continuous data in NOSEP scoring system did not improve the accuracy of the global scoring system [49].

### Table 4 Odds Ratio for Clinical Conditions, Laboratory Data, and Treatment Modalities between Sepsis and Non-Sepsis Patients

| Variables                                      | OR    | 95% CI Lower | 95% CI Upper | p value |
|------------------------------------------------|-------|--------------|--------------|---------|
| **Clinical Condition**                         |       |              |              |         |
| Lethargy (Yes: No)                             | 15.9  | 7.385–33.737 | < 0.001      |         |
| Poor feeding (Yes: No)                         | 18.4  | 8.477–39.927 | < 0.001      |         |
| Cyanosis (Yes: No)                             | 1.4   | 0.463–4.243  | 0.554        |         |
| Icterus in hospital (Yes: No)                  | 0.2   | 0.087–0.620  | 0.002        |         |
| Abnormal heart rate (Yes: No)                  | 40.8  | 9.029–184.041< 0.001 |       |         |
| Abnormal temperature (Yes: No)                 | 24.8  | 6.880–89.195 | < 0.001      |         |
| Apnea episode (Yes: No)                        | 17.0  | 4.616–62.608 | < 0.001      |         |
| Respiratory insufficiency (Yes: No)            | 11.0  | 5.090–23.931 | < 0.001      |         |
| Hypoxemia (Yes: No)                            | 14.2  | 6.273–32.309 | < 0.001      |         |
| **Laboratory Data**                            |       |              |              |         |
| Anemia (Yes: No)                               | 5.1   | 2.593–10.067 | < 0.001      |         |
| Abnormal leucocytes (Manroe’s) (Yes: No)       | 17.4  | 6.056–49.921 | < 0.001      |         |
| Thrombocytopenia (Yes: No)                     | 9.5   | 4.260–21.264 | < 0.001      |         |
| Abnormal pH (Yes: No)                          | 18.3  | 8.437–39.577 | < 0.001      |         |
| Abnormal base excess (Yes: No)                 | 10.5  | 4.597–23.778 | < 0.001      |         |
| Abnormal potassium (Yes: No)                   | 4.1   | 2.108–7.869  | < 0.001      |         |
| Abnormal sodium (Yes: No)                      | 4.4   | 1.764–10.872 | 0.001        |         |
| Abnormal BUN* (Yes: No)                        | 10.1  | 3.681–27.095 | < 0.001      |         |
| Abnormal CSF* WBC* (YN)                        | 14.7  | 5.068–42.456 | < 0.001      |         |
| Abnormal CSF glucose (YN)                      | 24.8  | 6.880–89.195 | < 0.001      |         |
| Abnormal CSF protein (YN)                      | 11.1  | 3.774–32.806 | < 0.001      |         |
| **Treatment Modalities**                       |       |              |              |         |
| Central/Umbilical catheter (Yes: No)           | 6.6   | 3.332–13.161 | < 0.001      |         |
| Total parenteral nutrition (Yes: No)           | 5.8   | 2.810–11.927 | < 0.001      |         |

Notes: *Chi-Square test
*Blood urea nitrogen
*Cerebrospinal fluid
*White blood cell

### Table 5 Odds Ratio and Adjusted Odds of Variables Used in The Final Equation

| Variables                                      | OR     | 95% CI Lower | 95% CI Upper | Adjustodds OR | 95% CI Lower | 95% CI Upper | p value |
|------------------------------------------------|--------|--------------|--------------|---------------|--------------|--------------|---------|
| **Clinical**                                   |        |              |              |               |              |              |         |
| Poor feeding                                   | 18.397 | 8.477        | 39.927       | 10.996        | 3.201        | 37.776       | < 0.001 |
| Abnormal heart rate                            | 40.765 | 9.029        | 184.041      | 21.920        | 2.235        | 215.020      | 0.008   |
| Abnormal temperature                           | 24.771 | 6.880        | 89.195       | 54.334        | 6.997        | 421.915      | < 0.001 |
| Abnormal oxygen saturation                     | 14.236 | 6.273        | 32.309       | 4.004         | 1.019        | 15.729       | 0.047   |
| **Laboratory**                                 |        |              |              |               |              |              |         |
| Abnormal leucocytes                            | 17.388 | 6.056        | 49.921       | 5.967         | 1.162        | 30.645       | 0.032   |
| Abnormal pH                                    | 18.273 | 8.437        | 39.577       | 11.924        | 3.583        | 39.690       | < 0.001 |

Notes: * = multiple logistic regression analysis
The final model was selected based on the variable composition, clinical judgment, and performance of the area under the ROC curve [16, 60].

The final equation used 6 variables (4 clinical conditions and 2 laboratory data). Abnormal heart rate had the second-highest adjusted OR after abnormal temperature. Abnormal heart rate characteristics (reduced variability and transient decelerations) occurred early in neonatal sepsis. These abnormalities were present 12–24 h before the clinical diagnosis of sepsis. This method was studied extensively by Griffin et al. in 2001 and 2003 (external validation) [61]. In this study, the normal value was simpler and not calculated using a sophisticated method. Reduced variability and transient decelerations in heart rate may be an early indicator of clinical instability [62, 63].

Abnormal temperature had the highest adjusted OR in the model. This was the most frequent clinical feature in some studies [16, 49]. For term infants, hyperthermia was a high predictive parameter. Some studies showed that more than 50% of the sepsis patients had a fever, while hypothermia was only found among 15% of the infants [64]. In this study, no infant with hypothermia developed late-onset sepsis. This is like the results by Okascharoen et al. (2005). The mortality rate was high among mild and moderate hypothermia in another study and the proportion of hyperthermia and hypothermia was 13 and 13.5%, respectively [65].

Abnormal leucocytes were determined according to Manroe’s criteria [66]. Leucocytes (total white blood cell (WBC) count) are one of the most common tests for evaluating bacterial infections. The criteria by Manroe were still used by some reference books despite its weaknesses, such as depending on the infant’s age, gestational age, and the blood vessels [66, 67]. Abnormal pH—mostly acidosis—would accompany hypoxemia. Metabolic acidosis is, most commonly, a consequence of lactate acid accumulation from anaerobic metabolism in hypoxic infants.

The NOSEP score had 5 final variables (1 risk factor, 1 clinical condition, and 3 laboratory data). The model from Okascharoen et al. had 6 variables (1 risk factor, 3 clinical conditions, and 2 laboratory data), and Singh et al. used 7 final variables (all clinical conditions) [16, 49, 50]. Later, the Hematology Scoring System was revalidated in India using 110 neonates with a good result [68]. Tollner in 1982 used seven clinical parameters, skin color, capillary refill, muscular hypotonia, apnea, respiratory distress, hepatomegaly, and gastrointestinal symptoms [52]. NEO-KISS was a score based on the German national surveillance scoring system. It includes clinical, biochemical, and hematological criteria [69].

Changing the equation into the scoring system will make the usage of the model easier. In comparison with the probability of the equation, the scoring system had a good result. The regression coefficients were used to determine the score [70]. At least 4 possibilities of rounding the coefficients were tried for each group. A different score would produce a different performance of the result. The best system was chosen based on the area under the curve (AUC) of the ROC curve and other performance indicators. The final scoring system for late-onset neonatal sepsis had AUC of 96.6%. The maximum score for this model was 23.

In this study, the AUC was 95.6% for the equation and 95.5% for the score. The sensitivity and specificity of the equation were above 80% for the probability cutoff of 20–40% (equation), or 2–3 (score) However, the choice of cutoff (including the PPV, NPV, LR+, and LR(-)) depends on the purpose of usage. For the balanced sensitivity and specificity, the choice would have to be above 70% of the value.

In the real clinical setting, the score proposes the use of antibiotics for “high” and “very high” groups. In contrast, no antibiotic is required for the “low” group of neonates. For the medium group, the antibiotic decision...
should be made individually by the attending physician. The clinical prediction rule is not a replacement for clinical judgment and should complement rather than supplant clinical opinion and intuition. Accurate clinical decision making is a central component of patient care [36, 37]. This clinical prediction rule can help the clinician diagnose late-onset neonatal sepsis.

Although some steps in the development were comparable, proper comparison with some other models could not be made easily since each model differs from each other in terms of age criteria, type of variables, validation process, and the purpose of the score. The NOSEP score and Okascharoen’s score use the age criteria of 3 days to determine early- or late-onset sepsis. Rodwell et al. only used the hematology parameter, while Singh et al. (2003) used just clinical conditions [16, 49, 50, 71].

The primary limitation of this study was its retrospective design. Information bias cannot be avoided using that design and data from medical records. The sample size of the study was limited since the total sample had to be divided into 2 groups. The missing data (as an unavoidable part of retrospective design study) was another limitation since any method, however perfect, can lead to biased estimates of the odds ratio and the model performance in predictive models [72]. Regarding the “worst” laboratory results, notably, several biochemistry results might be normal in a septic condition. The choice of patients in the control group (non-sepsis) may also affect the result of the study. For example, in this study, most of the non-sepsis cases had hyperbilirubinemia. The result for the icterus variable might be different if the predominant diagnoses were other diseases. This study also did not use a new data set. However, when our results were compared to the more recent literature, we considered our study to still be appropriate for some settings, especially underdeveloped and developing countries.

The chosen outcome was only proven sepsis. This could result in an underestimation of the true incidence. However, including unproven sepsis would cause incorporation bias. Lastly, validation of a new sample set was needed, either in the same setting or others. It is recommended to prospectively perform the validation process.

Conclusion
In conclusion, our study developed two predictive models for late-onset neonatal sepsis. One as an equation and another as a scoring system. The predictive models enable clinicians, especially in the resource-limited setting, to have an alternative for microbiological culture. External validation should be done soon to assess the real performance of the other institutions.

Supplementary information
Supplementary information accompanies this paper at https://doi.org/10.1186/s12879-020-4875-5.

Additional file 1: Table S1. The Coding for the Final Equation.
Table S2. Performance of The Equation for Bacterial Late-Onset Neonatal Sepsis. Table S3. Performance of The Scoring System for Bacterial Late-Onset Neonatal Sepsis.

Abbreviations
AUC: Area Under the Curve; BUN: Blood Urea Nitrogen; CFR: Case Fatality Rate; CONS: Coagulase-negative Staphylococcus; CSF: Cerebrospinal Fluid; ELISA: Enzyme-linked Immunosorbent Assay; EOS: Early Onset Sepsis; ICD: International Classification of Diseases; LMIC: Low and Middle-Income Countries; LOS: Length of Stay; LR: Likelihood Ratio; NICU: Neonatal Intensive Care Unit; NPV: Negative Predictive Value; OR: Odds Ratio; PCR: Polymerase Chain Reaction; PPV: Positive Predictive Value; QSNICH: Queen Sirikit National Institute of Child Health; ROC: Receiver Operating Characteristics; SD: Standard Deviation; SEAMEO: Southeast Asian Ministers of Education Organization; SEM: Standard Error of the Mean; SPSS: Statistical Package for the Social Sciences; TPN: Total Parenteral Nutrition; VLBW: Very Low Birth Weight; WBC: White Blood Cell

Acknowledgments
We thank Prof. Pratap Singhasivanon and Prof. Tawee Chotpithayasunondh for the great support. Data collection was a significant collaboration with Dr. Supichaya Netsawang. We also thank Dr. Kiengsak Limkittikul, Prof. Ismoedijanto, and Prof. Parwati Basuki, for very valuable suggestions.

Ethical approval and consent to participate
Ethical approval was obtained from the Faculty of Tropical Medicine Mahidol University Bangkok and Queen Sirikit National Institute of Child Health Bangkok. The data used in this study was anonymised before its use.

Author contributions
DH collected, analyzed, and interpreted the data, searched the literature, and drafted the manuscript. PC designed the study, analyzed and interpreted the data, and drafted the manuscript. PS and UC directed the data collection at the hospital, supported data analysis and interpretation, added some literature, and reviewed the manuscript. KP, CS, and WC partly analyzed and interpreted the data and reviewed the manuscript. JK performed the statistical calculations and interpretation and reviewed the manuscript. All authors read and approved the final manuscript.

Funding
Funding for this study was provided by the Faculty of Tropical Medicine, Mahidol University, and Southeast Asian Ministers of Education Organization (SEAMEO) Tropmed in Bangkok. Both funders did not interfere with the whole process of the study.

Availability of data and materials
The datasets used and/or analysed during the current study are de-identified and available from the corresponding author on reasonable request.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
1Department of Child Health, School of Medicine Airlangga University/Dr. Soetomo Hospital, Surabaya 60286, Indonesia. 2Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand. 3Queen Sirikit National Institute of Child Health, Bangkok, Thailand.

Additional file 1: Table S1. The Coding for the Final Equation. Table S2. Performance of The Equation for Bacterial Late-Onset Neonatal Sepsis. Table S3. Performance of The Scoring System for Bacterial Late-Onset Neonatal Sepsis.
47. Ghotaslou R, Ghorashi Z, Nahaei MR. Klebsiella pneumoniae in neonatal sepsis: a 3 year study in the pediatric hospital of Tabriz, Iran. Jpn J Infect Dis 2007;60:126–128.

48. Couto RC, Canvalho EAA, Pedrosa TMG, Pedroso ER, Neto MC, Biscione FM. A 10-year prospective surveillance of nosocomial infections in neonatal intensive care units. Am J Infect Control. 2007;35:183–9.

49. Mahieu LM, De Muynck AO, De Dooy JJ, et al. Prediction of nosocomial sepsis in neonates by means of a computer-weighted bedside scoring system (NOSEP score). Crit Care Med. 2000;28:2026–33.

50. Singh SA, Dutta S, Narang A. Predictive clinical scores for diagnosis of late-onset neonatal septicemia. J Trop Pediatr. 2003;49:235–9.

51. Rosenberg RE, Ahmed ASMNU, Saha SK, Chowdhury MAKA, Ahmed S, Law PA, et al. Nosocomial sepsis risk score for preterm infants in low resource settings. J Trop Pediatr. 2010;56(2):82–9.

52. Tollner U. Early diagnosis of septicemia in the newborn. Clinical studies and tests in neonatal sepsis. Pediatrics. 2005;115:937–43.

53. Guyatt G. Determining prognosis and creating clinical decision rules. In: Haynes RB, Sackett DL, Guyatt GH, Tugwell P. Clinical epidemiology. Third edition. Philadelphia, Lippincott Williams & Wilkins 2008;323–55.

54. Smith PB, Benjamin DK. Clinical approach to the infected neonate. In: Cherry JD, Demmler-Harrison GJ, Kaplan SL, editors. Feigin & Cherry's textbook of pediatric infectious diseases. 6th ed. Philadelphia: Saunders Elsevier; 2009. p. 530–2.

55. Pontrelli G, De Crescenzo F, Buzzetti R, Jenkner A, Balduzzi S, Carducci FC, et al. Accuracy of serum procalcitonin for the diagnosis of sepsis in neonates and children with systemic inflammatory syndrome: a meta-analysis. BMC Infect Dis. 2017;17:302.DOI https://doi.org/10.1186/s12879-017-2396-7.

56. Wagner TA, Gravett CA, Healy S, Soma V, Patterson JC, Gravett MG, et al. Emerging biomarkers for the diagnosis of severe neonatal infections applicable to low-resource settings. J Glob Health. 2011;1(2):210–23.

57. Meem M, Modak JK, Mortuza R, Morshed M, Islam MS, Saha SK. Biomarkers for diagnosis of neonatal infections: a systematic analysis of their potential as a point of care diagnostics. J Glob Health. 2011;1(2):201–9.

58. Ng PC, Lam HS. Biomarkers for late-onset neonatal sepsis: cytokines and beyond. Clin Perinatol. 2010;37:599–610.

59. Bellizzi S, Bassat Q, Ali MM, Sobel HL, Temmerman M. Effect of puerperal fever on early neonatal mortality: a secondary analysis of six demographic and health surveys. PLoS One. 2017;12(1):e0170856.

60. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. J Clin Epidemiol. 2008;61:1085–94.

61. Griffin MP, Lake DE, Moorman JE. Heart rate characteristics and laboratory tests in neonatal sepsis. Pediatrics. 2005;115:937–41.

62. Moorman JR, Carlo WA, Kattwinkel J, Schelonka RL, Porcelli PJ, Navarette CT, et al. Mortality reduction by heart rate characteristics monitoring in very low birth weight neonates: a randomized trial. J Pediatr. 2011;159(6):900–6.

63. Bhatti M, Chu A, Hageman JR, Schreiber M, Alexander K. Future directions in the evaluation and management of neonatal sepsis. Neoreviews. 2012;13(2):e103–11.

64. Estripeaut D, Saez-Llorenz X. Perinatal bacterial diseases. In: Feigin RD, Cherry JD, Demmler-Harrison GJ, Kaplan SL, editors. Feigin & Cherry's textbook of pediatric infectious diseases. 6th ed. Philadelphia: Saunders Elsevier; 2009. p. 979–1020.

65. Ahmad MS, Ali N, Mehbob N, Mehmood R, Ahmad M, Wahid A. Temperature on admission among cases of neonatal sepsis and its association with mortality. J Pak Med Assoc. 2016;66:1303.

66. Manroe BL, Weinberg AG, Rosenfeld CR, Tudehope DI. Early diagnosis of neonatal sepsis using a hematologic scoring system. J Pediatr. 1988;112:761–6.

67. Rodwell RL, Leslie AL, Tudehope DI. Early diagnosis of neonatal sepsis using a hematologic scoring system. J Pediatr. 1988;112:761–6.

68. Gorelick MH. Bias arising from missing data in predictive models. J Clin Epidemiol. 2006;59:1092–101.

69. Meem M, Modak JK, Mortuza R, Morshed M, Islam MS, Saha SK. Biomarkers for diagnosis of neonatal infections: a systematic analysis of their potential as a point of care diagnostics. J Glob Health. 2011;1(2):201–9.

70. Moons KGM, Donders RART, Stijnen T, Harrell FE. Using the outcome for imputation of missing predictor values was preferred. J Clin Epidemiol. 2006;59:1092–101.

71. Rodwell RL, Leslie AL, Tudehope DI. Early diagnosis of neonatal sepsis using a hematologic scoring system. J Pediatr. 1988;112:761–7.

72. Ghotaslou R, Ghorashi Z, Nahaei MR. Klebsiella pneumoniae in neonatal sepsis: a 3 year study in the pediatric hospital of Tabriz, Iran. Jpn J Infect Dis 2007;60:126–128.

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