Predicting risk of serious bacterial infections in febrile children in the Emergency Department.

Adam D Irwin, PhD MRCPCH[1], Alison Grant, BSc[2], Rhian Williams, BSc[2], Ruwanthi Kolamunnage-Dona, PhD[3], Richard J Drew, MD FRCPath[45], Stephane Paulus, MD FRCPCH[6], Graham Jeffers[7], Kim Williams[2], Rachel Breen, PhD[8], Jennifer Preston[7], Duncan Appelbe, PhD[8], Christine Chesters[9], Paul Newland, MPhil[9], Omnia Marzouk, MD FRCPCH[2], Paul S McNamara, PhD FRCPCH[7], Peter J Diggle, PhD[1,10], Enitan D Carrol, MD FRCPCH[1]

Affiliations:

[1]Institute of Infection and Global Health, University of Liverpool, Liverpool, UK
[2]Emergency Department, Alder Hey Children's Hospital NHS Foundation Trust, Liverpool, UK
[3]Department of Biostatistics, Institute of Translational Medicine, University of Liverpool, Liverpool, UK
[4]Department of Microbiology, Rotunda Hospital, Dublin, Ireland
[5]Department of Microbiology, Royal College of Surgeons in Ireland, Dublin, Ireland
[6]Department of Infectious Disease, Alder Hey Children's Hospital NHS Foundation Trust, Liverpool, UK
[7]Institute of Translational Medicine, University of Liverpool, Liverpool, UK
[8]Clinical Trials Research Centre, University of Liverpool, Liverpool, UK
[9]Department of Biochemistry, Alder Hey Children's Hospital NHS Foundation Trust, Liverpool, UK
[10]Medical School, Lancaster University, Lancaster, UK

Corresponding author:
Dr Adam Irwin
Institute of Infection and Global Health
University of Liverpool
Ronald Ross Building
8 West Derby Street
Liverpool L69 7BE
adam.irwin@nhs.net
+44 7859 063222

Table of Contents Summary:
Multinomial regression is used to model risk of serious bacterial infection in febrile children in the Emergency Department

Short title:
Risk prediction in febrile children in ED

Financial Disclosure:
The authors have no financial relationships relevant to this article to disclose

Funding

This work was supported by the National Institute for Health Research, Research for Innovation, Speculation and Creativity Programme [grant number RC-PG-0309-10053 to EDC], and the Alder Hey Charity funding [grant number 8037 to ADI and EDC]. The funding bodies had no role in the design or undertaking of the study.

This article presents independent research funded by the National Institute for Health Research (NIHR) Research for Innovation Speculation and Creativity (RISC) Programme. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Potential conflicts of Interest:

The authors have no conflicts of interest relevant to this article to disclose

Transparency declaration

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Professor Carrol affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

What's known on this subject:

Failure to identify serious infections in children results in adverse outcomes whilst a failure to rule-out serious infections results in unnecessary antibiotic use and hospital admission. Multivariable clinical risk prediction models appear to discriminate well between serious and self-limiting infections.

What this study adds:

In a study of 1101 children of all ages, risk prediction models discriminated well between pneumonia, other serious bacterial infections and none. A published model performed well on external validation and model extension with Procalcitonin and Resistin improved discrimination.

Author contributions

Dr Irwin oversaw the running of the study, collected the data, determined outcome diagnoses, performed laboratory assays, and statistical analysis, wrote the first draft of the manuscript, and revised and approved the final manuscript as submitted.

Ms Grant and Ms R Williams supervised collection of data, contributed to writing the manuscript and approved the final manuscript as submitted.

Dr Kolamunnage-Dona oversaw the running of the study, performed statistical analysis, contributed to writing the manuscript and approved the final manuscript as submitted.

Dr Drew contributed to study design, writing the manuscript, and approved the final manuscript as submitted.

Dr Paulus determined outcome diagnoses, oversaw the running of the study, contributed to writing the manuscript and approved the final manuscript as submitted.

Mr Jeffers and Ms Chesters performed laboratory assays, acquired and interpreted data, contributed to writing the manuscript and approved the final manuscript as submitted.

Ms K Williams and Dr Marzouk designed and oversaw the study, collected the data, contributed to writing the manuscript and approved the final manuscript as submitted.

Dr Breen and Ms Preston oversaw the running of the study, contributed to writing the manuscript, and approved the final manuscript as submitted.

Dr Appelbe supported study design and data acquisition, contributed to writing the manuscript, and approved the final manuscript as submitted.

Dr Newland and Professor McNamara designed the study, contributed to writing the manuscript and approved the final manuscript as submitted.

Professor Diggle performed statistical analysis, contributed to writing the manuscript, reviewed and approved the final manuscript as submitted.

Professor Carrol designed and oversaw the running of the study, determined outcome diagnoses, contributed to writing the manuscript, reviewed and approved the final manuscript as submitted.

All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

ABSTRACT

Background

Improving the diagnosis of serious bacterial infections (SBI) in the children's Emergency Department (ED) is a clinical priority. Early recognition reduces morbidity and mortality, while supporting clinicians to rule out SBI may limit unnecessary admissions and antibiotic use.

Methods

A prospective diagnostic accuracy study of clinical and biomarker variables for the diagnosis of SBI (pneumonia or 'other SBI') in febrile children <16 years. A diagnostic model was derived using multinomial logistic regression, and internally validated. External validation of a published model was undertaken followed by model updating and extension by the inclusion of Procalcitonin and Resistin.

Results

1101 children were studied, of whom 264 had SBI. A diagnostic model discriminated well between pneumonia and no SBI (*c* statistic 0.84, 95%CI 0.78 to 0.90) and between other SBIs and no SBI (0.77, 95% CI 0.71 to 0.83) on internal validation. A published multivariable model discriminated well on external validation. Model updating yielded good calibration with good performance at both high risk (Positive Likelihood Ratios 6.46 and 5.13 for pneumonia and other SBI respectively) and low risk (Negative Likelihood Ratios 0.16 and 0.13) thresholds. Extending the model with the addition of Procalcitonin and Resistin yielded improvements in discrimination.

Conclusions

Diagnostic models discriminated well between pneumonia, other SBIs and no SBI in febrile children in the ED. Improvements in classification of non-events have the potential to reduce unnecessary hospital admission, and improve antibiotic prescribing. The benefits of this improved risk prediction should be further evaluated in robust impact studies.

INTRODUCTION

Acute febrile illness is among the most common of all presentations to the children's Emergency Department (ED).[1] In this context, the probability of serious bacterial infection (SBI) is estimated to be 7% - predominantly lower respiratory or urinary tract infection.[2]

The prompt recognition of SBI is fundamental to effective management. Children with meningococcal disease are frequently missed at initial presentation,[3] and delayed recognition increases mortality.[4, 5] Though rates of invasive infection have declined with the introduction of conjugate vaccines,[6-8] SBI remains an important contributor to childhood morbidity and mortality.[9]

In the UK, as rates of invasive infections have declined, the number of children admitted to hospital has increased.[10] The greatest increase is in young children with uncomplicated admissions for acute infections.[11] Supporting clinicians to rule out SBI may reduce unnecessary hospital admissions in children.[12]

A number of studies have reported the diagnostic accuracy of clinical[13] and laboratory[14] variables in febrile children. More recently, risk prediction models combining clinical variables have been evaluated,[2, 15] and in one the addition of CRP improved diagnostic accuracy.[16] We ourselves have previously reported the combined performance of Procalcitonin, Resistin and Neutrophil Gelatinase-associated Lipocalin (NGAL) in Malawian children.[17]

Diagnostic accuracy studies in febrile children have so far failed to impact clinical practice. Restrictive inclusion criteria, such as age, temperature, or clinical syndrome[18] have limited their external validity and few have progressed to validation in external populations. We therefore set out to derive and internally validate a multivariable risk prediction model, and to

externally validate a previously published model[16] for the diagnosis of SBI in febrile children of all ages.

METHODS

A prospective diagnostic accuracy study of clinical and biomarker variables for the diagnosis of SBI in children presenting to the Alder Hey Children's Hospital ED. This is the busiest children's ED in the UK, managing 60000 attendances each year. Recruitment was undertaken between November 2010 and April 2012. The study is reported in line with the Standards for Reporting of Diagnostic Accuracy (STARD) and Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD) guidelines.[19, 20]

Participants

Children less than 16 years of age with fever (>38°C) or history of fever were eligible if they required blood tests as part of clinical management. Children with primary immunodeficiency were excluded. Using prior estimates of sensitivity and specificity of 65% and 90% respectively, and a rate of SBI of 15%, a sample size of 2300 was proposed. For skin and soft tissue infections, the reference standard for SBI was that children were deemed by the clinical team to require intravenous antibiotics. As the outcome diagnosis was solely based upon a clinical decision, and as this was true of all such cases, these children (n=82) were excluded (Figure 1).

Patient involvement

The GenerationR Young Person's Advisory Group (www.generationr.org.uk), initiated by the National Institute for Health Research (NIHR) helped design patient information leaflets for

young people and families. In the course of the study the group explored improvements in the recognition of serious infection, and the discussed diagnostic tests using various samples (such as saliva or blood). This involvement has informed the design of subsequent studies.

Data

Relevant clinical and biomarker variables were identified from the literature, including two large systematic reviews.[13, 14] Clinical data were entered onto a proforma at the time of the clinical assessment. Where possible, this was done by the attending clinician. When the proforma was incomplete, missing clinical information was retrieved from the clinical notes, where explicitly referenced. Paper proformas were collected by the study team daily. All proformas were cross checked against the clinical notes which were electronically scanned and stored. Missing or ambiguous data were recorded as missing. Data collection and entry into the database was blinded to final outcomes.

Samples

Tests performed in study subjects are recorded in supplementary Table 1. All samples were processed in Clinical Pathology Accredited laboratories. Blood (0.5 to 1ml) inoculated into culture bottles was monitored using the BacT/ALERT 3D system. Positive cultures were processed in line with UK standards for Microbiology investigations developed by Public Health England.[21] Specific *Streptococcus pneumoniae* and *Neisseria meningitidis* PCR assays were performed at the Meningococcal Reference Unit in Manchester.[22, 23] Urine and CSF underwent microscopy and culture on agar gel plates, and were processed in line with UK standards. Multiplex PCR was performed on respiratory (RSV, Influenza A and B, Parainfluenza 1-3, Adenovirus, Rhinovirus and Human Metapneumovirus) and CSF (HSV 1 and 2, Varicella Zoster and Enterovirus) samples at the regional laboratory in Manchester. From April 2011, respiratory PCRs were performed using the FilmArray respiratory viral

panel (Biomerieux) and additionally identified Parainfluenza 4, Rhino/Enterovirus and Coronavirus 1-4.

Blood (0.5-1ml) was collected into Lithium Heparin and plasma stored in Sarstedt microtubes at -80°C within 1 hour. Prior to analysis samples were thawed, vortex mixed and centrifuged to remove bubbles and particulate matter. Procalcitonin analysis was undertaken on the B.R.A.H.M.S. Kryptor according to manufacturer's instructions. Quality control samples were analysed with each run. NGAL and Resistin were analysed using validated commercial ELISA.

Reference tests

In common with other published studies, outcome diagnoses were determined by a composite reference standard incorporating clinical, microbiological and radiological features (supplementary Table 2).[14, 15, 24, 25] Using these pre-defined criteria, a paediatric research fellow and a paediatric infectious disease consultant independently attributed outcome diagnosis. In the case of disagreement, a second paediatric infectious disease consultant determined final outcome. Children who failed to meet the pre-defined criteria for SBI were considered to have 'No SBI'. Subjects were followed up to 28 days to reduce misclassification.

Statistical methods

Analysis was undertaken in R, version 3.0.1.[26] Missing data were handled by ten-fold multiple imputation using fully conditional specification implemented by the MICE package.[27] In this method, missing values are replaced by values drawn from a conditional distribution specific to each individual predictor variable and defined by its own imputation

model. Data were assumed to be 'missing at random'. The proportion of missing data relating to each clinical variable is recorded in supplementary table 3.

Model derivation, validation and updating

The dataset was randomised into a split sample "derivation", and "validation" set. Univariate analysis of clinical and biomarker variables was undertaken using logistic regression, for the outcome of SBI. Explanatory variables were examined for evidence of collinearity. Scatter plots and generalised additive model (GAM) plots,[28] fitted using the gam() function in the mgcv package, were examined for evidence of non-linearity on the log-odds scale. Piecewise and polynomial transformations were undertaken where appropriate. Plausible interaction terms were explored, including interactions between age, heart rate and respiratory rate. A multivariable model was derived using a forwards stepwise method. Improvements in model fit were tested by means of a likelihood ratio test (α=0.05) and variables associated with a significant improvement were retained. Having identified a parsimonious model for SBI, these variables were then included in a multinomial regression model for the categorical outcomes "pneumonia", "other SBI", and "no SBI".

External validation of the model published by Nijman *et al*[16] was undertaken using the published coefficients. A comparison of study participants is given in supplementary table 4. The model was updated by re-fitting variables and estimating the individual co-efficients, then extended by the inclusion of Procalcitonin and Resistin. This strategy preserved the original model structure and avoided deriving an entirely new model. The biomarkers were chosen having observed their value in our earlier model derivation. Additional clinical variables were not investigated as they appeared less predictive in our model derivation, and plausible clinical variables were adequately represented by the published model.

Model evaluation

Performance characteristics of the fitted models at various risk thresholds were estimated using the epiR package.[29] Discrimination was measured using the concordance (*c*) statistic, and illustrated by Receiver Operating Characteristic (ROC) curves using the pROC package.[30] The *c* statistic estimates the probability that a randomly selected subject with the outcome of interest has a higher predicted probability than a randomly selected subject without. Comparison of the *c* statistic was undertaken using the DeLong method.[31] For the multinomial regression model, the *c* statistic estimated discrimination between pairs of patients –a patient with pneumonia and a patient with no SBI, or patient with "other SBI" and a patient with no SBI. Confidence intervals (95%) were estimated by a bootstrapping process using 2000 bootstrap replicates. Calibration of the models (how closely risk predictions fit observed cases) was illustrated using multinomial calibration plots.[32]

In the absence of established methods to report classification in multinomial risk prediction models, we compared crude classification (that is, the most likely diagnosis predicted by the multinomial models) in the updated and extended models. To investigate potential clinical utility, we estimated the ability of the models to 'rule-out' SBI (predictions for both categories of SBI <5%), or to 'rule-in' SBI (prediction of either category >20%). These thresholds represent approximately half and double the observed event rate in the study population.

Ethics

Approval for the study was granted by the Greater Manchester West Research Ethics Committee (10/H1014/53), and by the Alder Hey Children's Hospital R&D department.

RESULTS

Between 1st November 2010 and 3rd April 2012, 7949 children presented to the Alder Hey Children's ED with fever. Of these, 1872 were eligible for inclusion, and 1101 recruited to the study (Figure 1). Median age was 2.4 years (IQR 0.9-5.7 years), and 55% were boys. Approximately one third of children had significant comorbidities (Table 1). 264 children (24.0%) were diagnosed with SBI (supplementary figure 1).

The probability of pneumonia and other SBIs increased linearly with heart rate, respiratory rate and temperature. Consistent with other studies, increased work of breathing (odds ratio 10.4, 95% confidence interval 6.69 to 16.2), hypoxia (9.29, 95%CI 5.35 to 16.1), and other respiratory variables were significantly associated with pneumonia. These features reduced the probability of other SBIs. Neck stiffness , a bulging fontanelle, irritability and dysuria were associated with other SBIs. Prolonged capillary refill time was associated with other SBIs (1.43, 95%CI 1.05 to 1.97) but not pneumoniawhile the presence of a rash reduced the probability of both pneumonia and other SBIs. Univariate odds ratios are presented in supplementary figure 2. CRP, Procalcitonin, NGAL and Resistin were all associated with SBI (supplementary table 5).

Model derivation and internal validation

The derived model included the variables "Respiratory rate", and "Normal Air Entry" alongside CRP, PCT, and Resistin (supplementary table 6). Fitting CRP as a piecewise term improved the model fit. The model discriminated well on internal validation (*c* statistic 0.84, 95%CI 0.78 to 0.90 for pneumonia, and 0.77, 95%CI 0.71 to 0.83 for other SBIs). Calibration plots suggested that the model overestimated the risk of pneumonia (Figure 2).

External validation and updating of Nijman model

The published model of Nijman *et al* was validated in the complete dataset (n=1101). Using the published coefficients, the model discriminated well between pneumonia and no SBI, though less well between other SBIs and no SBI (*c* statistic 0.85 and 0.76 respectively, supplementary figure 3). Model calibration was poor though calibration plots indicated that predicted risks and observed outcomes were highly correlated (Figure 3).

Observing the correlation between predicted probabilities and observed outcomes in the poorly calibrated model, we updated the model by re-estimating the individual co-efficients. No attempt was made to adjust the functional form of predictor variables. The re-fitted model discriminated well (*c* statistic 0.88 and 0.82 for pneumonia and other SBIs respectively), and was well calibrated (Figure 4).The model was then extended by the inclusion of PCT and Resistin. This improved discrimination of the pneumonia (*c* statistic increased from 0.88 to 0.90, p=0.03), and other SBI models (from 0.82 to 0.84, p=0.03) and calibration remained good (supplementary figure 4).

The performance characteristics of the updated and extended models are summarised in Table 2. At a low-risk threshold of 5%, the extended pneumonia model had a sensitivity of 92% (95%CI 85 to 96%) and negative likelihood ratio (NLR) of 0.12 (0.06 to 0.23). For other SBIs, model sensitivity was 92% (86 to 95%), and NLR 0.21 (0.12 to 0.35). At a high-risk threshold (>20%), specificity was 89% (95%CI 87 to 91%) for pneumonia, with a positive likelihood ratio (PLR) of 6.69 (5.30 to 8.44), and 86% (83 to 88%), PLR of 4.96 (4.07 to 6.03) for other SBIs.

Classification (determined by likeliest outcome category) was similar between the updated and extended models (893/1101 v 917/1101, 2.2% improvement, 95%CI -1.1 to 5.4%, supplementary Table 7). Using the extended model, SBI was correctly 'ruled out' in 31 additional children (3.7%, 95%CI -1.0 to 8.4%) and there were five fewer potentially missed SBI diagnoses (14/264 v 19/264, 1.8% reduction, 95%CI -2.6 to 6.4%, Table 3).

DISCUSSION

Main findings

In this large, prospective study of febrile children of all ages presenting to the ED, multinomial risk prediction models discriminated well between pneumonia, other SBIs and none. A newly derived model performed well on internal validation, and identified Procalcitonin and Resistin along with CRP as biomarkers of potential value. A published model performed well on external validation and the addition of PCT and Resistin improved discrimination. At a low-risk threshold (<5%), a NLR of 0.12 (pneumonia) or 0.21 (other SBIs) may help to rule out SBI, whilst at a high-risk threshold (>20%) PLRs of 6.69 and 4.96 may expedite treatment.

Strengths:

We present data on multiple biomarkers of SBI in more than 1000 children. We have evaluated children irrespective of age, past medical history or clinical syndrome, and obtained comparable discrimination to other studies with more restrictive inclusion criteria. In common with other recent data,[2, 16] we have demonstrated the value of combining clinical and biomarker variables.

This is the first broad external validation of the published multivariable model by Nijman *et al*. The model discriminated well, but was poorly calibrated. Specifically, there was a problem of calibration in the large – the model predicted too few cases in our population. Correlation between model predictions and observed cases suggested the overall structure of the model was appropriate to our dataset however and our approach of re-estimating the model coefficients resulted in a well-calibrated model.

Limitations:

This is a single centre study, and whilst we have performed internal validation of our derived model, external validity would require demonstration in an alternative setting. We have grouped 'other SBI' into a single outcome category. It would be preferable to model outcomes such as septicaemia and meningitis separately, but the infrequency of these outcomes makes this challenging. A pragmatic response is to advocate further diagnostic testing (including urgent urine or CSF microscopy) in children considered at high risk of 'other SBIs'.

Diagnostic studies with imperfect reference standards require a pragmatic approach to determine outcomes. An established approach to this is to use pre-defined composite reference standards as we have done. The universal application of respiratory viral assays may have yielded additional evidence upon which to base classification but such testing was undertaken at the discretion of the clinical team, and not applied systematically. Our use of a radiological diagnosis of 'pneumonia', despite its limitations, is common in this setting.[15, 33] We included a category of 'probable SBI' to account for the lack of sensitivity of conventional diagnostic testing in children. This category accounted for only a small number

of cases (8), and was defined in advance. By establishing clear criteria for each outcome diagnosis, we have sought to minimise verification bias.

We studied children already considered at risk of SBI, in whom the clinical team had initiated further investigation. This unmeasured risk evaluation limits the external validity of our findings. The proportion of SBI (24%) is significantly higher than that observed in all febrile children in the ED and we agree with previous authors who have stressed the importance of diagnostics research in low-risk populations (such as all children attending the ED, or primary care).[18] Almost 80% of our sample were admitted to hospital and received antibiotics, including 60% of those who did not have SBI. Decision-making based on a low-risk threshold of 5% may reduce admissions and antibiotic use but does not (by definition) eliminate risk. Clinicians would need to combine risk evaluation with appropriate safety-netting.

Comparison with published studies

Our finding that clinical variables such as hypoxia, abnormal respiratory findings, irritability and dehydration increase the probability of SBI is consistent with similar studies.[2, 13, 16] We failed to demonstrate the value of more subjective assessments, such as 'ill appearance', and 'parental concern', though for each there was a significant problem of missing data.

Next steps:

Our results support a growing body of research to suggest that risk prediction models improve the identification of SBI in the children's ED. Such models have yet to translate into improved clinical decision-making. Two recent impact studies challenge the assumption that accurate risk prediction will necessarily improve decision-making. In the first, the use of the 'Lab Score' - a decision rule combining CRP, PCT and urinalysis - failed to reduce antibiotic

prescriptions in children in the ED.[34] A second evaluated the use of the Nijman risk prediction model to guide decisions, and no impact on antibiotic prescribing, or hospital admission was observed[35].

Future impact studies need to evaluate the behaviours associated with decision-making. This has been of considerable importance in evaluating interventions to rationalise antibiotic prescribing[36]. In order to translate estimates of risk into safe clinical decisions and improve the management of children in the ED, it will be necessary to involve clinicians and families . The risk thresholds we have proposed are not yet established in the context of SBI in the children's ED, and more work is necessary to determine whether they, and the clinical decisions they guide, are appropriate.

CONCLUSION

A diagnostic model combining clinical and biomarker variables discriminated well between pneumonia, other SBIs and no SBI in febrile children of all ages in the ED. External validation of a previously derived risk model yielded encouraging diagnostic accuracy and was improved by the addition of PCT and Resistin. Future work should establish the value of decision rules based upon risk prediction models in robust impact studies. Such studies must address the complex behaviours associated with clinical decisions in order to yield clinical benefit.

Acknowledgments

We thank all children and their parents for agreeing to participate, and we thank the significant number of ED and paediatric clinical and nursing staff working in the Alder Hey Children's ED for their substantial contribution to the study. In particular we thank the children of the Medicine's for Children Research Network Young Person's Advisory Group. We thank Professor Matthew Peak and Ms Dot Lambert for logistical support from the Alder Hey Research Department. We thank Ms Elaine Hanmer, Ms Laura Medway, and Ms Gemma Boydell for informatics support, and Mrs Sarah Olsen for administrative support in the running of the study.

The University of Liverpool acknowledges the support of the National Institute for Health Research, through the Comprehensive Clinical Research Network

1. Sands R, Shanmugavadivel D, Stephenson T, Wood D. Medical problems presenting to paediatric emergency departments: 10 years on. *Emerg Med J*. 2012;29(5):379-382.
2. Craig JC, Williams GJ, Jones M, Codarini M, Macaskill P, Hayen A, et al. The accuracy of clinical symptoms and signs for the diagnosis of serious bacterial infection in young febrile children: prospective cohort study of 15 781 febrile illnesses. *BMJ*. 2010;340:c1594.
3. Thompson MJ, Ninis N, Perera R, Mayon-White R, Phillips C, Bailey L, et al. Clinical recognition of meningococcal disease in children and adolescents. *Lancet*. 2006;367(9508):397-403.
4. Launay E, Gras-Le Guen C, Martinot A, Assathiany R, Blanchais T, Mourdi N, 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(4):469-474.
5. Inwald DP, Tasker RC, Peters MJ, Nadel S, Paediatric Intensive Care Society Study G. Emergency management of children with severe sepsis in the United Kingdom: the results of the Paediatric Intensive Care Society sepsis audit. *Arch Dis Child*. 2009;94(5):348-353.
6. Koshy E, Murray J, Bottle A, Sharland M, Saxena S. Impact of the seven-valent pneumococcal conjugate vaccination (PCV7) programme on childhood hospital admissions for bacterial pneumonia and empyema in England: national time-trends study, 1997-2008. *Thorax*. 2010;65(9):770-774.
7. Whitney CG, Farley MM, Hadler J, Harrison LH, Bennett NM, Lynfield R, et al. Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *N Engl J Med*. 2003;348(18):1737-1746.
8. Balmer P, Borrow R, Miller E. Impact of meningococcal C conjugate vaccine in the UK. *J Med Microbiol*. 2002;51(9):717-722.
9. Ladhani S, Pebody RG, Ramsay ME, Lamagni TL, Johnson AP, Sharland M. Continuing impact of infectious diseases on childhood deaths in England and Wales, 2003-2005. *Pediatr Infect Dis J*. 2010;29(4):310-313.
10. Saxena S, Bottle A, Gilbert R, Sharland M. Increasing short-stay unplanned hospital admissions among children in England; time trends analysis '97-'06. *PLoS One*. 2009;4(10):e7484.
11. Gill PJ, Goldacre MJ, Mant D, Heneghan C, Thomson A, Seagroatt V, et al. Increase in emergency admissions to hospital for children aged under 15 in England, 1999-2010: national database analysis. *Arch Dis Child*. 2013;98(5):328-334.
12. Irwin AD, Wickenden J, Le Doare K, Ladhani S, Sharland M. Supporting decisions to increase the safe discharge of children with febrile illness from the emergency department: a systematic review and meta-analysis. *Arch Dis Child*. 2016;101(3):259-266.
13. Van den Bruel A, Haj-Hassan T, Thompson M, Buntinx F, Mant D, European Research Network on Recognising Serious Infection i. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet*. 2010;375(9717):834-845.
14. Van den Bruel A, Thompson MJ, Haj-Hassan T, Stevens R, Moll H, Lakhanpaul M, et al. Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review. *BMJ*. 2011;342:d3082.
15. Brent AJ, Lakhanpaul M, Thompson M, Collier J, Ray S, Ninis N, et al. Risk score to stratify children with suspected serious bacterial infection: observational cohort study. *Arch Dis Child*. 2011;96(4):361-367.
16. Nijman RG, Vergouwe Y, Thompson M, van Veen M, van Meurs AH, van der Lei J, et al. Clinical prediction model to aid emergency doctors managing febrile children at risk of serious bacterial infections: diagnostic study. *BMJ*. 2013;346:f1706.
17. Irwin AD, Marriage F, Mankhambo LA, Group IS, Jeffers G, Kolamunnage-Dona R, et al. Novel biomarker combination improves the diagnosis of serious bacterial infections in Malawian children. *BMC Med Genomics*. 2012.

18. Thompson M, Van den Bruel A, Verbakel J, Lakhanpaul M, Haj-Hassan T, Stevens R, et al. Systematic review and validation of prediction rules for identifying children with serious infections in emergency departments and urgent-access primary care. *Health Technol Assess*. 2012;16(15):1-100.
19. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med*. 2003;138(1):40-44.
20. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594.
21. Investigation of Blood Cultures (for Organisms other then *Mycobacterium* species). UK Standards for Microbiology Investigations. Public Health England; 2014.
22. Corless CE, Guiver M, Borrow R, Edwards-Jones V, Fox AJ, Kaczmarski EB. Simultaneous detection of Neisseria meningitidis, Haemophilus influenzae, and Streptococcus pneumoniae in suspected cases of meningitis and septicemia using real-time PCR. *J Clin Microbiol*. 2001;39(4):1553-1558.
23. Meningococcal Reference U, Gray SJ, Trotter CL, Ramsay ME, Guiver M, Fox AJ, et al. Epidemiology of meningococcal disease in England and Wales 1993/94 to 2003/04: contribution and experiences of the Meningococcal Reference Unit. *J Med Microbiol*. 2006;55(Pt 7):887-896.
24. Rutjes AW, Reitsma JB, Coomarasamy A, Khan KS, Bossuyt PM. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess*. 2007;11(50):iii, ix-51.
25. Galetto-Lacour A, Zamora SA, Andreola B, Bressan S, Lacroix L, Da Dalt L, et al. Validation of a laboratory risk index score for the identification of severe bacterial infection in children with fever without source. *Arch Dis Child*. 2010;95(12):968-973.
26. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2013.
27. van Buuren S. Multiple imputation of discrete and continuous data by fully conditional specification. *Stat Methods Med Res*. 2007;16(3):219-242.
28. Hastie T, Tibshirani R. Generalized additive models for medical research. *Stat Methods Med Res*. 1995;4(3):187-196.
29. M S. epiR: Tools for the analysis of Epidemiological Data. R package version 0.9-62 ed2015.
30. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77.
31. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837-845.
32. Van Hoorde K, Vergouwe Y, Timmerman D, Van Huffel S, Steyerberg EW, Van Calster B. Assessing calibration of multinomial risk prediction models. *Stat Med*. 2014;33(15):2585-2596.
33. Nijman RG, Zwinkels RL, van Veen M, Steyerberg EW, van der Lei J, Moll HA, et al. Can urgency classification of the Manchester triage system predict serious bacterial infections in febrile children? *Arch Dis Child*. 2011;96(8):715-722.
34. Lacroix L, Manzano S, Vandertuin L, Hugon F, Galetto-Lacour A, Gervaix A. Impact of the lab-score on antibiotic prescription rate in children with fever without source: a randomized controlled trial. *PLoS One*. 2014;9(12):e115061.
35. de Vos-Kerkhof E, Nijman RG, Vergouwe Y, Polinder S, Steyerberg EW, van der Lei J, et al. Impact of a clinical decision model for febrile children at risk for serious bacterial infections at the emergency department: a randomized controlled trial. *PLoS One*. 2015;10(5):e0127620.

524 36. Little P, Stuart B, Francis N, Douglas E, Tonkin-Crine S, Anthierens S, et al. Effects of internet-
525 based training on antibiotic prescribing rates for acute respiratory-tract infections: a
526 multinational, cluster, randomised, factorial, controlled trial. *Lancet*. 2013;382(9899):1175-
527 1182.

528

	Overall n=1101		Pneumonia n=108		Other SBI n=156		No SBI n=837	
Demographics	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Age	2.39	0.88-5.73	3.51†	1.60-6.29	2.28	0.43-7.54	2.21	0.92-5.35
	Proportion	95%CI	Proportion	95%CI	Median	IQR	Proportion	95%CI
Male sex	0.55	0.52-0.58	0.48	0.39-0.57	0.59	0.51-0.66	0.56	0.52-0.59
PMH	0.31	0.28-0.34	0.47†	0.38-0.57	0.26	0.19-0.33	0.30	0.27-0.33
Clinical variables	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Temperature	37.8	37.0-38.6	37.9*	37.1-38.9	38.0*	37.2-38.8	37.7	36.9-38.6
Heart Rate	140	121-166	147*	132-170	148*	122-175	139	120-163
Respiratory Rate	30	24-38	38†	28-48	30	24-38	28	24-36
Biomarkers	Median	IQR	Median	IQR	Median	IQR	Median	IQR
CRP / mg/l	19.6	5.8-54.0	49.0†	21.1-119	68.3†	28.9-137	14.3	4.0-36.5
WCC / $x10^9$/l	11.5	7.9-15.8	11.8*	8.4-18.5	15.0†	10.9-20.5	10.8	7.7-14.7
Neutrophils / $x10^9$/l	6.9	3.8-10.8	8.0†	4.8-13.4	10.0†	5.9-14.8	6.2	3.4-9.7
NGAL / ng/l	77.1	52.5-121	92.1†	65.9-162	120†	74.4-170	69.7	49.5-103
PCT / µg/l	0.23	0.10-0.83	0.49†	0.12-2.85	1.10†	0.15-5.85	0.18	0.09-0.53
Resistin / ng/l	40.3	21.1-73.4	67.3†	31.4-107	60.6†	29.7-113	35.7	19.8-64.3
Outcomes	Median	IQR	Median	IQR	Median	IQR	Median	IQR
Length of stay/ days	2	0-3	3†	2-6	4.5†	2-7	1	0-2
	n (%)	95%CI	n (%)	95%CI	n (%)	95%CI	n (%)	95%CI
Antibiotic use	855 (78)	75-80	108† (100)	96-100	156† (100)	97-100	509 (61)	57-64
Hospital admission	844 (77)	74-79	102† (94)	88-98	148† (95)	90-98	516 (62)	58-65
PICU	19 (1.73)	1.11-2.68	5* (4.63)	2.00-10.4	10*(6.41)	3.52-11.4	4 (0.48)	0.19-1.22
Mortality	1 (0.09)	0.01-0.51	0	0-3.40	1 (0.65)	0.12-3.55	0	0-0.46

Table 1: Characteristics of study subjects. IQR – interquartile range. Statistical comparisons between Pneumonia, or Other SBI and No SBI. Continuous data were compared by means of the Kruskal Wallis test, proportions were compared by means of the Pearson's Chi squared statistic. Rare events such as admission to PICU or death were compared by means of a Monte Carlo simulation. †$p<0.001$ *$p<0.05$

Updated Nijman model: Pneumonia												
	Sensitivity	95% CI	Specificity	95% CI	PPV	95% CI	NPV	95% CI	PLR	95% CI	NLR	95% CI
2.5%	0.93	(0.86 - 0.97)	0.51	(0.47 - 0.55)	0.20	(0.16 - 0.24)	0.98	(0.96 - 0.99)	1.91	(1.75 - 2.08)	0.14	(0.07 - 0.28)
5%	0.89	(0.81 - 0.94)	0.70	(0.67 - 0.73)	0.28	(0.23 - 0.33)	0.98	(0.97 - 0.99)	2.98	(2.63 - 3.37)	0.16	(0.09 - 0.27)
10%	0.81	(0.79 - 0.88)	0.82	(0.79 - 0.84)	0.36	(0.30 - 0.43)	0.97	(0.95 - 0.98)	4.41	(3.72 - 5.23)	0.24	(0.16 - 0.35)
20%	0.69	(0.60 - 0.78)	0.89	(0.87 - 0.91)	0.45	(0.38 - 0.53)	0.96	(0.94 - 0.97)	6.46	(5.12 - 8.14)	0.34	(0.26 - 0.46)
30%	0.60	(0.92 - 0.96)	0.94	(0.92 - 0.96)	0.58	(0.48 - 0.67)	0.95	(0.93 - 0.96)	10.5	(7.66 -14.4)	0.42	(0.33 - 0.53)
Other SBI												
2.5%	0.99	(0.96 - 1.0)	0.09	(0.07 - 0.11)	0.17	(0.15 - 0.19)	0.99	(0.93 - 1.00)	1.09	(1.06 - 1.12)	0.07	(0.01 - 0.50)
5%	0.97	(0.93 - 0.99)	0.24	(0.21 - 0.27)	0.19	(0.17 - 0.22)	0.98	(0.95 - 0.99)	1.28	(1.22 - 1.34)	0.13	(0.06 - 0.31)
10%	0.83	(0.77 - 0.89)	0.58	(0.55 - 0.62)	0.27	(0.23 - 0.31)	0.95	(0.93 - 0.97)	1.99	(1.79 - 2.21)	0.29	(0.20 - 0.41)
20%	0.56	(0.48 - 0.64)	0.89	(0.87 - 0.91)	0.49	(0.41 - 0.56)	0.92	(0.90 - 0.93)	5.13	(4.04 - 6.50)	0.49	(0.41 - 0.59)
30%	0.40	(0.32 - 0.48)	0.95	(0.94 - 0.97)	0.61	(0.51 - 0.70)	0.89	(0.87 - 0.91)	8.31	(5.80 - 11.9)	0.63	(0.56 - 0.72)
Extended Nijman model (including PCT and Resistin): Pneumonia												
2.5%	0.94	(0.87 - 0.97)	0.52	(0.49 - 0.56)	0.20	(0.17 - 0.24)	0.98	(0.97 - 0.99)	1.96	(1.79 - 2.13)	0.12	(0.06 - 0.25)
5%	0.92	(0.85 - 0.96)	0.69	(0.66 - 0.72)	0.28	(0.23 - 0.33)	0.98	(0.97 - 0.99)	2.96	(2.64 - 3.33)	0.12	(0.06 - 0.23)
10%	0.85	(0.77 - 0.91)	0.82	(0.79 - 0.84)	0.38	(0.31 - 0.44)	0.98	(0.96 - 0.99)	4.66	(3.96 - 5.49)	0.18	(0.12 - 0.29)
20%	0.70	(0.61 - 0.79)	0.89	(0.87 - 0.91)	0.46	(0.39 - 0.54)	0.96	(0.94 - 0.97)	6.69	(5.30 - 8.44)	0.33	(0.25 - 0.44)
30%	0.62	(0.52 - 0.71)	0.94	(0.92 - 0.95)	0.56	(0.47 - 0.65)	0.95	(0.93 - 0.96)	9.99	(7.38 - 13.5)	0.4	(0.32 - 0.52)
Other SBI												
2.5%	0.97	(0.94 - 0.99)	0.18	(0.15 - 0.20)	0.18	(0.16 - 0.21)	0.97	(0.93 - 0.99)	1.18	(1.14 - 1.23)	0.15	(0.05 - 0.39)
5%	0.92	(0.86 - 0.95)	0.40	(0.37 - 0.44)	0.22	(0.19 - 0.26)	0.96	(0.94 - 0.98)	1.54	(1.43 - 1.65)	0.21	(0.12 - 0.35)
10%	0.85	(0.79 - 0.90)	0.61	(0.58 - 0.65)	0.29	(0.25 -0.34)	0.96	(0.94 - 0.97)	2.21	(1.98 - 2.46)	0.24	(0.16 - 0.35)
20%	0.70	(0.62 - 0.77)	0.86	(0.83 - 0.88)	0.48	(0.41 - 0.55)	0.94	(0.92 - 0.95)	4.96	(4.07 - 6.03)	0.35	(0.28 - 0.45)
30%	0.53	(0.45 - 0.61)	0.94	(0.92 - 0.95)	0.61	(0.52 - 0.69)	0.91	(0.89 - 0.93)	8.40	(6.23 - 11.3)	0.50	(0.42 - 0.59)

Table 2: Performance characteristics of the updated (top) and the extended Nijman models (bottom) including the biomarkers Procalcitonin and Resistin (bottom) PPV: positive predictive value, NPV: negative predictive value, PLR: positive likelihood ratio, NLR: negative likelihood ratio.

Outcome category	Updated				Extended				n
	Rule-out	Inter-mediate	Rule-in		Rule-out	Inter-mediate	Rule-in		
			Pneu	Other			Pneu	Other	
No SBI	269	355	76	137	300	352	74	111	837
Pneumonia	6	19	70	13	5	16	71	16	108
Other SBI	13	29	7	107	9	33	7	107	156
Total	288	403	153	257	314	401	152	234	1101

Table 3: Outcomes according to risk classification for the updated and extended models. SBI was considered 'ruled-out' if the predicted probabilities of both pneumonia ("Pneu") and other SBI ("Other") were <5%, while SBI was considered 'ruled-in' if the probability of either outcome was >20%. All other subjects were considered to be at intermediate risk.

Figure legends:

Figure 1: Flow diagram of the study. PID: Primary immunodeficiency. ED: Emergency Department. SBI: Serious Bacterial Infection. Excluded children with a 'clinical reference standard' are explained in the text.

Figure 2: Parametric nominal calibration plot of predicted risks and observed outcomes in the validation set.

Figure 3: Parametric nominal calibration plot of the original Nijman model on external validation.

Figure 4: Parametric nominal calibration plot of the Nijman model with co-efficients re-fitted to the validation dataset.

	Derivation group (n=532)			Validation group (n=569)		
	No SBI (401)	Pneu (63)	Other SBI (68)	No SBI (436)	Pneu (45)	Other SBI (88)
FBC	391 (98)	62 (98)	68 (100)	427 (98)	45 (100)	87 (99)
Urinalysis	99 (25)	10 (16)	17 (25)	97 (22)	7 (16)	25 (28)
Blood culture	257 (64)	54 (86)	56 (82)	280 (64)	42 (93)	77 (88)
CXR	168 (42)	61 (97)	24 (35)	195 (45)	44 (98)	38 (43)

Supplementary Table 1: Number (%) of diagnostic tests performed in each group. FBC: Full blood count, CXR: Chest X-ray

Diagnosis		Criteria
Pneumonia		Respiratory symptoms and signs and focal consolidation on X-ray reported by a paediatric radiologist.
Other SBI		
	Bacteraemia	Identification of a significant bacterial pathogen in blood using culture or molecular methods.
	Urinary tract infection	Growth of a single bacterial urinary tract pathogen at $\geq 10^5$ colony-forming units/ml in a normally sterile urine sample in the context of clinical signs of systemic involvement.
	Meningitis	Identification of a bacterial pathogen in CSF using culture or molecular methods, or clinical meningitis plus a cerebrospinal fluid polymorphonuclear leucocytosis in the absence of an alternative aetiological diagnosis.
	Osteomyelitis	Clinical signs, and radiological confirmation or identification of a pathogen in the bloodstream.
	Septic arthritis	Isolation of a bacterial pathogen from a joint.
	Probable SBI	Prolonged admission, and administration of intravenous antibiotics beyond 72h despite negative culture results.

Supplementary Table 2: Pre-defined criteria for the diagnosis of SBI[14, 15, 25]

Variable	Observations	Missing	
		n	%
Neck stiffness	652	449	40.8
Normal air entry	1069	32	2.9
Chest clear	1069	32	2.9
Bulging fontanelle	246	855	77.7
Rash	1009	92	8.4
Abdominal pain	306	795	72.2
Parental concern	159	942	85.6
History of myalgia	151	950	86.3
Irritability	256	845	76.7
Abnormal ENT signs	921	180	16.3
Heart rate	1058	43	3.9
History of diarrhoea	899	202	18.3
Respiratory rate	907	194	17.6
Duration of fever (day)	1101	0	0.0
Temperature	1092	9	0.8
Prolonged Capillary Refill (>2s)	909	192	17.4
History of dysuria	270	831	75.5
Dehydration	480	621	56.4
Pallor	469	632	57.4
Comorbidity	1101	0	0.0
History of drowsiness	295	806	73.2
Prior antibiotics	1097	4	0.4
Wheeze	1074	27	2.5
Ill appearance	108	993	90.2
History of chest pain	118	983	89.3
Chest crackles	1071	30	2.7
History of cough	847	254	23.1
Hypoxia (Sats <92%)	963	138	12.5
Decreased Breath Sounds	1068	33	3.0
Increased Work of Breathing	1071	30	2.7

Supplementary Table 3: Proportion of missing data for each observed clinical variable.

Characteristics	Derivation		Validation
	Erasmus (n=1750)	Haga-Juliana (n=967)	Liverpool (n=1101)
Median age/years (IQR)	1.8 (0.9-3.7)	1.5 (0.7-3.2)	2.4 (0.9-5.7)
Male sex	0.57	0.55	0.55
Median (IQR) duration of fever (days)	n=1185	n=807	n=1052
	2 (1-3)	2 (1-3)	2 (0-3)
Median temperature/°C (IQR)	n=1699	n=967	n=1092
	39.0 (38.3-39.7)	38.8 (38.3-39.4)	37.8 (37.0-38.6)
Median heart rate (IQR)	n=914	n=473	n=1058
	140 (120-160)	156 (140-172)	140 (121-166)
Median respiratory rate (IQR)	n=819	n=183	n=907
	36 (28-48)	48 (40-60)	30 (24-38)
Oxygen saturations <94%	n=914	n=473	n=963
	41	43	82
Cap refill time >3s	n=914	n=473	n=909
	96	9	40
Increased work of breathing	n=914	n=473	n=1071
	97	108	218
Ill appearance	n=914	n=473	n=108
	520	317	64
Median CRP (IQR)	n=780	n=317	n=1072
	21 (7-54)	22 (7-56)	20 (6-54)
Outcomes			
SBI/ n (%)	222 (13)	119 (12)	264 (24)
Pneumonia	105 (6)	66 (7)	108 (10)
UTI	50 (3)	38 (4)	58 (5)
Septicaemia/meningitis	21 (1)	1 (0)	49 (4)
Other	46 (3)	14 (1)	49 (4)

Supplementary table 4: Comparison of characteristics of study participants used in the derivation of the Nijman risk prediction model, and the Liverpool validation group

Biomarkers	n	Pneumonia			Other SBI		
		OR	LCI	UCI	OR	LCI	UCI
Procalcitonin	1034	1.22	1.15	1.29	1.23	1.16	1.30
Neutrophils	1059	1.09	1.06	1.13	1.12	1.09	1.15
WCC	1059	1.05	1.02	1.08	1.08	1.06	1.11
CRP	1072	1.02	1.01	1.02	1.02	1.02	1.02
Resistin	1045	1.01	1.00	1.01	1.01	1.01	1.01
NGAL	1046	1.00	1.00	1.01	1.01	1.00	1.01
Blood glucose	123	0.78	0.54	1.12	1.03	0.89	1.20
Lactate	167	0.67	0.42	1.09	1.12	0.84	1.50

Supplementary Table 5: Odds ratios of biomarker variables significantly associated with pneumonia and other SBI in univariate multinomial regression analysis. OR: Odds ratio. LCI: Lower (95%) confidence interval. UCI: Upper (95%) confidence interval.

	Pneumonia				Other SBI			
	Est	OR	LCI	UCI	Est	OR	LCI	UCI
(Intercept)	-2.516	0.081	0.025	0.260	-2.779	0.062	0.016	0.239
CRP / mg/l (<30)	0.025	1.025	0.990	1.060	0.045	1.046	1.011	1.081
CRP / mg/l (>30)	0.010	1.010	1.003	1.018	0.012	1.012	1.005	1.019
Respiratory rate	0.047	1.048	1.021	1.076	0.009	1.009	0.980	1.039
PCT / µg/l	0.173	1.189	1.079	1.310	0.168	1.183	1.074	1.303
Normal air entry	-2.387	0.092	0.046	0.182	0.240	1.271	0.514	3.142
Resistin / ng/ml	0.003	1.003	0.999	1.008	0.004	1.004	1.000	1.007

Supplementary table 6: Summary output of the derived polynomial models for the diagnosis of pneumonia and other SBIs. Est: Estimate of the regression co-efficient.

Outcome diagnosis	n	Updated model			Extended model		
		No SBI	Pneumonia	Other SBI	No SBI	Pneumonia	Other SBI
No SBI	837	801	18	18	807	17	13
Pneumonia	108	61	44	3	58	45	5
Other SBI	156	106	2	48	89	2	65

Supplementary table 7: Observed and predicted outcomes as determined by the highest risk category predicted by the updated and extended multinomial models.