

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

Jonathan C Craig, senior staff specialist,^{1,2} Gabrielle J Williams, clinical researcher,¹ Mike Jones, senior lecturer,^{1,3} Miriam Codarini, PhD student,¹ Petra Macaskill, associate professor of biostatistics,¹ Andrew Hayen, senior lecturer in biostatistics,¹ Les Irwig, professor of epidemiology,¹ Dominic A Fitzgerald, senior staff specialist,⁴ David Isaacs, senior staff specialist,⁵ Mary McCaskill, senior staff specialist⁶

¹Screening and Test Evaluation Program, School of Public Health, University of Sydney, Sydney, Australia

²Department of Nephrology, The Children's Hospital at Westmead, Westmead, Australia

³Psychology Department, Macquarie University, North Ryde, Australia

⁴Department of Respiratory Medicine, The Children's Hospital at Westmead, Westmead, Australia

⁵Department of Infectious Disease, The Children's Hospital at Westmead, Westmead, Australia

⁶Department of Emergency Medicine, The Children's Hospital at Westmead, Westmead, Australia

Correspondence to: G J Williams gabrielw4@chw.edu.au

Cite this as: *BMJ* 2010;**340**:c1594 [doi:10.1136/bmj.c1594](https://doi.org/10.1136/bmj.c1594)

ABSTRACT

Objectives To evaluate current processes by which young children presenting with a febrile illness but suspected of having serious bacterial infection are diagnosed and treated, and to develop and test a multivariable model to distinguish serious bacterial infections from self limiting non-bacterial illnesses.

Design Two year prospective cohort study.

Setting The emergency department of The Children's Hospital at Westmead, Westmead, Australia.

Participants Children aged less than 5 years presenting with a febrile illness between 1 July 2004 and 30 June 2006.

Intervention A standardised clinical evaluation that included mandatory entry of 40 clinical features into the hospital's electronic record keeping system was performed by physicians. Serious bacterial infections were confirmed or excluded using standard radiological and microbiological tests and follow-up.

Main outcome measures Diagnosis of one of three key types of serious bacterial infection (urinary tract infection, pneumonia, and bacteraemia), and the accuracy of both our clinical decision making model and clinician judgment in making these diagnoses.

Results We had follow-up data for 93% of the 15 781 instances of febrile illnesses recorded during the study period. The combined prevalence of any of the three infections of interest (urinary tract infection, pneumonia, or bacteraemia) was 7.2% (1120/15 781, 95% confidence interval (CI) 6.7% to 7.5%), with urinary tract infection the diagnosis in 543 (3.4%) cases of febrile illness (95% CI 3.2% to 3.7%), pneumonia in 533 (3.4%) cases (95% CI 3.1% to 3.7%), and bacteraemia in 64 (0.4%) cases (95% CI 0.3% to 0.5%). Almost all (94%) of the children with serious bacterial infections had the appropriate test (urine culture, chest radiograph, or blood culture). Antibiotics were prescribed acutely in 66% (359/543) of children with urinary tract infection, 69% (366/533) with pneumonia, and 81% (52/64) with bacteraemia. However, 20% (2686/13 557) of children without bacterial infection were also prescribed antibiotics. On the basis of the data from

the clinical evaluations and the confirmed diagnosis, a diagnostic model was developed using multinomial logistic regression methods. Physicians' diagnoses of bacterial infection had low sensitivity (10-50%) and high specificity (90-100%), whereas the clinical diagnostic model provided a broad range of values for sensitivity and specificity.

Conclusions Emergency department physicians tend to underestimate the likelihood of serious bacterial infection in young children with fever, leading to undertreatment with antibiotics. A clinical diagnostic model could improve decision making by increasing sensitivity for detecting serious bacterial infection, thereby improving early treatment.

INTRODUCTION

Febrile illnesses are common in children, especially in children under 5 years of age. On average, young children experience three to six febrile illnesses per year.¹ Parental anxiety over fever is common²: 20-40% of parents seek medical assessment when their child develops a fever.¹ Febrile illness is the single most common reason for young children to be seen by primary care practitioners³ and to present to emergency departments for acute care.^{4,5} In the United States alone, these figures translate to around 25 million episodes of febrile illness in young children managed by clinicians each year.⁶

In most cases, the cause of a febrile illness is a self limiting, presumed viral, infection, but 5-10% of febrile children have serious bacterial infections such as pneumonia, urinary tract infection, meningitis, bacteraemia, or bone or joint infection (which usually present with fever).⁷⁻⁹ These conditions can be difficult to distinguish from viral infections and benefit from early antibiotic therapy. The consequences of a delayed or missed diagnosis can be serious and, occasionally, fatal.¹⁰

When a child presents with fever, the decision to commence antibiotics is based on a thorough clinical evaluation by the treating physician, supplemented by rapid tests such as a urine dipstick. The reference

standard for most serious bacterial diseases is a positive microbiological culture from a sterile sample such as blood, urine, or cerebrospinal fluid. Obtaining a result takes 24-48 hours, however, so culture cannot be used to guide management in the acute setting.

Many clinical scoring systems have been developed for predicting bacterial infection in febrile children,¹¹⁻¹³ but they are not widely used in practice. Most scores apply to young infants only, are not sufficiently sensitive for routine use, and have been developed from studies that were small, did not verify bacterial disease in all children, and did not adjust for correlated clinical features. A recent and well designed Belgian study¹⁴ of children presenting for assessment of an acute illness suggests clinical signs have value in excluding serious disease, but the limited sample size of children with serious infection (n=31) creates some doubt over the precision and statistical power of the study.

Clinical assessment, including history taking for symptoms and physical examination for signs, is regarded as the cornerstone of clinical practice,¹⁵ but has rarely been evaluated. The few evaluations that have been conducted are usually limited to individual items for a specific condition and rarely cover a combined set of symptoms and signs, as used in clinical practice. Recently, the problems of missed or delayed diagnosis have been emphasised, but no clear solutions provided.¹⁶ The aims of this study were to evaluate the current diagnostic processes by which febrile young children with suspected serious bacterial infection are investigated and treated, and to develop and test the accuracy

of a clinical diagnostic model in which the essential elements of clinical evaluation are statistically combined.

METHODS

At the start of the study, an electronic template was introduced to the record keeping system in the emergency department of The Children's Hospital at Westmead, which standardised the mandatory entry of 40 symptoms and signs for all children presenting with febrile illness. Doctors were required to fill in this template and were also asked to estimate the probability that their patient had any of 10 potential diagnoses. There was no study mandated investigation or treatment algorithm.

Children were then diagnosed as having a serious bacterial infection (urinary tract infection, pneumonia, and bacteraemia), clinically diagnosed infection, or no bacterial infection using standard radiological and microbiological tests; this classification was verified by a final diagnosis committee. All eligible febrile children were followed up until they fulfilled the case definition for serious bacterial infection or until the fever had resolved for ≥ 24 hours.

A model was then developed according to the clinical symptoms and signs data in the electronic records and the case definition. A preliminary analysis was used to select variables for inclusion in the multinomial model. The selected variables were then fitted jointly in a multinomial logistic regression model, and variables that were no longer statistically significant were removed using backward elimination.

The performance of the model was assessed for each type of serious bacterial infection by constructing a receiver operating characteristic (ROC) curve. Finally, the clinical diagnoses estimated by clinicians were compared against the model to test the accuracy of clinician judgment when attempting to identify bacterial infection in children with fever.

Recruitment

Consecutive children under 5 years of age presenting to the emergency department of The Children's Hospital at Westmead with a febrile illness between 1 July 2004 and 30 June 2006 were eligible for the study. The emergency department is a large unit that sees nearly 50 000 presentations each year across the full scope of acute paediatric illness, including medical and surgical conditions. The Children's Hospital is a tertiary level university teaching hospital with about 220 beds. The composition of the clinical staff is typical of many emergency departments. The triage staff are experienced paediatric emergency nurses, with the physicians comprising permanent senior staff and 30-40 rotating junior staff at registrar and resident level.

Febrile illness was defined as any illness with one or more of the following elements: a measured axillary temperature of $\geq 38.0^{\circ}\text{C}$; parental report of a temperature of $\geq 38.0^{\circ}\text{C}$ measured at home within the previous 24 hours; a parental report that the child "felt hot" in the previous 24 hours; or a presenting problem related to fever (10th revision of the international classification of diseases,

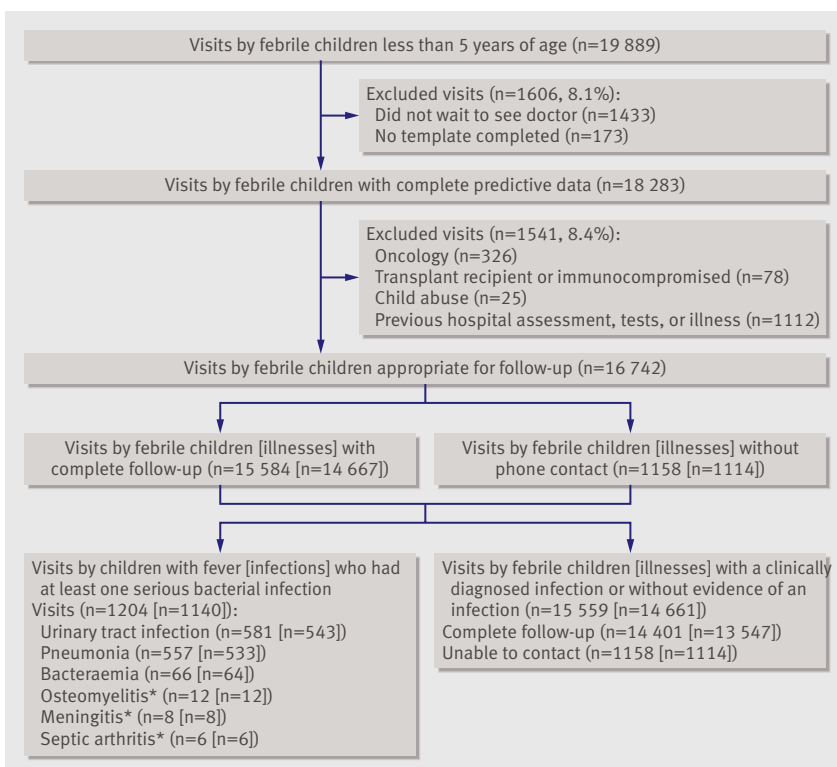


Fig 1 | Patient flow chart

Table 1 | Demographic and clinical characteristics of children presenting to the emergency department with febrile illness (n=12 807)

	Number of illnesses (n=15 781)	Proportion of illnesses (%)
Age (months)		
<3	748	4.7
>3-6	925	5.9
>6-9	1421	9.0
>9-12	1574	10.0
>12-24	4865	30.8
>24-36	2855	18.1
>36-48	1971	12.5
>48-60	1422	9.0
Gender		
Male	6967	44.1
Female	8814	55.9
Highest temperature recorded in the emergency department (°C)*		
<38	3446	21.8
38-38.9	5624	35.6
39-39.9	4994	31.6
40+	1717	10.9
Meningococcal vaccination		
Unvaccinated	5333	33.8
Vaccinated	5583	35.4
Unknown	4865	30.8
Pneumococcal vaccination		
Unvaccinated	5325	33.7
Vaccinated	5403	34.2
Unknown	5053	32.0
Illnesses per child		
1	10647	83.1
2	1595	12.5
3	404	3.2
4-7	159	1.2
Emergency department visits per illness		
1	14876	94.3
2	855	5.4
3-5	50	0.3
Provisional diagnosis on leaving the emergency department		
Viral infection	3191	20.2
Viral upper respiratory tract infection	2611	16.5
Gastroenteritis	2186	13.9
Fever (no focus)	1285	8.1
Other	974	6.2
Tonsillitis	852	5.4
Bronchitis or bronchiolitis	799	5.1
Asthma or wheeze	712	4.5
Pneumonia	630	4.0
Croup	614	3.9
Ear disease (otitis media or externa)	486	3.1
Febrile convulsion	420	2.7
Urinary tract infection	295	1.9
Allergy	244	1.5
Skin infection (cellulitis and impetigo)	158	1.0
Viral stomatitis (including herpes)	138	0.9
Lymphadenitis	62	0.4
Sepsis	38	0.2
Meningitis or meningoencephalitis	22	0.1
Scarlet fever	18	0.1
Kawasaki disease	13	0.1
Mycoplasma pneumoniae	11	0.1
Pertussis	10	0.1
Meningococcal infection	7	0.0
Osteomyelitis	5	0.0

*Highest temperature recorded by a physician. If missing, replaced by highest temperature recorded by a triage nurse. If temperature not taken by physician or nurse, child assumed to have temperature <38°C.

Australian modification codes R50, R50.0, R50.1, R50.9, and R56.0), as determined by a triage nurse.

Children were excluded if they had been transferred to The Children's Hospital from another hospital. In this setting, the details of the initial assessment at the referring hospital, which may have included a reference standard test for bacterial infection or other laboratory result, would be available, and it would not be possible for the clinician performing the clinical evaluation at The Children's Hospital to be blinded to this previous assessment. Children with cancer and transplant recipients were excluded because disease frequency, clinical evaluation, and threshold for treatment are substantially different from those of children with normal immune function. If the medical records of a child indicated the possibility of physical or sexual abuse, no follow-up contact was undertaken for ethical reasons.

The unit of analysis was an instance of febrile illness. In the scenario of multiple presentations to the emergency department with the same illness, history and clinical evaluation data from the first visit only were used in diagnostic modelling. Our case definition for "same illness" was if the child presented within 24 hours of a previous visit, or if the fever had persisted between visits to the emergency department without a fever-free period of at least 24 hours (as recorded in the medical record or by parental report at phone follow-up).

Data collection

The medical record keeping system in the emergency department at The Children's Hospital is electronic. At study commencement, an electronic template was introduced that standardised the mandatory entry of 40 clinical features (including symptoms and signs) on all eligible children, recorded prospectively and in real time by the treating physician. The template was compiled from a review of the published literature on assessment tools for febrile children^{11-13 17-20} and from data routinely recorded by emergency department staff, with expert opinion from specialists in paediatric emergency medicine, infectious disease, respiratory medicine, nephrology, and general paediatrics.

Draft versions of the assessment tool were piloted and re-piloted in the emergency department and revised by research staff over a six month period until there was agreement that all fields were essential, no essential fields were missing, and that data collection was feasible. The finalised febrile assessment template consisted of 40 stem and 25 conditional data fields and is provided in full in web table A.

This template was completed following the initial physician assessment but before test results were obtained. Any tests could be ordered during the clinical assessment on the basis of clinician judgment.

Physicians were also asked to estimate the probability that their patient had one of 10 potential diagnoses, which covered serious bacterial infections, clinically apparent infections, and viral illness. Clinicians were asked to perform this task after their initial clinical assessment of each patient and before test results

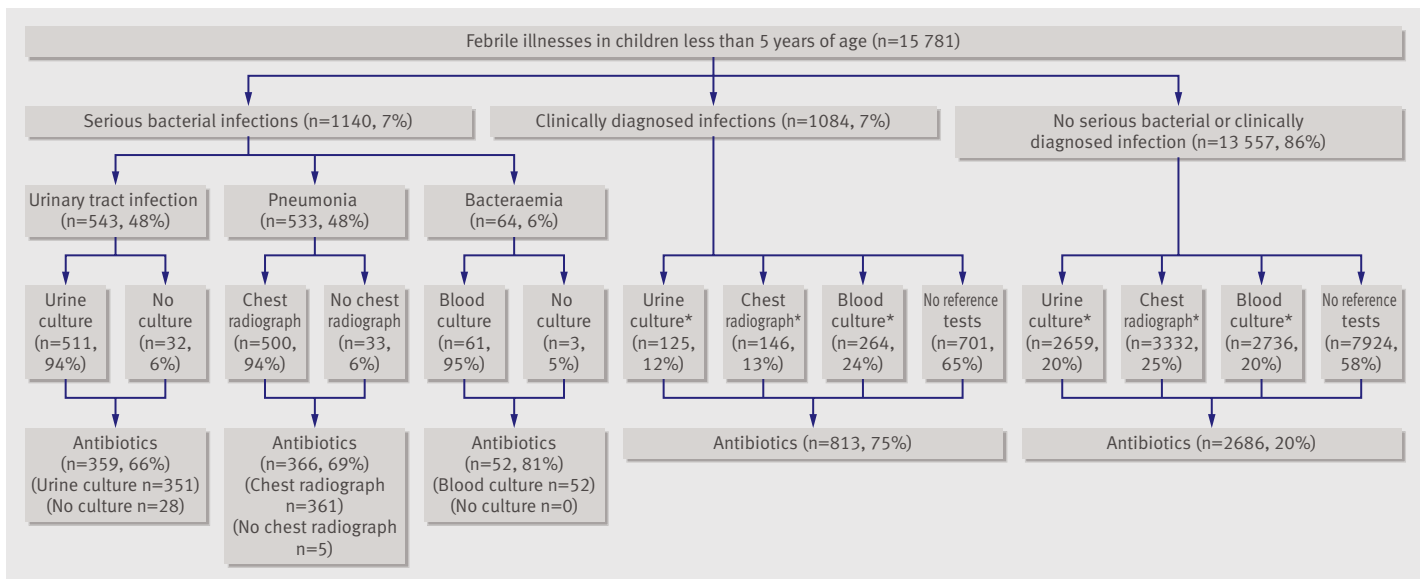


Fig 2 | Frequency of testing and antibiotic administration in the emergency department among children less than 5 years of age with febrile illness. *Children with more than one reference test are counted in each test specific box

were available. The sum of all probabilities had to equal 100% and was tallied for them.

Diagnosis of serious bacterial infection

Definite urinary tract infection was defined as $\geq 10^8$ colony forming units (cfu) per litre of a single type of organism in a voided sample, $\geq 10^7$ cfu/l of a single organism in a catheter sample, or any growth of a single organism in a suprapubic bladder tap sample.²¹ Probable urinary tract infection was defined as $\geq 10^7$ cfu/l of a single organism in a voided sample, $\geq 10^6$ cfu/l of a single organism in a catheter sample, $\geq 10^8$ cfu/l of two organisms in a voided sample, or $\geq 10^7$ cfu/l of two organisms from a catheter sample.

Definite pneumonia was defined as consolidation on chest radiograph plus any of the following signs: a positive blood culture for a pathogenic organism; culture of a pathogenic organism from pleural fluid; detection of an antigen for a pathogenic organism in a pleural fluid sample; or a positive serology result for *Mycoplasma* species. Probable pneumonia was defined as consolidation alone.

Definite bacteraemia was defined as a pure growth of a single pathogenic micro-organism on blood culture, or polymerase chain reaction positive for *Meningococcus*. Probable bacteraemia was defined as the growth of two or more types of organism (with at least one not a likely contaminant), growth of *Streptococci viridans* or other likely contaminant if the final diagnosis was endocarditis, or persistently positive blood cultures (that is, ≥ 2 blood cultures positive for same type of organism). Given the very low frequency of meningitis, osteomyelitis, and septic arthritis, we decided not to undertake diagnostic modelling for these outcomes because the risk of unreliable and potentially misleading models

was unacceptably high. Children with one of these diagnoses were not included in the analysis.

All probable cases of serious bacterial infection were reviewed by a final diagnosis committee composed of two specialist paediatricians (with experience in paediatrics, infectious disease, and respiratory medicine) and, in cases of pneumonia, a radiologist. The presence or absence of bacterial infection was decided blinded to clinical information and based on consensus. Where a reference standard test (urine or blood culture, or chest radiograph) was performed outside the hospital, the report and films were obtained for review. Uncertain results and all abnormal chest radiographs were reviewed by the final diagnosis committee.

Children were classified into three categories: serious bacterial infection; clinically diagnosed infection; and no bacterial infection. Children were regarded as having a clinically diagnosed infection when a clinical test was positive (for example, red, bulging ear drums for otitis media) and it was possible that the cause was bacterial (for example, abscess, balanitis, cellulitis, epididymitis, epiglottitis, impetigo, lymphadenitis, lymphangitis, mastoiditis, otitis media, periorbital cellulitis, scarlet fever, sinusitis, tonsillitis, or wound infection). These children were treated separately for the purpose of estimating the proportion of children who were investigated and prescribed antibiotics, but for the purposes of model development, these children were included in the group categorised as without serious bacterial infection. The classification of “no bacterial infection” was assigned to children whose reference standard tests for serious bacterial infection (chest radiograph, urine culture, blood culture) were negative or the tests were not done, but whose illness spontaneously resolved before the 10-14 day follow-up without antibiotic administration.

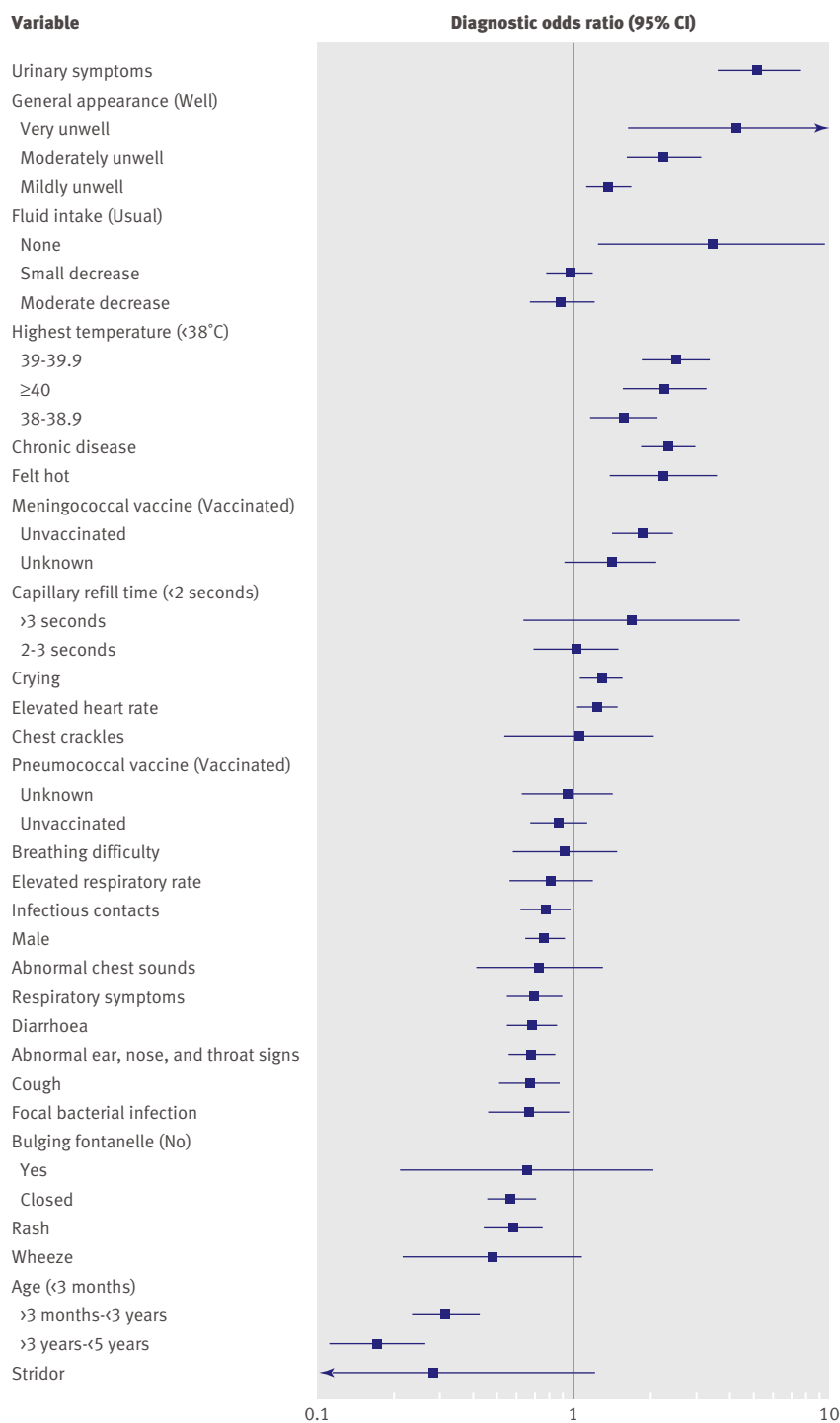


Fig 3 | Clinical indicators of urinary tract infection, displayed as logs of the odds ratios from the multinomial model

Investigation and treatment outcomes

Data on antibiotic administration in the emergency department were obtained by manual review of electronic medical records and from parental reports. In the case of a contradiction in findings between the two sources, we regarded the medical record results as the reference standard. The frequency of testing and antibiotic administration was tabulated for each category of serious bacterial infection.

Follow-up

All eligible febrile children were followed up until they fulfilled the case definition for serious bacterial infection or until the fever had resolved for ≥ 24 hours. Follow-up was achieved using multiple sources. Data collected electronically from hospital databases included test details and results; duration of admission; final diagnoses; and procedures during admission. Data fields manually extracted from review of electronic medical notes included exclusion criteria; antibiotic use in emergency department and wards; and presence of related episodes. Data collected during parental phone contact at day 10-14 included well-being of the child; presence or absence of fever; current symptoms; antibiotic use; attendance at another healthcare facility; tests performed; antibiotic prescription and diagnosis at other facility (and permission to contact the other healthcare facility and request relevant test results); and details on febrile periods if a possible related episode existed.

Sample size

We determined that data collection over a two year period would cover approximately 16 000 eligible emergency department visits. Assuming a prevalence of pneumonia and urinary tract infection of 5%, 800 cases of each were expected. On the basis of guidelines requiring at least 10 events per predictor variable in a regression model,²² the expected number of events was adequate for the number of variables recorded in the template, except for the much rarer outcome of bacteraemia.

Model development

Demographic variables and clinical symptoms and signs elicited by the treating physician in the emergency department were considered as possible indicators of serious bacterial infection. Laboratory test results were not included. The computer software SAS was used for all statistical analysis.

A multinomial logistic regression model was developed to allow the simultaneous estimation of the probability of the multiple differential diagnoses (urinary tract infection, pneumonia, bacteraemia, and no serious bacterial illness (the reference category)). This method fits simultaneous logistic regression models, each with its own intercept, to compare each of the three illnesses with the reference category. The estimated coefficients for the explanatory variables differ by illness. The unit of analysis in all models was the febrile illness; children with more than one illness were included separately for each disease.

A preliminary analysis was used to select variables for inclusion in the multinomial model. For each individual infection, forward stepwise selection of variables (using $P < 0.05$ as the criterion for statistical significance) was applied in each of 50 bootstrap samples of the data. Only variables that were selected in at least half of the resulting models were considered for inclusion in the next stage of the modelling.²³⁻²⁵ The selected variables were then fitted jointly in a multinomial logistic regression model, and variables that were

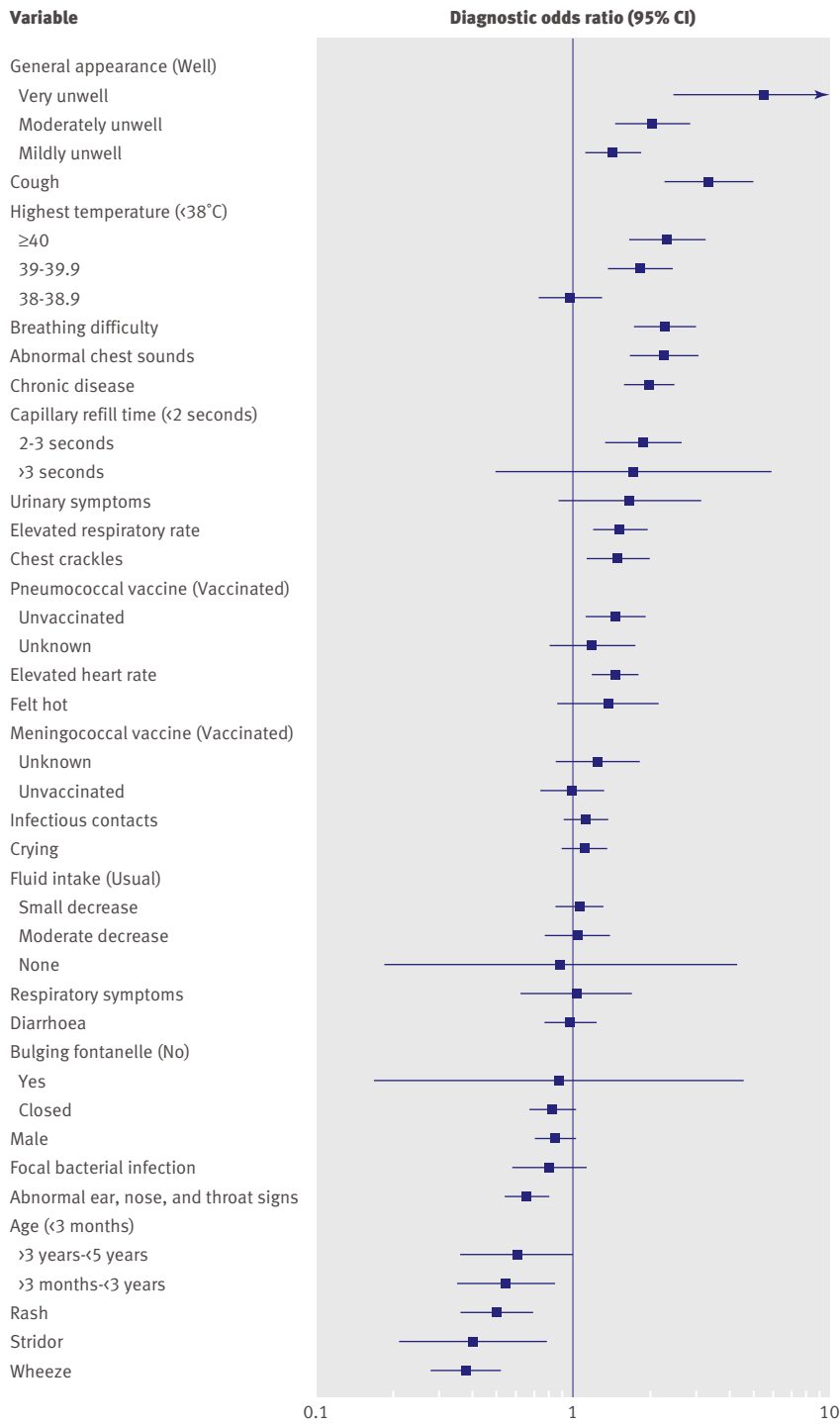


Fig 4 | Clinical indicators of pneumonia, displayed as logs of the odds ratios from the multinomial model

no longer statistically significant were removed using backward elimination.

The effect of children whose parents were unable to be contacted (loss to follow-up) was explored by the inclusion of their data in the analysis with the assumption of no serious bacterial infection as well as by their exclusion from the analyses (sensitivity analysis). There was no substantive change in the model and thus these children were included in the final model.

Given that we were modelling a paediatric population and some participants were less than 3 months of age, we considered the possibility that the strength of some diagnostic items might vary with age. However, the interactions we considered in statistical models failed to identify any evidence of effect modification by age.

Results for near patient tests, such as oxygen saturation, that are only undertaken on a select subset of children were not considered for inclusion in the model on the grounds that performance of such tests implies that a tentative diagnosis has already been made. The urine dipstick test is a widely used near patient test for the diagnosis of urinary tract infection, so revised probabilities of urinary tract infection were obtained for those children with dipstick results (details are given in web table B). These revised probabilities for urinary tract infection were used for the model.

Model assessment and validation

The diagnostic performance of the model was assessed for each type of serious bacterial infection (urinary tract infection, pneumonia, and bacteraemia) by constructing a receiver operating characteristic (ROC) curve based on the estimated probability of the relevant infection and the reference standard result for each child. The area under the curve was used as an overall indicator of test performance. The possibility of overestimation of the area under the curve was assessed by applying the model to 200 bootstrap samples of the data and computing the area under the curve for each.

The model was validated in the same setting between February 2007 and February 2009. A randomly selected group of doctors completed identical items of the template. The reference standards for serious bacterial infection were the same as those used in the main study. Estimated probabilities of each of the three serious bacterial infections were calculated, along with areas under the ROC curve.

The accuracy of early clinician judgment was compared with the model by computing the sensitivity and specificity for both across all available thresholds, compared with the reference standard for each type of serious bacterial infection.

RESULTS

Characteristics of included children and illnesses

From 1 July 2004 to 30 June 2006, there were 19 889 visits by febrile children under 5 years of age at the emergency department of The Children’s Hospital. Of these, 3147 visits were excluded for reasons shown in figure 1, which left 16 742 visits by febrile children eligible for follow-up. The presentations with fever consisted of 15 781 separate illnesses in 12 807 children. A total of 14 876 (94.3%) febrile illnesses involved one presentation to the emergency department, 855 (5.4%) involved two presentations, and 50 (0.3%) involved three to five visits. The characteristics of the children at the time of presentation are shown in table 1. Most illnesses were in children who were under 3 years of age, with a peak between 1 and

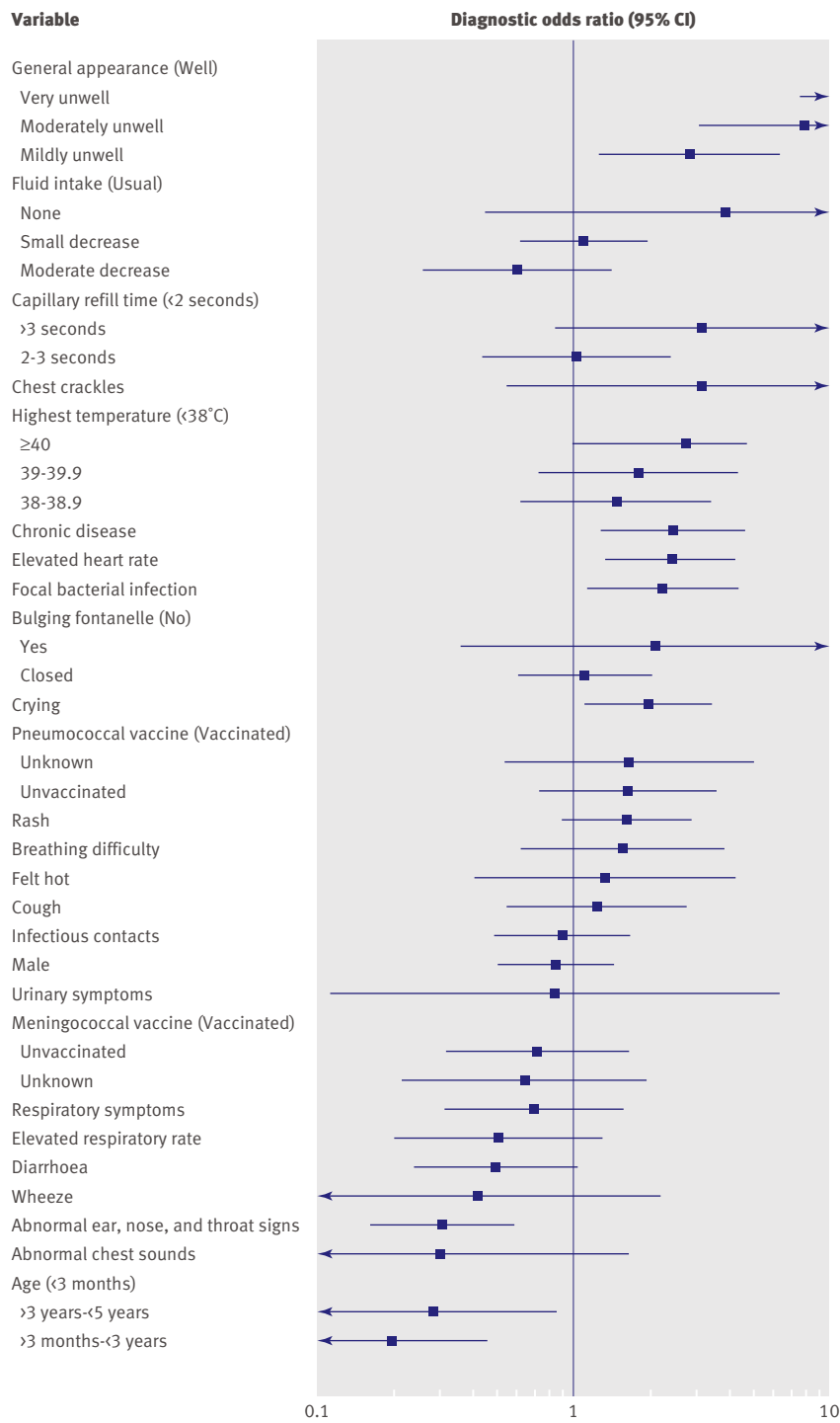


Fig 5 | Clinical indicators of bacteraemia, displayed as logs of the odds ratios from the multinomial model

2 years of age, and 55.9% (8814/15781) occurred in girls. The most common clinical diagnoses were viral infection (3191/15 781 (20.2%)), viral upper respiratory tract infection (2611/15 781 (16.5%)), and gastroenteritis (2186/15 781 (13.9%)).

Frequency of serious bacterial infection

The combined prevalence of any of the three infections of interest (urinary tract infection, pneumonia or

bacteraemia) was 7.2% (1140/15 781, 95% confidence interval (CI) 6.7% to 7.5%), with urinary tract infection in 543 (3.4%) of cases of febrile illness (95% CI 3.2% to 3.7%), pneumonia in 533 (3.4%) of cases (95% CI 3.1% to 3.7%), and bacteraemia in 64 (0.4%) instances (95% CI 0.3% to 0.5%). There were 1066 illnesses caused by a single serious bacterial infection of interest and 37 illnesses caused by two of these infections. Thus 1140 serious bacterial infections of interest occurred in 1103 separate illnesses in 1054 children. Osteomyelitis, septic arthritis, and meningitis were rare, responsible for 12, eight, and six illnesses, respectively; 10 of these infections occurred in illnesses involving more than one infection.

Testing and antibiotic administration

Of the 1140 cases of serious bacterial infection, almost all had the relevant reference standard test performed during their emergency department assessment (fig 2). A total of 511 (94%) of the 543 children with a urinary tract infection had urine collected for culture, whereas 500 (94%) of the 533 children with pneumonia had a chest radiograph, and 95% (61/64) of children with bacteraemia had a blood culture ordered on their first presentation. In comparison, about 383 (35%) of the 1084 children with a clinically diagnosed infection and 5633 (42%) of the 13 557 children without a serious bacterial infection or a clinically diagnosed infection had at least one reference test performed.

A total of 67% (725/1076) of children with a urinary tract infection or pneumonia were prescribed antibiotics during their emergency department presentation, as were 81% (52/64) of children with bacteraemia and 75% (813/1084) of children with clinically diagnosed infections. In children without an identified bacterial infection, 20% (2686/13 557) were prescribed antibiotics.

Follow-up of the 363 children with a serious bacterial infection in whom antibiotics were not prescribed showed that 230 (63%) were subsequently treated with antibiotics and 105 (29%) were not treated; however, for 28 episodes (7.7%), the child's parents were uncontactable. A total of 30% (108/363) of these children returned to the emergency department, usually within 48 hours of the initial visit. Of the 363 children with bacterial infection not treated, only eight were unwell at follow-up an average of 10.2 days later and none was febrile. One death occurred in a child with a lethal congenital disorder on end of life management.

There were no serious health outcomes related to the decision not to start antibiotics in the emergency department. Possible explanations for the absence of harm from underuse of antibiotics include misclassification of disease, because no definition can be expected to cover every possible variant, or clearing the bacteria by the natural immune response.

Selection of items for the diagnostic model

More than 40 clinical signs and symptoms that are routinely elicited in children with fever suspected of

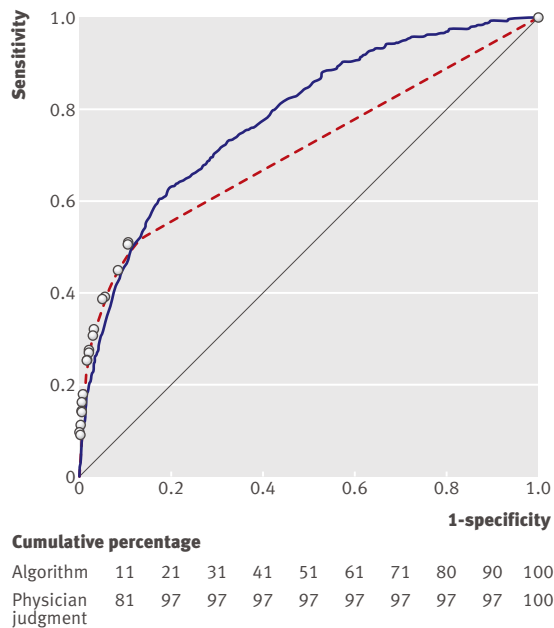


Fig 6 | Receiver operating characteristic curve of the model (solid line) compared with clinician estimation for urinary tract infection (circles (data points) and broken line). Cumulative percentage displays the distribution of the false positive rates

having a serious bacterial infection were considered for our model.

The 26 items selected in the final multinomial logistic regression model have considerable clinical validity because many are intuitively important for diagnosis. Variables that were not selected in the final model are listed in the web extra file. Figures 3, 4, and 5 show the model parameters, odds ratios, and their corresponding 95% confidence intervals for each of the three infections. Some items were positive indicators of disease presence and have an odds ratio greater than one, whereas other variables make the diagnosis of bacterial infection less likely and have an odds ratio of less than one. For example, for urinary tract infection, urinary symptoms and general appearance of the child are the two variables that make urinary tract infection most likely. The presence of a symptom not associated with urinary tract infection, such as stridor, makes diagnosis of a urinary tract infection very unlikely.

Duration of illness, the only continuous variable in the multinomial models, is not reported in the figures. The odds ratio for serious bacterial infection with each 48 hour increase in the duration of illness was 1.05 (95% CI 1.01 to 1.09) for urinary tract infection, 1.09 (95% CI 1.06 to 1.12) for pneumonia, and 1.09 (95% CI 1.02 to 1.16) for bacteraemia.

For all serious bacterial infections, appearing generally unwell was the strongest diagnostic marker, with raised temperature, no fluid intake in the previous 24 hours, increased capillary refill time, and chronic disease also predictive. The presence of localising symptoms and signs was important for pneumonia and urinary tract infection but not bacteraemia. In fact, localising symptoms, except chest crackles and urinary

symptoms, made the diagnosis of bacterial infection less likely. Children aged more than 3 months were less likely to have an infection.

Model performance and validation

The performance of the diagnostic models for each infection was acceptable, with all area under the curve calculations between 0.8 and 0.9. The area under the curve for urinary tract infection was 0.80 (95% CI 0.78 to 0.82), for pneumonia it was 0.84 (95% CI 0.83 to 0.86), and for bacteraemia 0.88 (95% CI 0.84 to 0.92). The performance of the model for urinary tract infection in the whole sample was lower than for the subset for whom dipstick results were available—the area under the curve in this subgroup was 0.89 (95% CI 0.85 to 0.93).

We then validated the model in 5584 illnesses. The validation study included 224 urinary tract infections (prevalence 4.0%), and the area under the curve for the urinary tract infection model was 0.78 (95% CI 0.74 to 0.81), similar to that found in the main study (area under the curve 0.80). There were 193 instances of pneumonia (prevalence 3.5%), and the area under the curve for the relevant model was 0.84 (95% CI 0.82 to 0.87), which is the same area under the curve as is found in the main study. There were 33 cases of bacteraemia (prevalence 0.6%), and the area under the curve was 0.74 (95% CI 0.66 to 0.82), which is lower than that found in the main study (area under the curve 0.88). In the subset of 423 instances where dipstick results were available, the area under the curve was 0.83 (95% CI 0.77 to 0.90), compared with 0.89 in the main study.

Early physician estimation of the likelihood of serious infection produced a different ROC curve from the probability curves developed by the model, mainly because the majority (85% to 95%) of physicians estimated that roughly 95% of all patients would not have a serious bacterial infection (figures 6, 7, and 8). From the figures, it is evident that physician estimates of the probability of infection correspond predominantly to the low sensitivity (0.1 to 0.5) and high specificity (0.9 to 1.0 (0 to 0.1 of the 1-specificity axis)) region of the ROC curve. Across these ranges of sensitivity and specificity, physician judgment and the diagnosis from the model were very similar. Beyond these values, clinician judgment after initial assessment was less accurate than the model for the majority of children that had a serious bacterial infection, because physicians did not estimate a high likelihood of disease. We did not calculate area under the curve values for the physicians' estimates because the distribution of data for clinical judgment was restricted to the lower quadrant of the ROC curve.

DISCUSSION

This large prospective cohort study of young children presenting with a febrile illness to an emergency department in Australia has shown that urinary tract infection, pneumonia, and bacteraemia occur in about 7% of illnesses in children who present with fever, but in only

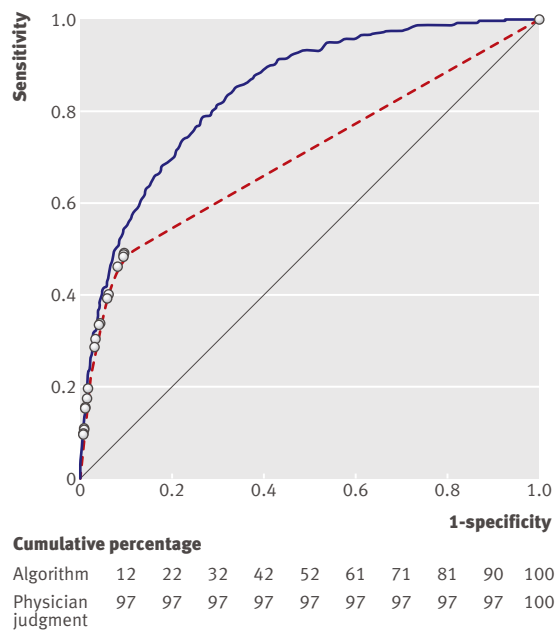


Fig 7 | Receiver operating characteristic curve of the model (solid line) compared with the clinician estimation for pneumonia (circles and broken line). Cumulative percentage displays the distribution of the false positive rates

66-81% of cases are the children administered antibiotics at the time of their first presentation.

Our study has shown that the low rate of antibiotic administration is not caused by failure to consider the diagnosis of serious bacterial infection. Urinary tract infection, pneumonia, and bacteraemia were almost always considered as possible causes of febrile illness, as shown by the ordering of the appropriate tests in approximately 95% of children with serious bacterial infection. Nor is the lack of treatment the result of the symptoms and signs elicited by physicians in this setting having poor diagnostic value. We found that many clinical features widely considered to be indicative of serious bacterial infection—such as overall appearance of the child, urinary symptoms for urinary tract infection, and cough for pneumonia—were highly discriminatory for both ruling in and ruling out infection.

We identified two major, potentially correctable, difficulties in the current diagnostic decision making process. Firstly, in combining the demographic items and clinical symptoms and signs related to febrile illness, the physicians tended to underestimate the likelihood of serious bacterial infection. There are too many relevant signs and symptoms for doctors to assimilate effectively; instead, they tended to discount the information and underestimate the probability of serious disease. As such, the full diagnostic value of current clinical tests was often not reached. Secondly, where near patient tests were available, such as urinalysis for urinary tract infection and chest radiograph for pneumonia, errors in interpretation meant that serious bacterial infection was left untreated at the initial presentation.

It is likely that in some cases physicians correctly identified pneumonia on the chest radiograph or urinary tract infection using the urinalysis results but decided not to treat the child with antibiotics. Consolidation on chest radiograph in association with viral symptoms, for example, may have been considered indicative of viral pneumonia by the treating physician. In suspected urinary tract infection, the importance of collecting an uncontaminated specimen of urine means that some practitioners routinely delay the prescription of antibiotics until the culture result is known. Within an emergency department setting, this approach is not unreasonable and all children are referred back to their local healthcare provider. However, we found that about two thirds of children who were not treated at their first presentation were subsequently prescribed antibiotics; therefore, early antibiotic administration would be expected to shorten the duration of the febrile illness and prevent unnecessary re-presentations.

One third of children with serious bacterial infection appeared to recover spontaneously without antibiotics. This finding could be explained by some degree of misclassification in our definitions of urinary tract infection and pneumonia, the two infections responsible for most cases in this study. Some level of misclassification is inevitable because no definition will encompass every possible variant of either of these infections; however, our definitions were based on published information and clinical consensus, and are clearly reported. There is evidence that between 28% and 40% of adult women with urinary tract infection are cured within 7 days without having received treatment,^{26,27} which suggests that our rate of spontaneous cure is not unprecedented.

Our data suggest that two interventions could be used to improve clinical decision making with respect to children who present with fever. Firstly, a computer assisted diagnostic decision tool such as that developed in this study could be used to determine the likelihood of serious bacterial infection. Doctors in the emergency department would enter the clinical findings into a computer program and the risk calculation would be generated for them. On the basis of the level of risk, treatment could commence or be withheld until further information is available. The model developed in our study could be calibrated to a low threshold of diagnosis (high sensitivity), for example, which should result in more children with serious bacterial infection receiving antibiotics and at an earlier stage of their illness. The trade-off would be a lower specificity, however, resulting in additional testing and treatment for children who do not have serious bacterial infection. Secondly, for urinary tract infection, where urinalysis can assist diagnostic decision making, and for pneumonia, where chest radiograph is available as the reference standard, better understanding of the results and the use of decision rules with lower decision thresholds than are currently used could improve rates of antibiotic prescription on initial presentation to emergency departments.

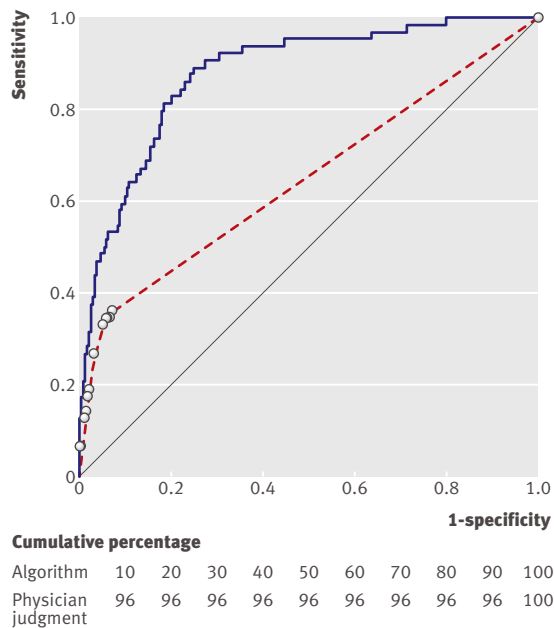


Fig 8 | Receiver operating characteristic curve of the model (solid line) compared with the clinician estimation for bacteraemia (circles and broken line). Cumulative percentage displays the distribution of the false positive rates

Strengths and limitations of study

Our study has a number of important features. The eligibility criteria were based on a common reason for presentation—a febrile illness rather than a specific disease—which enabled us to evaluate the accuracy of clinical symptoms and signs for the main types of relevant bacterial infections. The size of the study provided sufficient power to allow us to develop a robust and reasonably precise model to discriminate between the most common serious bacterial infections (urinary tract infection, pneumonia, and bacteraemia) rather than having to consider them as a single group of “serious bacterial infections,” which is critical given that the treatment for each is different. Validation of the model in a different group of patients showed that the performance of the model was very stable for both pneumonia and urinary tract infection, the most common serious bacterial illnesses

We collected data on the decision making steps after the initial assessment of a febrile child—including physician estimates of the probability of serious bacterial infection on the basis of initial assessment, test ordering, and antibiotic administration—and not simply on the final diagnosis for each child. We also used a final diagnosis committee to verify probable cases of bacterial infection, members of which were blinded to other clinical data like the reference test results, which minimised possible random and systematic error in the classification of children. Finally, we were able to follow up and verify the diagnosis in 93% of children.

We believe the findings of this study are generalisable to other acute paediatric settings because the prevalence of serious bacterial infection in this research is

very similar to that in other studies^{17,28,29} and the frequency of antibiotic administration is also similar.^{30,31}

Our study does have some potential limitations. We did not have microbiological and radiological verification in all children, so we may not have detected some bacterial infections that spontaneously resolved. This would have required urine, blood, and chest radiography in all children up to 5 years of age, which was not justifiable clinically or ethically. Instead, in keeping with standards for the evaluation of diagnostic tests where applying the reference standard test in all patients is not feasible, we used a double reference standard.^{32,33} Children were classified as “negative” for serious bacterial infection if all reference standard tests that were done were negative (which was the case in about 25% of eligible children) and if, on follow-up, a parent reported resolution of the child’s illness by days 10–14. Any cases of serious bacterial infection missed because no tests were done and that resolved spontaneously within the follow-up time frame would not greatly benefit from earlier antibiotic administration. In addition, the validity of the physician estimates of disease may be uncertain, particularly at this early stage of assessment, and the estimates had no influence on subsequent clinical care decisions.

Comparison with other studies

The evaluation of young febrile children is a major management dilemma worldwide and has attracted considerable research and policy attention. This is reflected in the recently published National Institute for Health and Clinical Excellence (NICE) guidelines for the assessment and initial management of febrile illness in children under 5 years.³⁴

These guidelines summarise the existing clinical assessment scales that combine multiple signs and symptoms in this setting. The guideline authors conducted a literature review and identified two scoring systems: the Yale observational scale¹² and the young infant observational scale.¹¹ Problems in the development of each scale were detected, however, and neither scale was considered sensitive for the detection of serious bacterial infection. Furthermore, the review of studies of individual signs and symptoms noted that most of the data were in infants younger than 6 months of age but authors concluded it was reasonable to extrapolate to older children.

The NICE group subsequently developed a “traffic light” assessment tool that can be used to classify children into low risk, intermediate risk, and high risk of serious infection; however, the tool has not yet been evaluated. The elements in the assessment tool do have considerable overlap with fields collected in the current study, with omission only of some comparatively rare and specific events: status epilepticus, swelling of limb (non-weight bearing), bile stained vomit, and “new lump.”

Policy implications

Diagnostic decision making has received little attention compared with therapeutics, despite the

WHAT IS ALREADY KNOWN ON THIS TOPIC

In children with fever, it can be difficult to distinguish serious bacterial infections from viral infections

Current diagnostic processes and clinical scoring systems are inadequate in an acute setting and lead to a substantial proportion of children with serious bacterial infection not being treated with antibiotics when they first present

There is a need for an accurate acute clinical decision making tool that takes into account all the signs and symptoms associated with serious causes of febrile illness

WHAT THIS STUDY ADDS

Urinary tract infection, pneumonia, and bacteraemia occur in about 7% of children who present to an emergency department with a febrile illness, but only 70-80% of such children are prescribed antibiotics on initial consultation

Our computerised diagnostic model successfully combined clinical history, signs, and symptoms to provide a high to moderately accurate estimate of the risk of serious bacterial infection in children presenting with febrile illness

Our clinical diagnostic model outperformed clinical judgment for the diagnosis of fever in young children

universally accepted “medical mantra” of the importance of history taking and physical examination for all patients. Given the complexity of this process and the sheer number of clinical symptoms and signs elicited (including multiple thresholds at multiple and varying times during the illness), it is highly likely that errors in judgment occur when combining these clinical features. In many situations, inefficiencies, costs, and harms may occur by the needless ordering of additional tests and by overtreatment or undertreatment. By combining routinely collected clinical information into a statistical model, we have demonstrated that a clinical diagnostic model may improve the care of children presenting with fever who have suspected serious bacterial illness.

Conclusions

Urinary tract infection, pneumonia, and bacteraemia occur in about 7% of children who present to an emergency department with a febrile illness, but only 70-80% of such children are prescribed antibiotics on initial consultation and 20% of children without an identified bacterial infection are probably overtreated with antibiotics. Combining physician elicited symptoms and signs into a statistical model presented as a computer assisted diagnostic decision system provides scope to improve sensitivity compared with physician judgment, thereby improving early treatment.

We would like to thank all the emergency department physicians who diligently completed the structured medical assessments on nearly 20 000 children; without their contribution this project would not have been possible.

Contributors: JCC undertook the design of the study, funding applications, presentation of the results, and the writing of manuscript. GJW obtained ethics permission and took part in data collection, database design, monitoring and reporting, compilation of results, medical staff training, and reviewing the manuscript. MJ analysed and compiled the results, and reviewed manuscript. MC compiled and piloted the febrile assessment tool, undertook medical staff training, data collection, and interpretation, compiled the results, and reviewed the manuscript. PM contributed to the statistical analysis design, interpretation of the analysis, and manuscript review. AH undertook statistical analysis, presentation of results, and

review of the manuscript. LI contributed to the study design and review of the manuscript. DAF and DI formulated the disease definitions, acted as members of the final diagnosis review committee, and undertook clinical interpretation of results and review of the manuscript. DAF also monitored the study during JCC's absence. MMC facilitated the study in the emergency department, formulated the disease definitions, reviewed the febrile template, undertook training and support of emergency staff, and contributed to clinical interpretation of the results and review of the manuscript. All authors had full access to all data and analyses. JCC, GJW, and PM act as the guarantors.

Funding: This trial was funded by the National Health and Medical Research Council of Australia (programme grant numbers 211205 and 402764). The funding source had no influence on study design, data collection, analysis, interpretation of data, writing of the report, or on the decision to submit the paper for publication.

Conflicts of interest: All authors have completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: (1) No financial support for the submitted work from anyone other than their employer; (2) No financial relationships with commercial entities that might have an interest in the submitted work; (3) No spouses, partners, or children with relationships with commercial entities that might have an interest in the submitted work; (4) No non-financial interests that may be relevant to the submitted work.

Ethical approval: The study had ethics approval through the University of Sydney Human Research Ethics committee and the Royal Alexandra Hospital for Children Ethics Committee.

Data sharing: No additional data available.

- Hay AD. The prevalence of symptoms and consultations in pre-school children in the Avon Longitudinal Study of Parents and Children (ALSPAC): a prospective cohort study. *Fam Pract* 2005;22:367-74.
- Schmitt BD. Fever phobia. *Am J Dis Child* 1980;134:176-81.
- Heffernan R, Mostashari F, Das D, Karpati A, Kulldorf M, Weiss D. Syndromic surveillance in public health practice, New York City. *Emerg Infect Dis* 2004;10:858-64.
- Alpern ER, Stanley RM, Gorelick MH, Donaldson A, Knight S, Teach SJ, et al. Epidemiology of pediatric emergency medicine research network. *Pediatr Emerg Care* 2006;22:689-99.
- Massin MM, Montesanti J, Gerard P, Lepage P. Spectrum and frequency of illness presenting to a pediatric emergency department. *Acta Clin Bel* 2006;61:161-5.
- US Census Bureau. Population finder: United States. 2010. http://factfinder.census.gov/servlet/SAFFPopulation?_submenuld=population_08_sse=on.
- Hsiao AL, Chen L, Baker D. Incidence and predictors of serious bacterial infections among 57 to 180 day old infants *Pediatrics* 2006;117:1695-9.
- Trainor JL, Hampers LC, Krug SE, Listerick R. Children with first-time simple febrile seizures are at low risk of serious bacterial illness. *Acad Emerg Med* 2001;8:781-7.
- McGowan JE, Bratton L, Klein JO, Finaland M. Bacteraemia in febrile children seen in a “walk-in” pediatric clinic. *N Engl J Med* 1973;288:1309-12.
- Simpkins D, Woods N, Jelfs J, McIntyre PB, Menzies R, Lawrence G, Booy R. Modern trends in mortality from meningococcal disease in Australia. *Pediatr Infect Dis J* 2009;28:1119-20.
- Bonadio WA, Hennes H, Smith D, Ruffing R, Melzer-Lange M, Lye P, et al. Reliability of observation variables in distinguishing infectious outcome of febrile young infants. *Pediatr Infect Dis J* 1993;12:111-4.
- McCarthy PL, Sharpe MR, Spiesel SZ, Dolan TF, Forsyth BW, DeWitt TG, et al. Observation scales to identify serious illness in febrile children. *Pediatrics* 1982;70:802-9.
- Hewson PH, Poulakis Z, Jarman F, Kerr J, McMaster D, Goode J, et al. Clinical markers of serious illness in young infants: a multicentre follow-up study. *J Paediatr Child Health* 2000;36:221-5.
- Van den Bruel A, Aertgeerts B, Bruyninckx R, Aerts M, Buntinx F. Signs and symptoms for diagnosis of serious infections in children: a prospective study in primary care. *Br J Gen Pract* 2007;57:538-46.
- Sackett DL, Rennie D. The science of the art of the clinical examination. *JAMA* 1992;267:2650-2.
- Newman-Toker DE, Pronovost PJ. Diagnostic errors—the next frontier for patient safety. *JAMA* 2009;301:1060-2.
- Dagan R, Powell KR, Hall CB, Menugus MA. Identification of infants unlikely to have serious bacterial infection although hospitalized for suspected sepsis. *J Pediatr* 1985;107:855-60.
- Jaskiewicz JA, McCarthy CA, Richardson AC, White KC, Fisher DJ, Dagan R, et al. Febrile infants at low risk for serious bacterial infection—an appraisal of the Rochester criteria and implications for management. *Pediatrics* 1994;94:390-6.

- 19 Teach SJ, Fleisher GR, for the Occult Bacteraemia Study Group. Efficacy of an observation scale in detecting bacteraemia in febrile children three to thirty-six months of age, treated as outpatients. *J Pediatr* 1995;126:877-81.
- 20 Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteraemia in young febrile infants. *Ann Emerg Med* 1998;31:679-87.
- 21 Whiting P, Westwood M, Watt I, Cooper J, Kleijnen J. Rapid tests and urine sampling techniques for the diagnosis of urinary tract infection (UTI) in children under five years: a systematic review. *BMC Pediatrics* 2005;5:4.
- 22 Steyerberg E. Clinical prediction models, a practical approach to development, validation and updating. Springer, 2009.
- 23 Sauerbrei W, Schumacher M. A bootstrap resampling procedure for model building: application to the Cox regression model. *Stat Med* 1992;11:2093-109.
- 24 Sauerbrei W. The use of resampling methods to simplify regression models in medical statistics. *Appl Statist* 1999;48:313-29.
- 25 Royston P, Sauerbrei W. Stability of multivariable fractional polynomial models with selection of variables and transformations: a bootstrap investigation. *Stat Med* 2003;22:639-59.
- 26 Christiaens TCM, De Meyere MD, verschraegen G, Peersman W, Heytens S, De Maeseneer JM. Randomised controlled trial of nitrofurantoin versus placebo in the treatment of uncomplicated urinary tract infection in adult women. *Br J Gen Pract* 2002;52:729-34.
- 27 Ferry SA, Holm SE, Stenlund H, Lundholm R, Monsen TJ. The natural course of uncomplicated lower urinary tract infection in women illustrated by a randomized placebo controlled study. *Scand J Infect Dis* 2004;36: 296-301.
- 28 Pantell RH, Newman TB, Bernzweig J, Bergman DA, Takayama JL, Segal M, et al. Management and outcomes of care of fever in early infancy. *JAMA* 2004;291:1203-12.
- 29 Bachur RG, Harper MB. Predictive model for serious bacterial infections among infants younger than 3 months of age. *Pediatrics* 2001;108:311-6.
- 30 Perz JA, Craig AS, Coffey CS, Jorgensen DM, Mitchel E, Hall S, et al. Changes in antibiotic prescribing for children after a community-wide campaign. *JAMA* 2002;287:3103-9.
- 31 Nyquist A, Gonzales R, Steiner JF, Sande MA. Antibiotic prescribing for children with colds, upper respiratory tract infections, and bronchitis. *JAMA* 1998;279:875-7.
- 32 Kondratovich MV. Comparing two medical tests when results of reference standard are unavailable for those negative via both tests. *J Biopharm Stat* 2008;18:145-66.
- 33 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. *Clin Chem* 2003;49:1-6.
- 34 National Institute for Health and Clinical Excellence. *Feverish illness: assessment and initial management in children younger than 5 years*. NICE, 2007.

Accepted: 11 February 2010