

Systematic review and validation of prediction rules for identifying children with serious infections in emergency departments and urgent-access primary care

M Thompson, A Van den Bruel, J Verbakel, M Lakhanpaul, T Haj-Hassan, R Stevens, H Moll, F Buntinx, M Berger, B Aertgeerts, R Oostenbrink and D Mant



March 2012
10.3310/hta16150

Health Technology Assessment
NIHR HTA programme
www.hta.ac.uk





How to obtain copies of this and other HTA programme reports

An electronic version of this title, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (www.hta.ac.uk). A fully searchable DVD is also available (see below).

Printed copies of HTA journal series issues cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our despatch agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per issue and for the rest of the world £3 per issue.

How to order:

- fax (with **credit card details**)
- post (with **credit card details** or **cheque**)
- phone during office hours (**credit card** only).

Additionally the HTA website allows you to either print out your order or download a blank order form.

Contact details are as follows:

Synergie UK (HTA Department)
Digital House, The Loddon Centre
Wade Road
Basingstoke
Hants RG24 8QW

Email: orders@hta.ac.uk

Tel: 0845 812 4000 – ask for ‘HTA Payment Services’
(out-of-hours answer-phone service)

Fax: 0845 812 4001 – put ‘HTA Order’ on the fax header

Payment methods

Paying by cheque

If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *University of Southampton* and drawn on a bank with a UK address.

Paying by credit card

You can order using your credit card by phone, fax or post.

Subscriptions

NHS libraries can subscribe free of charge. Public libraries can subscribe at a reduced cost of £100 for each volume (normally comprising 40–50 titles). The commercial subscription rate is £400 per volume (addresses within the UK) and £600 per volume (addresses outside the UK). Please see our website for details. Subscriptions can be purchased only for the current or forthcoming volume.

How do I get a copy of HTA on DVD?

Please use the form on the HTA website (www.hta.ac.uk/htacd/index.shtml). *HTA on DVD* is currently free of charge worldwide.

The website also provides information about the HTA programme and lists the membership of the various committees.

Systematic review and validation of prediction rules for identifying children with serious infections in emergency departments and urgent-access primary care

M Thompson,^{1*} A Van den Bruel,¹ J Verbakel,²
M Lakhanpaul,³ T Haj-Hassan,¹ R Stevens,¹ H Moll,⁴
F Buntinx,² M Berger,⁵ B Aertgeerts,² R Oostenbrink⁴
and D Mant¹

¹Department of Primary Care Health Sciences, Oxford University, Oxford, UK

²Department of General Practice, Katholieke Universiteit Leuven, Leuven, Belgium

³Department of Paediatrics, Division of Medical Education and Social Care, University of Leicester, Leicester, UK

⁴Department of Pediatrics, Erasmus MC-Sophia, Rotterdam, the Netherlands

⁵Department of General Practice, Groningen, the Netherlands

*Corresponding author

Declared competing interests of authors: none

Published March 2012

DOI: 10.3310/hta16150

This report should be referenced as follows:

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).

Health Technology Assessment is indexed and abstracted in *Index Medicus/MEDLINE*, *Excerpta Medica/EMBASE*, *Science Citation Index Expanded (SciSearch®)* and *Current Contents®/Clinical Medicine*.

The Health Technology Assessment (HTA) programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The research findings from the HTA programme directly influence decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC). HTA findings also help to improve the quality of clinical practice in the NHS indirectly in that they form a key component of the 'National Knowledge Service'.

The HTA programme is needs led in that it fills gaps in the evidence needed by the NHS. There are three routes to the start of projects.

First is the commissioned route. Suggestions for research are actively sought from people working in the NHS, from the public and consumer groups and from professional bodies such as royal colleges and NHS trusts. These suggestions are carefully prioritised by panels of independent experts (including NHS service users). The HTA programme then commissions the research by competitive tender.

Second, the HTA programme provides grants for clinical trials for researchers who identify research questions. These are assessed for importance to patients and the NHS, and scientific rigour.

Third, through its Technology Assessment Report (TAR) call-off contract, the HTA programme commissions bespoke reports, principally for NICE, but also for other policy-makers. TARs bring together evidence on the value of specific technologies.

Some HTA research projects, including TARs, may take only months, others need several years. They can cost from as little as £40,000 to over £1 million, and may involve synthesising existing evidence, undertaking a trial, or other research collecting new data to answer a research problem.

The final reports from HTA projects are peer reviewed by a number of independent expert referees before publication in the widely read journal series *Health Technology Assessment*.

Criteria for inclusion in the HTA journal series

Reports are published in the HTA journal series if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this issue of the journal was commissioned by the HTA programme as project number 07/37/05. The contractual start date was in March 2009. The draft report began editorial review in December 2010 and was accepted for publication in May 2011. As the funder, by devising a commissioning brief, the HTA programme specified the research question and study design. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA programme or the Department of Health.

Editor-in-Chief: Professor Tom Walley CBE
 Series Editors: Dr Martin Ashton-Key, Professor Aileen Clarke, Dr Tom Marshall, Professor John Powell, Dr Rob Riemsmas and Professor Ken Stein
 Associate Editor: Dr Peter Davidson
 Editorial Contact: edit@southampton.ac.uk

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

ISSN 2046-4932 (DVD)

© Queen's Printer and Controller of HMSO 2012. This work was produced by Thompson *et al.* under the terms of a commissioning contract issued by the Secretary of State for Health.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (<http://www.publicationethics.org/>).

This journal may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NETSCC, Health Technology Assessment, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Published by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk), on behalf of NETSCC, HTA.

Printed on acid-free paper in the UK by the Charlesworth Group.

Abstract

Systematic review and validation of prediction rules for identifying children with serious infections in emergency departments and urgent-access primary care

M Thompson,^{1*} A Van den Bruel,¹ J Verbakel,² M Lakhanpaul,³
T Haj-Hassan,¹ R Stevens,¹ H Moll,⁴ F Buntinx,² M Berger,⁵ B Aertgeerts,²
R Oostenbrink⁴ and D Mant¹

¹Department of Primary Care Health Sciences, Oxford University, Oxford, UK

²Department of General Practice, Katholieke Universiteit Leuven, Leuven, Belgium

³Department of Paediatrics, Division of Medical Education and Social Care, University of Leicester, Leicester, UK

⁴Department of Pediatrics, Erasmus MC-Sophia, Rotterdam, the Netherlands

⁵Department of General Practice, Groningen, the Netherlands

*Corresponding author

Background: Although the vast majority of children with acute infections are managed at home, this is one of the most common problems encountered in children attending emergency departments (EDs) and primary care. Distinguishing children with serious infection from those with minor or self-limiting infection is difficult. This can result in misdiagnosis of children with serious infections, which results in a poorer health outcome, or a tendency to refer or admit children as a precaution; thus, inappropriately utilising secondary-care resources.

Objectives: We systematically identified clinical features and laboratory tests which identify serious infection in children attending the ED and primary care. We also identified clinical prediction rules and validated those using existing data sets.

Data sources: We searched MEDLINE, Medion, EMBASE, Cumulative Index to Nursing and Allied Health Literature and Database of Abstracts of Reviews of Effects in October 2008, with an update in June 2009, using search terms that included terms related to five components: serious infections, children, clinical history and examination, laboratory tests and ambulatory care settings. We also searched references of included studies, clinical content experts, and relevant National Institute for Health and Clinical Excellence guidelines to identify relevant studies. There were no language restrictions. Studies were eligible for inclusion if they were based in ambulatory settings in economically developed countries.

Review methods: Literature searching, selection and data extraction were carried out by two reviewers. We assessed quality using the quality assessment of diagnostic accuracy studies (QUADAS) instrument, and used spectrum bias and validity of the reference standard as exclusion criteria. We calculated the positive likelihood ratio (LR+) and negative likelihood ratio (LR-) of each feature along with the pre- and post-test probabilities of the outcome. Meta-analysis was performed using the bivariate method when appropriate. We externally validated clinical prediction rules identified from the systematic review using existing data from children attending ED or primary care.

Results: We identified 1939 articles, of which 35 were selected for inclusion in the review. There was only a single study from primary care; all others were performed in the ED. The quality of the included studies was modest. We also identified seven data sets (11,045 children) to use for external validation. The most useful clinical features for ruling in serious infection was parental or clinician overall concern that the illness was different from previous illnesses or that something was wrong. In low- or intermediate-prevalence settings, the presence of fever had some diagnostic value. Additional red flag features included cyanosis, poor peripheral circulation, rapid breathing, crackles on auscultation, diminished breath sounds, meningeal irritation, petechial rash, decreased consciousness and seizures. Procalcitonin (LR+ 1.75–2.96, LR– 0.08–0.35) and C-reactive protein (LR+ 2.53–3.79, LR– 0.25–0.61) were superior to white cell counts. The best performing clinical prediction rule was a five-stage decision tree rule, consisting of the physician’s gut feeling, dyspnoea, temperature $\geq 40^{\circ}\text{C}$, diarrhoea and age. It was able to decrease the likelihood of serious infections substantially, but on validation it provided good ruling out value only in low-to-intermediate-prevalence settings (LR– 0.11–0.28). We also identified and validated the Yale Observation Scale and prediction rules for pneumonia, meningitis and gastroenteritis.

Limitations: Only a single study was identified from primary-care settings, therefore results may lack generalisability.

Conclusions: Several clinical features are useful to increase or decrease the probability that a child has a serious infection. None is sufficient on its own to substantially raise or lower the risk of serious infection. Some are highly specific (‘red flags’), so when present should prompt a more thorough or repeated assessment. C-reactive protein and procalcitonin demonstrate similar diagnostic characteristics and are both superior to white cell counts. However, even in children with a serious infection, red flags will occur infrequently, and their absence does not lower the risk. The diagnostic gap is currently filled by using clinical ‘gut feeling’ and diagnostic safety-netting, which are still not well defined. Although two prediction rules for serious infection and one for meningitis provided some diagnostic value, we do not recommend widespread implementation at this time. Future research is needed to identify predictors of serious infection in children in primary-care settings, to validate prediction rules more widely, and determine the added value of blood tests in primary-care settings.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

List of abbreviations	vii
Executive summary	ix
1. Background	1
Importance	1
Diagnostic difficulties in children with acute infections	1
Clinical tests	1
Additional testing	2
Assessing the level of urgency	2
National Institute for Health and Clinical Excellence guidance	3
2. Research objectives	5
3. Methodology of the systematic review	7
Literature search	7
Selection process	7
Quality assessment	8
Data extraction and analysis	8
4. Number and characteristics of studies included in the systematic review	11
Numbers of included studies	11
Setting of included studies	11
Age and prevalence of serious infection	15
Outcomes reported	16
Quality of included studies	16
5. Results of systematic review of clinical features	17
Global assessment	17
Child behaviour	17
Circulatory and respiratory clinical features	17
Neurological signs or petechial rash	17
Fever	21
Clinical prediction rules	22
Features of limited help in ruling in or ruling out serious infections	22
6. Results of the systematic review of laboratory tests for serious infections	27
Diagnostic value of laboratory tests for composite outcome of serious infection	27
Diagnostic value of laboratory tests for meningitis and/or bacteraemia	32
7. Methods used for validation of prediction rules	35
Identification of data sets	35
Identification of clinical prediction rules	35
Exploring heterogeneity	36
Characteristics of included data sets	36
Statistical analysis	36

8. Results of external validation of clinical prediction rules	39
Description of included data sets	39
Clinical predictor variables included in data sets	39
Results of external validation	41
9. Discussion	47
Studies identified by systematic review	47
Predictive value of clinical features for serious infection	47
Predictive value of laboratory tests rules for serious infection	48
Predictive value of clinical prediction rules for serious infection	50
Limitations of the systematic review	51
Limitations of validation of existing clinical prediction rules	51
10. Conclusions	55
Overall clinical implications	55
Research implications	56
Acknowledgements	59
References	61
Appendix 1 Search terms used for systematic review	67
Appendix 2 Details of the clinical prediction rules identified in the systematic review	69
Appendix 3 Details of clinical features and cut-off values in prediction rules included in <i>Chapter 6</i>	71
Appendix 4 Characteristics of variables included in the data sets used to validate clinical prediction rules	73
Appendix 5 Data sets identified which allowed multiple external validation of clinical prediction rules	75
Appendix 6 Percentage of missing values of variables included in the clinical prediction rules	77
Appendix 7 Research protocol	79
Health Technology Assessment programme	95

List of abbreviations

CI	confidence interval
CRP	C-reactive protein
CSF	cerebrospinal fluid
ED	emergency department
ESR	erythrocyte sedimentation rate
GP	general practitioner
LR	likelihood ratio
LR+	positive likelihood ratio
LR-	negative likelihood ratio
NICE	National Institute for Health and Clinical Excellence
PCT	procalcitonin
QUADAS	quality assessment of diagnostic accuracy studies
WBC	white blood cell count
YOS	Yale Observation Scale

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

Executive summary

Background

Although the vast majority of children with acute infection are managed at home, this is one of the most common problems encountered in children attending emergency departments (EDs) and primary care (in and out of hours). Distinguishing children with serious infection (such as meningitis or complications from viral illnesses such as hypoxia due to bronchiolitis) from those with minor or self-limiting infection is difficult. Firstly, despite the high volume of acute paediatric illness, serious infections are rare in most settings, ranging from < 1% in primary-care settings to as high as 25% in children attending ED with fever without source. Secondly, children with serious illness may present at an early stage when severity is not apparent and deteriorate rapidly. Finally, assessment of children can be difficult and is often undertaken by staff with limited paediatric training. This can result in either misdiagnosis of children with serious infections, which results in a poorer health outcome, or a tendency to refer or admit children as a precaution, thus inappropriately utilising secondary-care resources.

The aim of this study was to identify clinical features, laboratory tests and clinical prediction rules which can be used to identify children with serious infection in acute paediatric settings, including paediatric ED and primary care. We also attempted to externally validate existing clinical prediction rules.

Methods

We used a systematic review of the literature to June 2009, not limited by language, to identify relevant studies of clinical and laboratory predictors of serious infection in children in ambulatory settings. We assessed quality using the quality assessment of diagnostic accuracy studies (QUADAS) instrument, and used two items as exclusion criteria: spectrum bias and validity of the reference standard. We calculated positive and negative likelihood ratios (LR+ and LR-, respectively) for each feature along with the pre- and post-test probabilities of the outcome. Diagnostic features were categorised either as red flags (LR+ > 5.0) or as rule-out features (LR- < 0.2) for serious illness. Setting was used to categorise studies, as a proxy for prevalence of serious infection. The diagnostic value of temperature was explored using a plot of post-test values against pre-test prevalence. Meta-analysis was performed using the bivariate method when appropriate.

We validated clinical prediction rules identified from the systematic review using existing data sets on populations of children attending ED or primary care. Variables used in each data set were translated and clarified. The accuracy of the clinical prediction rules identified in the systematic review was assessed in each of the data sets in which this was possible, using approximations when necessary.

Results

We identified 1939 articles, of which 35 were selected for inclusion in the review. Studies were performed in the USA (16), the UK (5), the Netherlands (4), Switzerland (3), Canada (2), and one each from Belgium, Italy, Australia, Denmark and Spain. There was only a single study from

primary care; all others were performed in ED. A total of 30 studies reported clinical features; 14 studies reported laboratory tests for the diagnosis of serious infections. Most studies included only children with fever, and most focused on the younger age groups. The quality of the included studies was modest.

Diagnostic value of clinical features

Parental concern that the illness is different from previous illnesses (LR+ 14) and the clinician's gut feeling that something is wrong (LR+ 23) provide the strongest rule-in value, based on a single study from a low-prevalence setting. Change in the child's crying pattern, drowsiness, moaning and inconsolability all had a LR+ > 5.0 from this study. However, these features all provided weaker likelihood ratios (LRs) in intermediate- or high-prevalence settings. Fever (temperature > 38.5 °C) had some rule-out value in three studies and a modest rule-in value in one single study. In the five studies with higher prevalence, temperature provided no rule-in ability. Cyanosis had LR+ ranging from 2.66 to 52.2, and poor peripheral circulation had LR+ ranging from 2.39 to 38.8. Rapid breathing and shortness of breath provided the greatest LR+ in the single low-prevalence study (9.3 and 9.70). Crackles on auscultation and diminished breath sounds again provided a LR+ > 5 in the low-prevalence setting, but little value in a single study in an intermediate prevalence setting study. Meningeal irritation, petechial rash, decreased consciousness and seizures had a LR+ > 5 in most of the studies which assessed these features. Loss of consciousness had a LR+ of 19.8–155.

We identified six clinical prediction rules. The Yale Observation Scale provided a LR- < 0.2 in two studies, whereas in five other studies it varied from 0.68 to 0.94. After meta-analysis, summary sensitivity was 32.5% [95% confidence interval (CI) 21.7% to 45.5%], and specificity was 78.9% (95% CI 73.9% to 83.1%). The rule that performed best for ruling out serious infection (LR- 0.04) involved the physician's gut feeling, dyspnoea, temperature $\geq 40^{\circ}\text{C}$ and diarrhoea in children between 1 and 2.5 years of age, but was assessed in only a single low-prevalence study. The same study reported two prediction rules for pneumonia (LR- 0.07), involving dyspnoea and either the physician's gut feeling or parental concern. Additionally, we identified two prediction rules for meningitis from intermediate settings; one had a very low LR- (LR- 0.05) and consisted of any neurological finding and seeking care within < 48 hours, whereas the other had high LR+ (LR+ 395) and consisted of petechiae, nuchal rigidity or coma. Finally, a single rule was identified for dehydration from gastroenteritis, which provided a modest LR+ (6.1) and LR- (0.24) from a single high-prevalence study. This rule consisted of any two of the following: absent tears, dry mucous membranes, ill appearance and decreased peripheral circulation.

Laboratory tests predictive of serious infections

Three studies which reported the results of procalcitonin (PCT) for composite outcome of serious infection demonstrated a LR+ of 1.75–2.96, with a LR- of 0.08–0.35. The five studies of C-reactive protein (CRP) for composite outcome of serious infection provided a LR+ of 2.53–3.79 and a LR- of 0.25–0.61. Meta-analysis of CRP yielded a pooled LR+ of 3.15 (95% CI 2.67 to 3.71) and a pooled LR- of 0.33 (95% CI 0.22 to 0.49) across all cut-offs. Both CRP and PCT had similarly shaped receiver operator characteristics curves with overlapping CIs. The one study which evaluated CRP for the diagnosis of meningitis and/or bacteraemia showed that CRP was able to exclude meningococcal disease (LR- 0.05). White blood cell count (WBC), absolute neutrophil count, band count or left shift all demonstrated little diagnostic value for composite outcome of serious infection: the minimum LR- was 0.61 with the 95% CI in most studies crossing 1.0, and LR+ was from 0.87 to 3.05. The summary sensitivity of six studies which evaluated WBC for bacteraemia was 62.71% (95% CI 52.60% to 71.81%) summary specificity 69.27% (95% CI 62.71% to 75.13%), summary LR+ 2.04 (95% CI 1.51 to 2.75), and summary LR- 0.54 (95% CI 0.40 to 0.73). Erythrocyte sedimentation rate was evaluated in a single study, in which it showed LR+ 2.49 and LR- 0.34. Combinations of inflammatory markers offered little

additional diagnostic value over the individual tests. A prediction rule consisting of CRP, PCT and urinalysis has good diagnostic performance for the composite outcome of serious infections, with LR+ 4.92 (95% CI 3.26 to 7.43) and LR- 0.07 (95% CI 0.02 to 0.27).

Results of validation of clinical prediction rules

We used seven data sets (11,045 children) to validate the prediction rules. The Yale Observation Scale was moderately useful to rule in serious infection in three studies (LR+ of 3.35–7.49 depending on cut-off and setting), but had no rule-out value. The five-stage decision tree had no rule-in value in any of the data sets, but in four it offered a marginally useful rule-out value (LR- 0.13–0.35). None of the data sets used to validate the pneumonia rule demonstrated clinically useful LR+, but in one the LR- was 0.22, suggesting some rule-out value. Validation of the meningitis rule demonstrated a clinically useful LR+ of 9.96–38.9 in three data sets from low-prevalence settings, but none provided a useful LR-. In contrast, based on one studying high-prevalence setting, it showed a poor LR+ (1.87), but an extremely small LR- (0.084). Being referred by a physician or not did not influence the LRs, with similar results in the referred and non-referred children.

Conclusions

Overall clinical implications

Our findings illustrate the diagnostic gap between the predictive value achievable by consideration of clinical features and the threshold of risk of serious infection. This gap is currently filled by using clinical 'gut feeling' and diagnostic safety-netting, which are still not well defined in primary care or ED settings. Clearly, a single abnormal clinical finding is insufficient on its own to substantially lower the risk of serious infection. We identified several clinical features which were highly specific 'red flags'. When present, these should prompt a more thorough assessment. However, even in children with a serious infection, red flags will occur infrequently owing to their low sensitivity; therefore, their absence does not lower the risk of a serious infection.

We identified several clinical prediction rules for identifying children with serious infection, but only one (Yale Observation Scale) had any published validation studies. By using existing data sets to validate these rules, we were able to draw additional conclusions. Firstly, clinical prediction rules offer different diagnostic value, depending particularly on the prevalence of serious infection. Secondly, in primary and ED settings, the five-stage decision tree offered a moderate rule-out value and the Yale Observation Scale had a moderate specificity offering some rule-in value. Thirdly, one rule for meningitis provided a high specificity and rule-in value.

Both CRP and PCT offer similar diagnostic performance and are superior to WBCs. However, neither CRP nor PCT has sufficient diagnostic value to either confirm or exclude a serious infection, and thus their results must be interpreted in the light of clinical findings. Moreover, different cut-off values are needed depending on whether these will be used as rule-in or rule-out, which may vary depending on setting in particular.

Research implications

There is a pressing need for:

1. Studies in primary care or low-prevalence ED settings where most children with acute infections are seen, but where we currently have least evidence to support clinical practice.

This research should include the diagnostic role of vital signs, the role of inflammatory markers, and content and implementation of safety-netting.

2. The value of repeated testing using single or combinations of inflammatory markers.
3. Research that involves collaboration at the national or international level which not only maximises study power and generalisability, but also is more efficient.
4. Improvements to the methodology of studies, such as avoiding restrictive selection criteria which involve age or temperature, considering outcomes that are appropriate to the setting, and ensuring that prediction rules are validated and that their impact on clinical practice can be assessed.

Funding

The National Institute for Health Research Health Technology Assessment programme.

Chapter 1

Background

Importance

Acute illness is one of the most common problems encountered in children attending emergency departments (EDs) as well as by primary-care services in the UK. Between 27% and 47% of patients who present to EDs in the UK do so for medical illness, rather than trauma.¹ In the case of children, the most common medical reasons for attending ED are breathing difficulty (31%), febrile illness (20%), diarrhoea/vomiting (16%), abdominal pain (6%), seizure (5%) and rash (5%).¹ Acute infections are responsible for four consultations per person-year in children aged < 1 year, and between one and two in children aged 1–15 years in the UK.² Children < 5 years of age also constitute a substantial part of the workload of urgent-access primary-care services. Indeed, the patient group which presents most commonly to out-of-hours clinics is children with acute infections.^{3,4} Similarly, acute illness in children is also a major component of the work of NHS Direct, in which 22% of all telephone calls are related to children < 5 years of age.⁵ Perhaps most importantly for the NHS, parents of unwell children are often concerned about their child and his/her risk of serious illness and need access to accurate and appropriate diagnostic clinical services.⁶

Diagnostic difficulties in children with acute infections

One of the key tasks in both hospital EDs and primary-care settings is therefore to distinguish children who may have serious infections (e.g. meningitis, bacteraemia) or complications from infection (e.g. hypoxia from bronchiolitis, dehydration from gastroenteritis) from the vast majority with self-limiting or minor infections who can safely be managed at home. This task is challenging. With increasing ED attendance rates in the UK, hospital admission of children is becoming more common despite a falling incidence of serious infection.^{7,8} At the same time, approximately half of children with meningococcal disease are missed at first consultation with a doctor, which results in a poorer health outcome.⁹ Infections were responsible for 20% of childhood deaths reviewed in the recent Confidential Enquiry into Maternal and Child Health report *Why children die*,¹⁰ with the greatest number in the 1- to 4-year age group. One reason why recognition of serious infection in children is difficult is the low prevalence of serious illness. In a primary-care setting, < 1% of children assessed have a serious infection,^{8,11} while in the ED setting this can be as high as 25% of children referred by their GP for fever without source.¹² In addition, infections are dynamic illnesses, and children with serious illness may present at an early stage when severity is not apparent and deteriorate rapidly. Finally, assessment of children (particularly pre-verbal children) may be undertaken by staff with limited paediatric training or under high pressure because of large patient volumes.^{13,14} Recent reductions in working hours may further reduce opportunities for trainees to gain experience of children with serious infections.

Clinical tests

Clinicians use several different combinations of clinical tools to identify children with serious infections in primary and emergency care. This primarily involves assessing symptoms reported

spontaneously or elicited from parents and patients, overall or global assessment of severity of illness, as well as measurement of vital signs and findings from physical examination.¹⁵ There has been no systematic assessment of which of these clinical features are most useful for identifying children with serious infection in ED or primary-care settings. Clinical prediction rules are a simple pragmatic technology that can potentially be used by clinical staff to assist them in assessment and clinical management. A widely implemented example which has been shown to reduce both resource use and missed diagnoses in EDs is the Ottawa Ankle Rule for ordering a radiograph.¹⁶ There has been no systematic assessment of prediction rules for children with serious infection and, in particular, how well these rules have been validated or implemented in different clinical settings. These are necessary steps in the development of prediction rules, before widespread introduction in clinical care.¹⁷

Additional testing

Some clinicians, particularly those working in EDs, have access to further diagnostic tests. These include urine dipsticks, blood tests for white blood cell counts (WBCs) and inflammatory markers, as well as imaging. Apart from urine dipsticks, most of these tests are carried out in hospital or ED settings, although several are now available as point-of-care tests which give an immediate result and could potentially be used in out-of-hours or other primary-care settings. EDs have the capability to use other tests in assessing children, for example additional biochemical and haematological tests, as well as microbiological cultures of blood, cerebrospinal fluid (CSF) and sputum, and a variety of imaging including plain radiographs, ultrasound, computerised tomography and magnetic resonance imaging. However, even the use of blood tests in identifying children with serious infections is not straightforward. In addition to the difficulties and discomfort of obtaining blood samples, particularly in younger children, the role of the test in the overall diagnostic pathway is important so that clinicians do not rely entirely on a test that does not have perfect discrimination.¹⁸ Previous systematic reviews have assessed the value of selected inflammatory markers in either children only¹⁹ or children and adults^{20–22} for the diagnosis of various outcomes such as distinguishing viral from bacterial pneumonia²³ or parenchymal involvement in children admitted with a urinary tract infection.²⁴ However, none has evaluated and compared all available laboratory tests for children with suspected serious infections in ambulatory care. In addition, most studies which evaluate blood tests neglect the available clinical information, which makes it difficult to determine their role in the clinical pathway, and their incremental value over clinical features.^{18,25} In this monograph we will focus on urine and blood tests, which can potentially be used as initial diagnostic tests in most ED settings, rather than the more extensive list of all possible laboratory and imaging investigations.

Assessing the level of urgency

There are several triage systems currently in use in EDs in the UK. These are primarily designed to assess level of urgency for care, rather than as diagnostic tests for serious infections. The Manchester Triage System assigns the patient to one of five categories based on the maximum time that he/she can wait for full assessment.^{26,27} It provides only modest sensitivity (63%) to detect emergency or very urgent cases, and is a generic instrument designed to deal with emergencies including trauma.^{28,29} Other triage systems used internationally include the Emergency Severity Index, the Paediatric Canadian Triage and Acuity Scale, the Paediatric Risk of Admission Score and the Paediatric Emergency Assessment Tool.^{30–34} A number of more specific 'scoring systems' for children presenting to EDs with medical illness have been developed. None has shown sufficient ability to rule out serious infection in children to be widely adopted in a NHS context.^{35–38}

National Institute for Health and Clinical Excellence guidance

The National Institute for Health and Clinical Excellence (NICE) guideline for the management of children < 5 years of age with a fever was published in 2007.³⁹ It is an important starting point because its recommendations are based on a literature review utilising stakeholders to identify key documents. One of the main outputs from this guideline was a 'traffic light' system of clinical features that are designed to be used to assess the risk of serious infection in children. This system assigned clinical features to green, amber and red categories based on the risk of serious infection, and provides clinical guidance for actions needed according to these categories. The current project aims to contribute to the guidelines for the management of children, in a wider context than the NICE guideline, namely by including all children rather than only those < 5 years of age or with fever, but will also be used to update the NICE feverish child guideline. Moreover, we will identify clinical prediction rules relevant to children with acute infection and validate their performance in different clinical settings.

Chapter 2

Research objectives

The overall aim of this research study was to systematically identify clinical features, laboratory tests and clinical prediction rules which can be used to identify children with possible serious infection in acute paediatric settings, including paediatric ED, out-of-hours primary care and other primary-care settings. The clinical role of these diagnostic features includes not only identifying children with serious infection, but also equally identifying children with self-limiting illness who can be safely discharged home from emergency and primary-care settings although not missing any cases of serious infection.

The specific objectives of the systematic review were:

1. to identify the clinical features and prediction rules which have already been shown to have predictive value for identifying (or excluding) children with severe infection
2. to identify and compare the best performing prediction rules from the literature
3. to explore the added value of laboratory tests and vital signs to prediction rules based solely on clinical history and observation.

The focus of the diagnostic tests in this review includes symptoms, vital signs, findings from physical examination, urine analysis tests and blood levels of inflammatory markers.

The diagnostic value of individual features is important to identify, clinicians often refer to these as 'red flags' for serious illnesses. Indeed, this is the categorisation that NICE used in its guideline on serious infection in children.³⁹ However, the combination of several features, referred to as a clinical prediction rule (also known as clinical prediction tool or clinical decision rule), may have better diagnostic value than individual features. In general, clinical prediction rules are most likely to be helpful in situations in which 'decision making is complex, the clinical stakes are high, or there are opportunities to achieve cost savings without compromising patient care'.⁴⁰ The management of children presenting to emergency and urgent-care settings with infections presents an ideal opportunity for application of a clinical prediction rule. The situation in which clinicians need to distinguish the very few seriously ill children from the vast majority of non-seriously ill children is very common; but individual experience with serious infections is more and more limited owing to the decreasing incidence and the consequences of missing a serious infection may be fatal. The ideal clinical prediction rules that this study aims to identify and validate will incorporate components of the history and basic examination findings including vital signs. This type of prediction rule is expected to be applicable to front-line clinicians, such as general practitioners (GPs), GP trainees, paramedics, practice nurses and ED medical and nursing staff. The advantage of the methods planned is that we intend to validate prediction rules in multiple clinical settings with varying prevalence of serious infection. This will form a test of the robustness of the prediction rules, and their generalisability and applicability to different acute paediatric settings in the NHS.

The marginal NHS cost of implementing a clinical prediction rule depends primarily on the cost of any additional staff time or investigations required. The prediction rules that we propose validating have very low marginal economic cost because the main components are an integral part of the standard clinical assessment of children that clinicians use in routine NHS practice (i.e. medical history, presenting complaints, vital signs and examination findings). In addition,

we anticipate that the systematic review will complement the current NICE guideline on the assessment of feverish children³⁹ by formally testing, simplifying and quantifying the accuracy of many of the clinical predictors used in that guideline.

This study focuses on children who are otherwise well and attending ED or primary-care settings. We will not address the important area of predicting infections in children with serious underlying health problems and particularly those who are immunocompromised and have markedly greater risks of morbidity and mortality from infections. In addition, we have limited this study to children older than 1 month as there are well-established guidelines for the management of infections in the neonatal period, and the clinical challenge of predicting infections in this period differs substantially in terms of clinical setting and expertise.

The main economic benefit to the NHS is the potential to reduce the need for urgent hospital admission by reliably identifying the vast majority of children who can safely be discharged home or to lower acuity care (e.g. GP follow-up). However, more effective triage using a formal prediction rule will also improve the care of children with serious infections (e.g. by signalling the need for emergency ambulance transfer to ED, for urgent diagnostic or therapeutic intervention, or for urgent paediatrician review), thus optimising the use and effectiveness of emergency services. In addition, identifying the optimal role of blood tests for inflammatory markers is important to ED settings in particular, as is balancing the diagnostic accuracy with the costs of these tests. If shown to discriminate effectively, such a prediction rule would be used at several levels of the emergency medical system in the UK, including by paramedics and in walk-in or out-of-hours surgeries, paediatric assessment units and EDs. Equally, of course, the consequence of a clinical prediction rule which is insufficiently specific is increasing inappropriate use of ED or secondary-care resources.

Chapter 3

Methodology of the systematic review

Literature search

We searched the literature in four electronic databases, MEDLINE, EMBASE, the Database of Abstracts of Reviews of Effects and the Cumulative Index to Nursing and Allied Health Literature, using search terms that included terms related to five components: serious infections, children, clinical history and examination, laboratory tests and ambulatory care settings (see *Appendix 1*). The searches were conducted in October 2008 with an update to the search in June 2009. We searched the Medion database (www.mediondatabase.nl) for systematic reviews using the 'signs and symptoms' component of our search strategy. Reference lists of included studies were searched for additional articles. We also checked reference lists of relevant NICE guidelines published prior to 2008^{39,41} and asked clinical experts to report any studies that had been omitted. We did not limit the search based on language or time.

Selection process

After a pilot selection on a sample of 20 studies, two independent reviewers (AVDB and THH), identified articles based on title and abstract for full-text review. Any discrepancies between the reviewers were resolved by a third independent reviewer (MJT). We used six criteria to determine inclusion and exclusion: study design, participants, setting, outcome, diagnostic tests and reporting of data (*Table 1*).

TABLE 1 Inclusion and exclusion criteria

Characteristic	Inclusion	Exclusion
Design	Studies assessing diagnostic accuracy or deriving prediction rules	Narrative reviews, letters, editorials, comments and case series of <20 subjects
Participants	Age between 1 month and 18 years Studies including children spanning this age range included if they reported age-stratified analyses (so that children aged <1 month or >18 years could be excluded) or if the proportion of children out of range was <50%	Children with pre-existing immune suppression (such as HIV infection or neutropenia due to chemotherapy) Outwith age range
Setting	Ambulatory care (defined as general or family practice, paediatric outpatient clinics, paediatric assessment units or EDs) Developed countries, defined using the United Nations list, which includes Europe, Canada, the USA, Australia, New Zealand and Japan	Studies conducted in developing countries
Outcome	Serious infection, defined as sepsis (including bacteraemia), meningitis, pneumonia, osteomyelitis, cellulitis, gastroenteritis with dehydration, complicated urinary tract infection (positive urine culture and systemic effects such as fever) and viral respiratory tract infections complicated by hypoxia (e.g. bronchiolitis)	Diagnosis other than serious infection
Diagnostic features	Clinical features Observation scales Clinical prediction rules Laboratory tests	Imaging Invasive tests (e.g. lumbar puncture, joint aspirates) Microbiological tests
Data reported	Reconstruction of the 2 × 2 tables possible	

HIV, human immunodeficiency virus.

Quality assessment

We assessed quality using the quality assessment of diagnostic accuracy studies (QUADAS) instrument.⁴² This was performed by one reviewer (AVDB) and checked by a second (THH), with discrepancies resolved by discussion involving additional researchers when necessary. In addition, we contacted study authors if additional clarification was necessary. Two items from the QUADAS criteria were used as study exclusion criteria:

- spectrum bias: this was considered present in case-control studies which used healthy controls or in studies in which participants were selected based on the performance of the reference standard
- validity of the reference standard: this was assessed by a clinical review committee consisting of a minimum of three researchers.

When there were insufficient data to be able to confidently assess whether or not a quality criterion had been met, we assessed it as not being met. The overall quality of included studies was rated from high (A) to low (D) using the following criteria:

- A: fulfilling all QUADAS criteria
- D: lacking total verification with the reference standard or with interpretation of the index feature unblinded to the results of the reference standard
- C: lacking an independent reference standard, with interpretation of the reference standard unblinded to the results of the index feature or with an unduly long time period between recording of the index feature and outcome
- B: all other studies.

Data extraction and analysis

Data were extracted by one reviewer (AVDB) and checked by a second (THH), and any errors identified were corrected. Key characteristics of included studies were extracted, including year, study design, setting or country, number of participants, proportion with serious illness, quality rating, age range, and inclusion and exclusion criteria. We reconstructed 2×2 tables based on information in the article or from additional information obtained from the authors. Where empty cells were present in the 2×2 tables, we added 0.5 to each cell. We calculated likelihood ratios (LRs) for the presence (LR+) or absence (LR-) of each feature along with the pre- and post-test probabilities of the outcome. Confidence intervals (CIs) were calculated on the basis of the standard error of a proportion using Stata version 9.2 (StataCorp LP, College Station, TX, USA).

Diagnostic features were categorised based on their diagnostic value as either red flags or as rule-out features for serious illness using the values of LR+ and LR-. Red flag (or rule-in) features were defined as those with a LR+ > 5.0. Rule-out features were defined as those with a LR- < 0.2.⁴³ In cases in which studies reported more than one result on the same feature using different cut-off points, we presented the result with the highest LR+ or lowest LR-. We included features if at least one study reported a LR+ > 5 or a LR- < 0.2. Setting was used to categorise studies, as a proxy for prevalence of serious infection: < 5%, low-prevalence setting; 5–20%, intermediate-prevalence setting; and > 20%, high-prevalence setting.

The results of the LR+, LR-, pre- and post-test probability and prevalence were plotted on dumbbell plots. We report both the pre-test and post-test probabilities of serious infection for each study in dumbbell plots. We grouped studies based on (a) global clinical assessment features,

(b) circulatory and respiratory features, (c) miscellaneous features, (d) inflammatory markers, (e) WBCs and (f) clinical prediction rules. The diagnostic value of temperature was explored using a plot of post-test values against pre-test prevalence on a log scale using R (www.r-project.org), using different cut-offs for temperature. Diagnostic meta-analysis of diagnostic markers across studies using multiple cut-points was carried out using the bivariate method of Dukic and Gatsonis,⁴⁴ implemented in R.

Chapter 4

Number and characteristics of studies included in the systematic review

Numbers of included studies

The electronic search of databases identified 1560 articles at the initial search date (October 2007) and a further 300 at the June 2009 update (*Figure 1*). After reviewing titles and abstracts, 176 of these articles were selected for full-text review. An additional 79 articles were identified based on reference lists of included studies, systematic reviews and NICE guidelines, and from content experts ('snowballing'). Following the review of the full text of these articles, we included 104 for quality assessment, of which 35 were selected for inclusion in the review. The most common reason for exclusion was spectrum bias.

One study⁷⁷ was excluded during the analysis stage because the 2×2 tables could not be reconstructed.

Setting of included studies

The majority of the studies were performed in the USA (16), with a further five from the UK, four from the Netherlands, three from Switzerland, two from Canada, and one each from Belgium, Italy, Australia, Denmark and Spain. Full details of the 35 included studies are shown in *Table 2*.

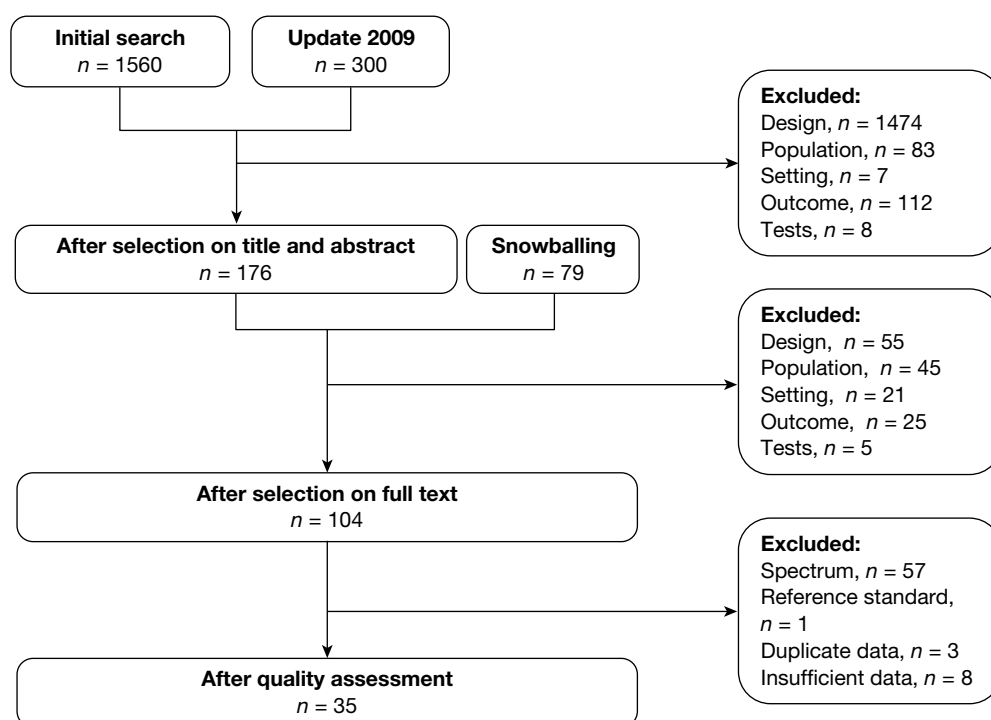


FIGURE 1 Flow of literature search.

TABLE 2 Characteristics of included studies

Study	Design	Setting, country	<i>n</i>	% serious illness	Quality rating	Age range	Inclusion criteria	Exclusion criteria
<i>Serious infections, composite outcome</i>								
1. Andreola 2007 ⁴⁵	Prospective, Cx, consecutive	ED, IT	408	23.0	C	<3 years	<i>Fever of uncertain source and increased risk of SBI, namely all infants aged 7 days to 3 months with rectal temperature >38 °C and children aged 3–36 months with ill/toxic appearance or with rectal temperature >39.5 °C</i>	Antibiotics or vaccination in 48 hours prior to enrolment, known immunodeficiencies, any chronic pathology, fever >5 days
2. Baker 1993 ⁴⁶	Prospective, Cx, consecutive	ED, USA	747	8.7	D	29–56 days	Temperature (rectal) ≥38.2 °C and immunocompetent	Not stated
3. Baker 1999 ⁴⁷	Prospective, Cx, consecutive	ED, USA	422	10.2	D	29–60 days	Temperature (rectal) ≥38.0 °C and immunocompetent	Not stated
4. Baker 1990 ⁴⁸	Prospective, consecutive	ED, USA	126	29.4	C	26–56 days	Temperature (rectal) >38.2 °C	Not stated
5. Berger 1996 ⁴⁹	Prospective, Cx, consecutive	ED, NL	138	23.9	B	2 weeks to 1 year	Temperature (rectal) ≥38.0 °C measured on the ward	Gestational age <37 weeks; perinatal complications; antibiotics or vaccination in previous 48 hours; known previous or underlying disease
6. Bleeker 2007 ⁵⁰	Prospective, Cx, consecutive	ED, NL	381	26.0	D	1–36 months	Referred to ED for fever <i>without</i> source, i.e. temperature ≥38 °C for which no clear focus could be identified after evaluation by the GP or after history taking by paediatrician	Not referred by GP; immune deficiencies
7. Bonadio 1993 ⁵¹	Prospective, Cx; consecutive	ED, USA	534	4.5	D	4–8 weeks	Temperature (rectal) ≥38 °C at triage, previously healthy	Not stated
8. Galetto-Lacour 2001 ⁵²	Prospective, Cx	ED, CH	124	22.6	D	7 days to 36 months	Temperature (rectal) >38.0 °C and <i>no</i> localising signs of infection from history or physical examination	Fever >7 days, neonates <1 week of age, children treated with antibiotics during the 2 preceding days, children with known immunodeficiencies
9. Galetto-Lacour 2003 ⁵³	Prospective, Cx	ED, CH	99	29.3	D	7 days to 36 months	Temperature (rectal) >38 °C and <i>without</i> localising signs of infection in their history or at physical examination	Fever >7 days, neonates <1 week of age, children treated with antibiotics during the 2 preceding days, children with known immunodeficiencies
10. Galetto-Lacour 2008 ⁵⁴	Prospective, Cx	ED, CH	202	26.7	D	7 days to 36 months	Temperature (rectal) >38.0 °C and <i>without</i> localising signs of infection in their history or at physical examination	Fever >7 days, neonates <1 week of age, children treated with antibiotics during the 2 preceding days, children with known immunodeficiencies

TABLE 2 Characteristics of included studies (continued)

Study	Design	Setting, country	<i>n</i>	% serious illness	Quality rating	Age range	Inclusion criteria	Exclusion criteria
11. Garra 2005 ⁵⁵	Prospective, Cx, consecutive	ED, USA	181	21.6	C	29–56 days	Temperature (rectal) ≥ 38.1 °C	Likely bacterial source for his or her fever on physical examination including cellulitis, mastitis, omphalitis, abscess, otitis media or septic arthritis
12. Grupo de Trabajo 2001 ⁵⁶	Prospective, Cx, consecutive	ED, ES	739	19.9	D	0–36 months	Temperature (rectal) ≥ 38 °C	Antibiotics or DTP within 48 hours or MMR within 10 days; systemic central nervous condition; concomitant analytical changes in blood that interfere with interpretation of full blood count; fever duration > 72 hours; chronic illness
13. Hsiao 2006 ⁵⁷	Prospective, Cx, consecutive	ED, USA	429	10.3	C	57–180 days	Temperature (rectal) > 37.9 °C	Not stated
14. McCarthy 1987 ⁵⁸	Prospective, Cx, consecutive	ED, USA	143	19.6	C	< 24 months	Temperature ≥ 38.3 °C	Not stated
15. McCarthy 1982 ³⁶	Prospective, Cx, consecutive	ED, USA	165	15.8	C	< 24 months	Temperature ≥ 38.3 °C	Not stated
16. Nademi 2001 ⁵⁹	Prospective, Cx, consecutive	PAU, UK	141	29.1	D	0–16 years	Temperature ≥ 38 °C	Temperature < 38 °C
17. Thayyil 2005 ⁶⁰	Prospective, Cx, consecutive	PD, UK	72	11.1	D	1–36 months	Temperature > 39 °C without localising signs	Antibiotics 72 hours prior to enrolment, immunodeficiencies, fever > 7 days
18. Thompson 2009 ²⁹	Prospective, Cx, consecutive	PAU, UK	700	55.3	C	3 months to 16 years	Suspicion of acute infection	Children with diseases liable to cause repeated serious bacterial infection, and infections resulting from penetrating trauma
19. Trautner 2006 ⁶¹	Prospective, Cx,	ED, USA	103	19.4	C	< 17 years	Temperature (rectal) ≥ 41.1 °C	None
20. Van den Bruel 2007 ¹¹	Prospective, Cx, consecutive	GP-APC-ED, BE	3981	0.78	C	< 17 years	Acute illness for a maximum of 5 days	Traumatic or neurological illness, intoxication, psychiatric or behavioural problems without somatic cause, exacerbation of a chronic condition
Bacteraemia								
21. Crocker 1985 ⁶²	Prospective, Cx, consecutive	ED, USA	201	10.5	C	6 months to 2 years	Temperature (rectal) ≥ 39.4 °C	Viral exanthem, enanthem, croup, vomiting, diarrhoea, admitted with a diagnosis of meningitis or sepsis
22. Haddon 1999 ⁶³	Prospective, Cx,	ED, AU	526	3.4	C	3–36 months	Temperature (tympanic) ≥ 39 °C	Varicella, croup or herpes gingivostomatitis

continued

TABLE 2 Characteristics of included studies (continued)

Study	Design	Setting, country	n	% serious illness	Quality rating	Age range	Inclusion criteria	Exclusion criteria
23. Jaffe 1991 ⁶⁴	Prospective, Cx	ED, CA	955	2.8	C	3–36 months	Temperature (rectal) $\geq 39.0^{\circ}\text{C}$	Focal infection requiring immediate antibiotic; clinical appearance necessitating immediate hospitalisation; specific viral infections; known immune-deficiency condition or chronic illness; antibiotic or DTP within preceding 48 hours
24. Osman 2002 ⁶⁵	Prospective, consecutive	ED, UK	1547	38.0	D	0–14 years	All children with an infectious illness	Not stated
25. Teele 1975 ⁶⁶	Prospective, Cx, consecutive	ED, USA	600	3.2	C	4 weeks to 2 years	Temperature (rectal) $\geq 38.3^{\circ}\text{C}$	Prior medical evaluation or referral from other physician or from other clinic
26. Waskerwitz 1981 ⁶⁷	Prospective, Cx, consecutive	ED, USA	292	5.8	B	<24 months	Temperature (rectal) $\geq 39.5^{\circ}\text{C}$	Not previously healthy; weight less than third percentile or known chronic disease
Gastroenteritis causing dehydration								
27. Gorelick 1997 ⁶⁸	Prospective, Cx	ED, USA	186	33.4	C	1 month to 5 years	Chief complaint of vomiting, diarrhoea or poor oral fluid intake	Symptoms >5 days; history of cardiac or renal disease or diabetes mellitus; malnutrition or failure to thrive; treatment within 12 hours in other health facility; hyponatraemia or hypernatraemia; tonsillectomy within 10 days; no telephone or beeper for follow-up
28. Shavit 2006 ⁶⁹	Prospective	ED, CA	83	15.7	C	1 month to 5 years	History of diarrhoea (with or without vomiting) for ≤ 5 days and judged by the ED triage nurse to have some degree of dehydration	History of cardiovascular or renal disease; judged by the triage nurse to require emergent medical intervention
Meningitis								
29. Joffe 1983 ⁷⁰	Retrospective	ED, USA	241	5.4	D	6 months to 6 years	First episode of fever and seizures	Did not undergo lumbar puncture and final outcome was not available; children with a predisposition to meningitis
30. Offringa 1992 ⁷¹	Retrospective, consecutive	ED, NL	309	7.4	C	3 months to 6 years	First episode of fever and seizures	Not stated
31. Oostenbrink 2001 ⁷²	Retrospective	ED, NL	256	38.7	C	1 month to 15 years	Signs of meningeal irritation	Patients with a history of severe neurological disease or ventricular drainage, and those referred from other hospitals

TABLE 2 Characteristics of included studies (*continued*)

Study	Design	Setting, country	<i>n</i>	% serious illness	Quality rating	Age range	Inclusion criteria	Exclusion criteria
<i>Pneumonia</i>								
32. Mahabee-Gittens 2005 ⁷³	Prospective, Cx	ED, USA	510	8.6	A	2–59 months	Cough and more than one of laboured, rapid or noisy breathing, chest or abdominal pain, or fever	Currently taking AB, smoke inhalation, foreign body aspiration or chest trauma; known diagnoses of asthma, bronchiolitis, sickle cell disease, cystic fibrosis, chronic cardiopulmonary disease
33. Taylor 1995 ⁷⁴	Prospective, Cx, consecutive	ED, USA	572	7.3	D	< 2 years	Temperature ≥ 38.0 °C	Acute wheezing and/or stridor, history of chronic pulmonary disease, chest radiograph interpreted as indeterminate by both radiologists ($n=2$), clinical diagnosis of pneumonia with no radiograph obtained ($n=2$)
<i>Meningococcal infection</i>								
34. Nielsen 2001 ⁷⁵	Prospective, Cx, consecutive	PD, DK	208	18.8	C	> 1 month to < 16 years	Haemorrhages in the skin detected at admission or during hospital stay, plus rectal temperature > 38 °C within 24 hours of admission	Second or more inclusion in the study
35. Wells 2001 ⁷⁶	Prospective, Cx, consecutive	ED, UK	218	11.0	C	≤ 15 years	Non-blanching rash	Not stated

APC, ambulatory paediatric care; AU, Australia; BE, Belgium; CA, Canada; CH, Switzerland; Cx, cross-sectional; DK, Denmark; DTP, diphtheria tetanus pertussis; ES, Spain; GP, general practice; IT, Italy; MMR, measles, mumps and rubella; NL, the Netherlands; PAU, paediatric assessment unit; PD, paediatric department; SBI, serious bacterial infections.

There was only a single trial from general practice (which also recruited non-referred patients from ambulatory paediatric care and the ED in Belgium).¹¹ All other studies were performed in EDs, with four of these in paediatric departments (one a paediatric assessment unit).^{29,59,60,75}

Age and prevalence of serious infection

The age of eligible children also varied among the studies, with an upper age limit ranging from 3 months to 18 years. The median prevalence of serious infection was 15.8% (interquartile range 8.66–22.06%), ranging from 0.8% in the general practice study¹¹ to 55.3% in one study at a paediatric assessment unit.²⁹

Outcomes reported

The majority (20/35) of studies reported on a composite outcome of serious infection, which included sepsis, bacteraemia, meningitis, pneumonia and urinary tract infection (and in some cases additional infections such as cellulitis, osteomyelitis and abscess).^{11,29,36,45-61} All but two^{18,20} of these included children based on the presence of fever and six cases^{45,50,52-54,60} were based on the absence of localising signs or focus of infection. A further six studies reported outcomes on bacteraemia,⁶²⁻⁶⁷ and five of them^{62-64,66,67} included children with fever. Three studies used an outcome of meningitis,⁷⁰⁻⁷² of which two included children with first episode of fever and seizures,^{70,71} and one based on signs of meningeal irritation.⁷² Two studies used an outcome of gastroenteritis causing dehydration,^{68,69} one of which included children with clinical gastroenteritis⁶⁸ and the other children with diarrhoea with or without vomiting and evidence of dehydration.⁶⁹ Two studies used an outcome of pneumonia,^{73,74} and included febrile children and children with clinical signs of respiratory infection. Finally, two studies^{75,76} used meningococcal infection as an outcome, both of which included haemorrhagic or non-blanching rash as inclusion criteria.

Quality of included studies

All but three of the studies used prospective designs; the remainder were retrospective. The majority (24) used consecutive recruitment. The quality of the included studies was modest (*Figure 2*). As representativeness of spectrum and valid reference standard were used as inclusion criteria, all studies met these quality criteria. The majority of studies were scored as yes or unclear for the criteria of index text interpreted blinded from reference standard, cut-offs were defined before study, definitions of positive results provided, treatment paradox absent, sample size exceeding 50, pre-specified objectives, and consecutive inclusion. The quality criteria filled by the fewest studies were blind reading of the reference standards which was explicitly mentioned in four studies, and scored as unclear in 21 studies.

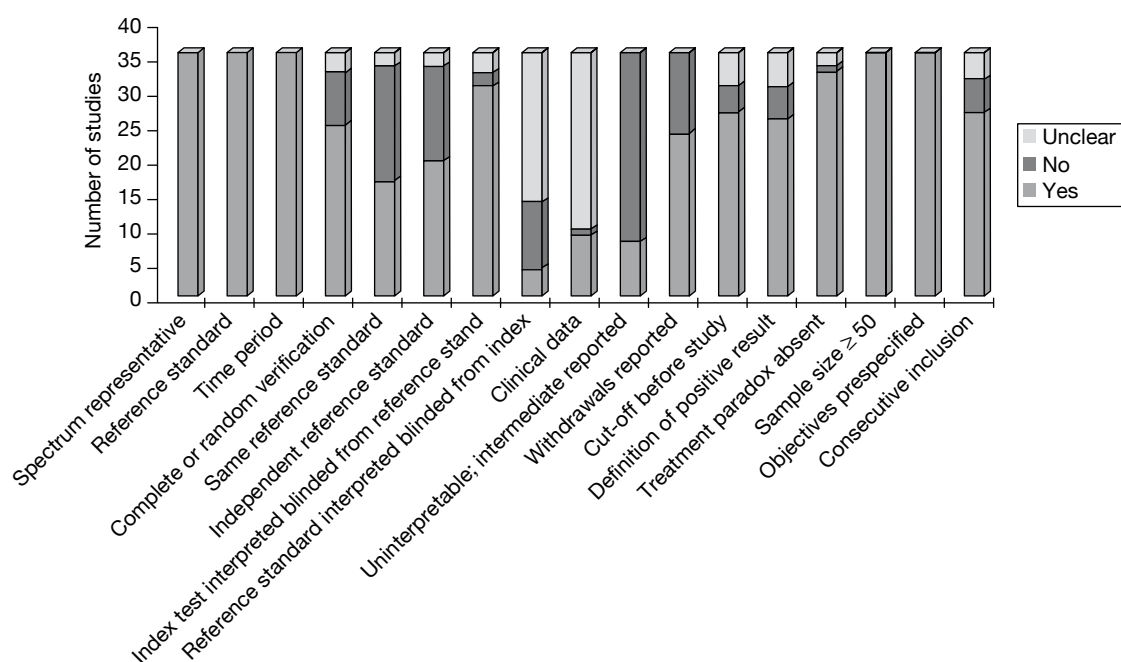


FIGURE 2 Summary of QUADAS features of studies included in systematic review.

Chapter 5

Results of systematic review of clinical features

A total of 30^{11,29,36,45,48–50,52,53,56–76} studies identified in the systematic review presented the diagnostic value of clinical features for serious infections as a composite outcome or for a specific infection. Clinical features included global assessment, child behaviour, circulatory and respiratory signs, neurological signs or petechial rash, fever, miscellaneous other clinical features and clinical prediction rules. The results of this analysis have also been published in *The Lancet*.⁷⁸

Global assessment

Parental concern that the illness is different from previous illnesses (LR+ 14) and the clinician's gut feeling that something is wrong (LR+ 23) were important red flags in a single study from a low-prevalence setting (*Figure 3*).¹¹ Overall assessment of severity based on clinical impression or the child appearing ill from six further studies^{50,56,63,67,69,76} (all but one in intermediate-prevalence setting) provided weaker LR+ and LR-.

Child behaviour

Change in the child's crying pattern, drowsiness, moaning or inconsolability all had a LR+ > 5.0 and thus potential red flags, but again all were from the sole low-prevalence study.¹¹ In an additional three studies from intermediate or high prevalence settings,^{50,71,72} these features provided weak LR+ and LR- and, paradoxically, changed crying reduced the probability of serious disease in a high-prevalence setting.

Circulatory and respiratory clinical features

The presence of cyanosis had a LR+ ranging from 2.66 to 52.2 in three studies,^{11,29,72} suggesting a rule-in value (*Figure 4*). Poor peripheral circulation was noted in six studies, with a LR+ ranging from 2.39 to 38.8.^{11,29,50,68,69,72} There did not appear to be a clear relationship between the prevalence of infection in the study setting and LR+. Rapid breathing and shortness of breath provided greatest LR+ in the single low-prevalence study,¹¹ with a LR+ of 9.3–9.7. In the four other studies that assessed these features, the LR+ was weaker, ranging from 1.11 to 3.6.^{29,50,73,74} The presence of crackles on auscultation and diminished breath sounds again provided a LR+ > 5 in the low-prevalence setting (LR+ 6.0–9.3),¹¹ but little value in the single study conducted in an intermediate-prevalence setting.⁷³

Neurological signs or petechial rash

The presence of meningeal irritation, petechial rash, decreased consciousness and seizures all had a LR+ > 5 in the majority of the eight studies which assessed these features (*Figure 5*).^{11,60,70–72,75,76} This effect appeared to be relatively consistent across all settings. Meningeal irritation had a

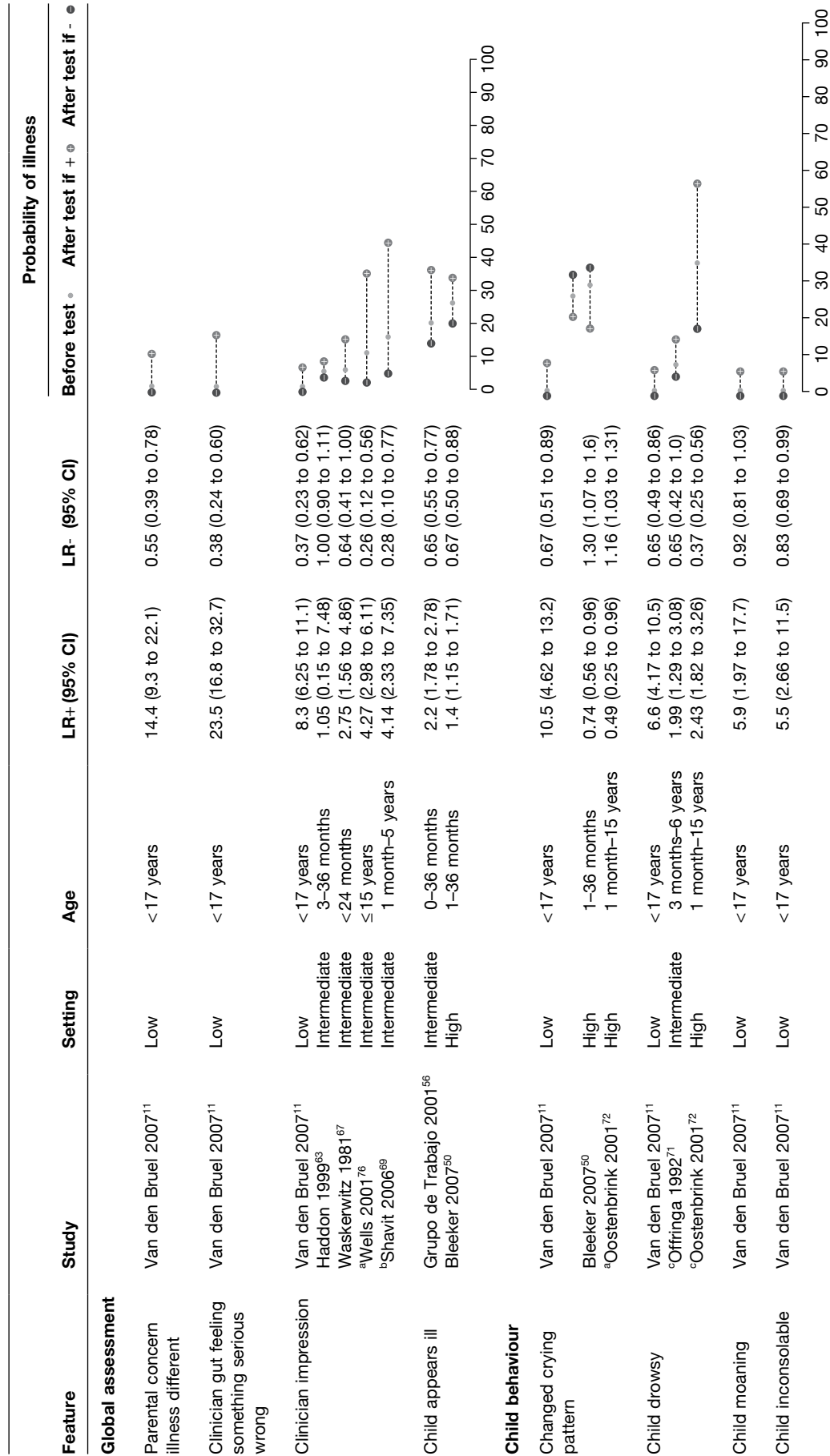


FIGURE 3 Potential red flag features for serious illness (LR+ > 5 in at least one study): global assessment and behavioural features. Setting: low, low prevalence of serious illness (<5%); intermediate, intermediate prevalence of serious illness (5–20%); high, high prevalence of serious illness (>20%). a, Meningococcal infection; b, gastroenteritis causing dehydration only; c, meningitis only.

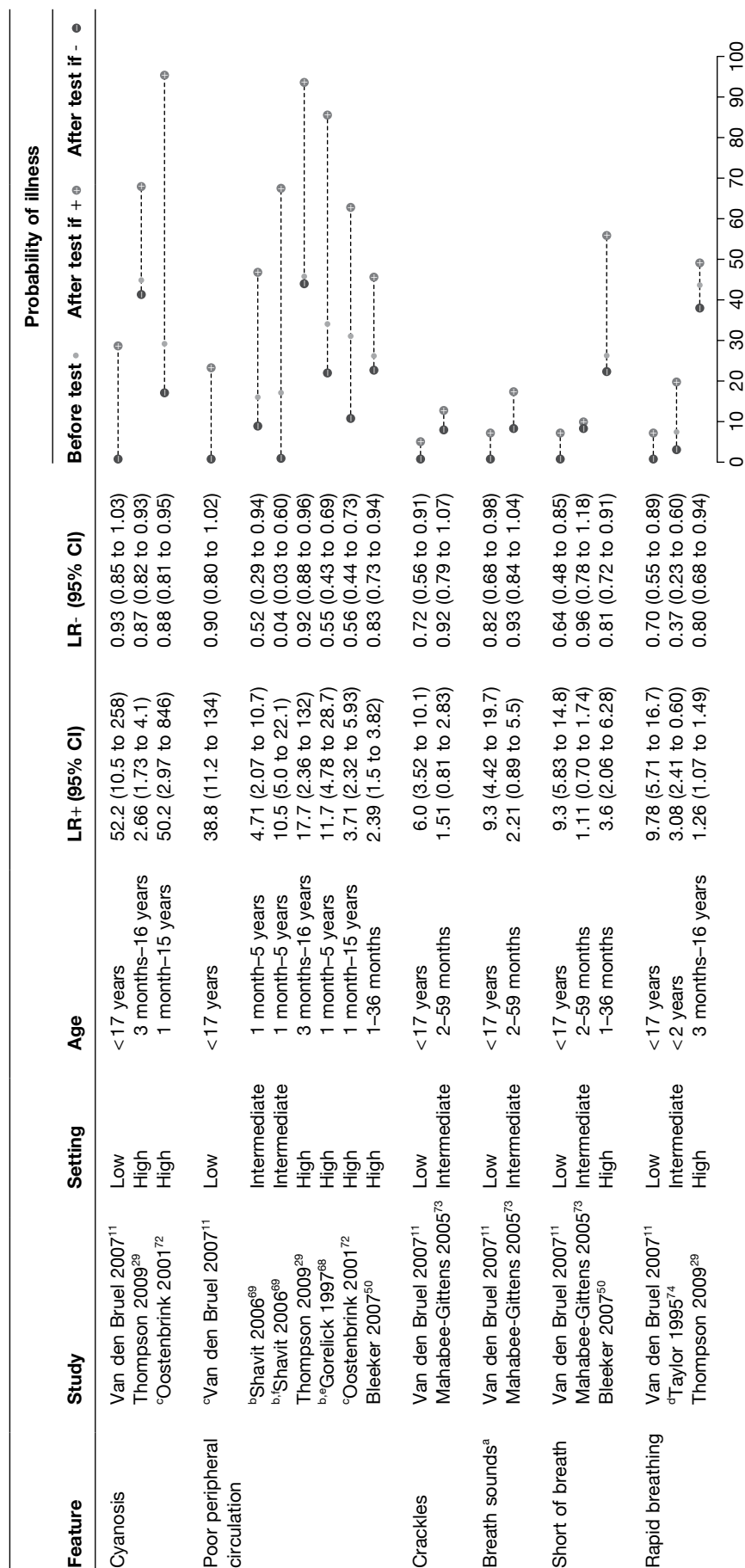


FIGURE 4 Potential red flag features for serious illness (LR+ >5 in at least one study); circulatory and respiratory features. Setting: low, low prevalence of serious illness (<5%); intermediate, intermediate prevalence of serious illness (5-20%); high, high prevalence of serious illness (>20%). a, Meningococcal infection; b, gastroenteritis causing dehydration only; c, meningitis only; d, pneumonia only; e, capillary refill > 2 seconds; f, digitally measured capillary refill.

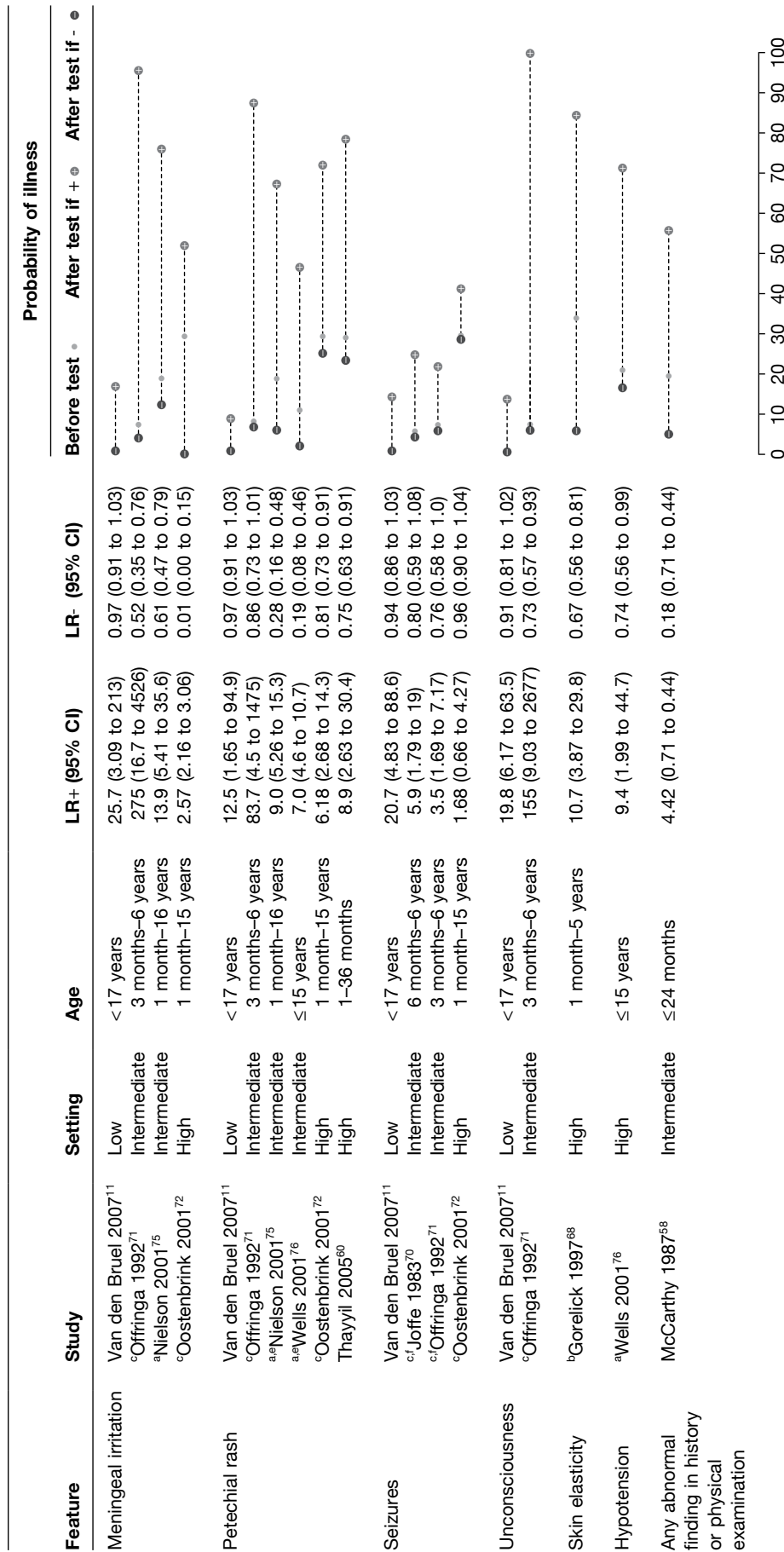


FIGURE 5 Potential red flag features for serious illness (LR+ > 5 in at least one study); neurological and miscellaneous features. Setting: low, low prevalence of serious illness (<5%); intermediate, intermediate prevalence of serious illness (5-20%); high, high prevalence of serious illness (>20%). a, Meningococcal infection; b, gastroenteritis causing dehydration only; c, meningitis only; d, pneumonia only; e, diameter >2 mm; f, during examination.

LR+ ranging from 2.57 to 275 in four studies.^{11,71,72,75} The presence of a petechial rash had a LR+ of 6.18–83.7 in six different studies.^{11,60,71,72,75,76} Seizures were a useful red flag in the study¹¹ in a low-prevalence setting (LR+ 20.7),¹¹ whereas in three studies in intermediate- or high-prevalence settings it was less useful (LR+ 1.68–5.9).^{70–72} Loss of consciousness was assessed in only two studies, in which it had a LR+ of 19.8–155.^{11,71}

Fever

Data from 11 studies^{11,29,49,50,59,64–66,72,74,76} were available to plot on a graph of pre-test probability (i.e. prevalence) versus post-test probability (*Figure 6*) using cut-offs with the highest LR+. The highest rule-in value was obtained in the setting with the lowest prevalence, where a temperature of 40°C or more increased the likelihood of disease from 0.8% to 5.0%.¹¹ By contrast, the absence of high temperature (< 38.5°C to 38.9°C) had greatest rule-out value in a study with prevalence of serious infection of 29.1%.⁵⁹ However, this rule-out potential was not seen in any of the other five studies^{29,49,50,65,72} with prevalence > 20% and temperature had no rule-in value in these high prevalence settings.

When we repeated the analysis using additional cut-offs (data not presented in *Figure 6*), results were very similar to the results presented in *Figure 2*. In the low-prevalence study,¹¹ temperature $\geq 38^\circ\text{C}$ had a LR+ of 1.5 and a LR- of 0.38 and temperature $\geq 39^\circ\text{C}$ has a LR+ of 2.3 and a LR- of 0.59.

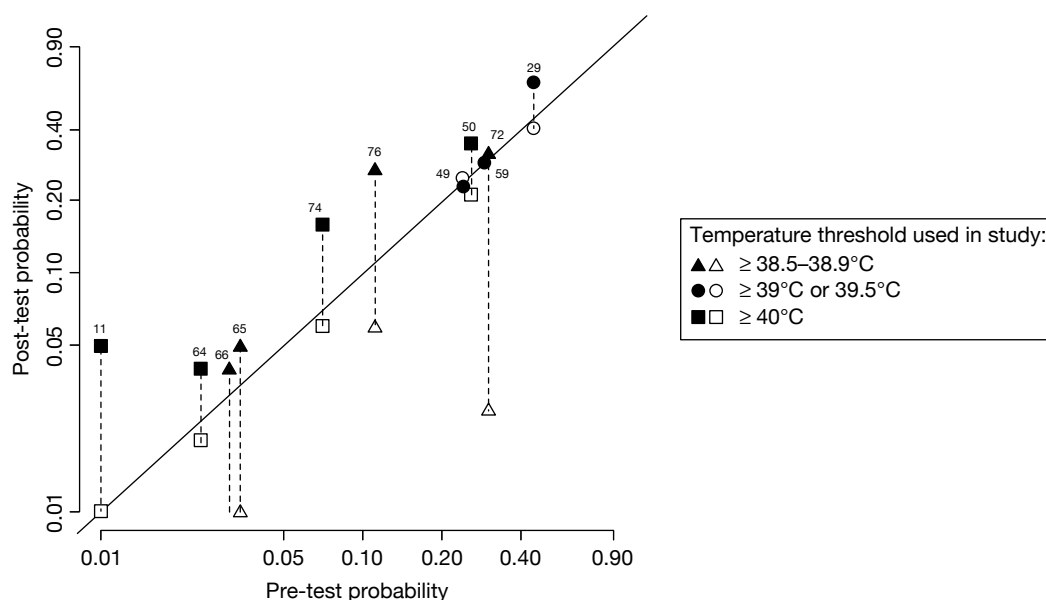


FIGURE 6 Probability of serious illness in children with temperature above (closed symbols) or below (open symbols) threshold in 11 studies^{11,29,49,50,59,64–66,72,74,76} carried out in health-care settings with different pre-test probabilities of serious infection and using different temperature thresholds. The distance of the symbol from the 45°-line indicates the diagnostic value of temperature measurement in the study (applying the specified threshold). The figure is plotted on a log scale to achieve visual separation of the studies carried out in low-prevalence settings (Teale *et al.*,⁶⁶ the estimated post-test probability if the temperature was < 38.5°C was 0%, which cannot be plotted on a log scale, so there is no lower symbol).

Clinical prediction rules

We identified clinical prediction rules which had been applied to a composite outcome of serious infection, pneumonia, meningitis and dehydration from gastroenteritis (see *Appendix 2*). The Yale Observation Scale (YOS) was assessed in seven studies (*Figure 7*). It provided a LR⁻ < 0.2 in two studies,^{36,60} whereas in five other studies it varied from 0.68 to 0.94,^{45,48,52,53,57} associated with post-test probabilities ranging from 10% to 28%. The YOS was derived in the McCarthy *et al.* 1982 study³⁶ at a time prior to routine vaccination against *Haemophilus influenzae* and pneumococcus, possibly explaining its better results. This would not explain the similar low LR⁻ of the Thayyil *et al.* study,⁶⁰ which was performed on a similar patient population in 2003. For meta-analysis, there was significant heterogeneity present ($p = 0.002$), which remained ($p = 0.026$) after exclusion of the McCarthy *et al.* study,³⁶ but disappeared ($p = 0.093$) after exclusion of both studies. The summary sensitivity of the five remaining studies^{45,48,52,53,57} was 32.5% (95% CI 21.7% to 45.5%) and specificity was 78.9% (95% CI 73.9% to 83.1%), theoretically corresponding to a LR⁺ of 1.54 and a LR⁻ of 0.86. An additional study⁵⁸ assessed the YOS or the presence of any abnormal finding on history or clinical examination in an intermediate population; again it provided a rule-out value with a LR⁻ of 0.17.

The clinical prediction rule that performed best for ruling out serious infection (LR⁻ 0.04) involved five sequential questions, but was assessed in only a single low-prevalence study.¹¹ Two prediction rules were identified which potentially ruled out pneumonia, with LR⁻ 0.07, both involved the absence of shortness of breath and either parental or clinician concern. Again, these were assessed in only a single low-prevalence study.¹¹ We identified two prediction rules for meningitis derived in intermediate settings. One provided a rule-out value (LR⁻ 0.05) if there was no abnormal neurological finding and care was not sought within 48 hours, whereas the other provided rule-in ability (LR⁺ 395) if any of petechiae, nuchal rigidity or coma was present. Finally, a single rule was identified for dehydration from gastroenteritis, which provided modest LR⁺ (6.10) and LR⁻ (0.24) from a single high-prevalence study.⁶⁸

Features of limited help in ruling in or ruling out serious infections

Features that had a LR⁺ < 5.0 and/or a LR⁻ > 0.2 (i.e. less helpful in either ruling in or ruling out serious infection) for any serious infection are listed in *Table 3*. The NICE traffic light system and the Manchester Triage System score were of little value in a single study with a high-prevalence setting. Symptoms that are common in children, such as cough, headache, tummy ache, vomiting, diarrhoea, poor feeding and coryza, had little diagnostic value. Two features, namely failure to smile (LR⁺ 4.2) and changed breathing pattern (LR⁺ 4.4), were just below the arbitrary LR⁺ cut-off of 5, but both had a weak LR⁻, suggesting no rule-out value (LR⁻ 0.6–0.7). The behavioural features of a reactive child (i.e. moving, reaching for objects, looking around the room), lack of irritability and lack of sleepiness did not provide a rule-in or rule-out value based on three studies.^{11,49,59} Although cyanosis and poor peripheral perfusion (which causes mottling and pallor) are red flag symptoms as described above, the report of abnormal skin colour (described as cyanotic, pallor or flushed/mottled) in three studies of low and high prevalence did not provide a useful LR.^{11,49,50}

The clinical features that were of limited rule-in or rule-out value for specific infections are listed in *Table 4*. Several clinical features in a single study of dehydration from gastroenteritis (low urine output, sunken eyes, dry mucous membranes, tachycardia, abnormal respiration) all had modest LR⁺ (LR⁺ 1.8–3.8). The presence of paralysis or paresis had only a modest LR⁺ for meningitis (3.48).

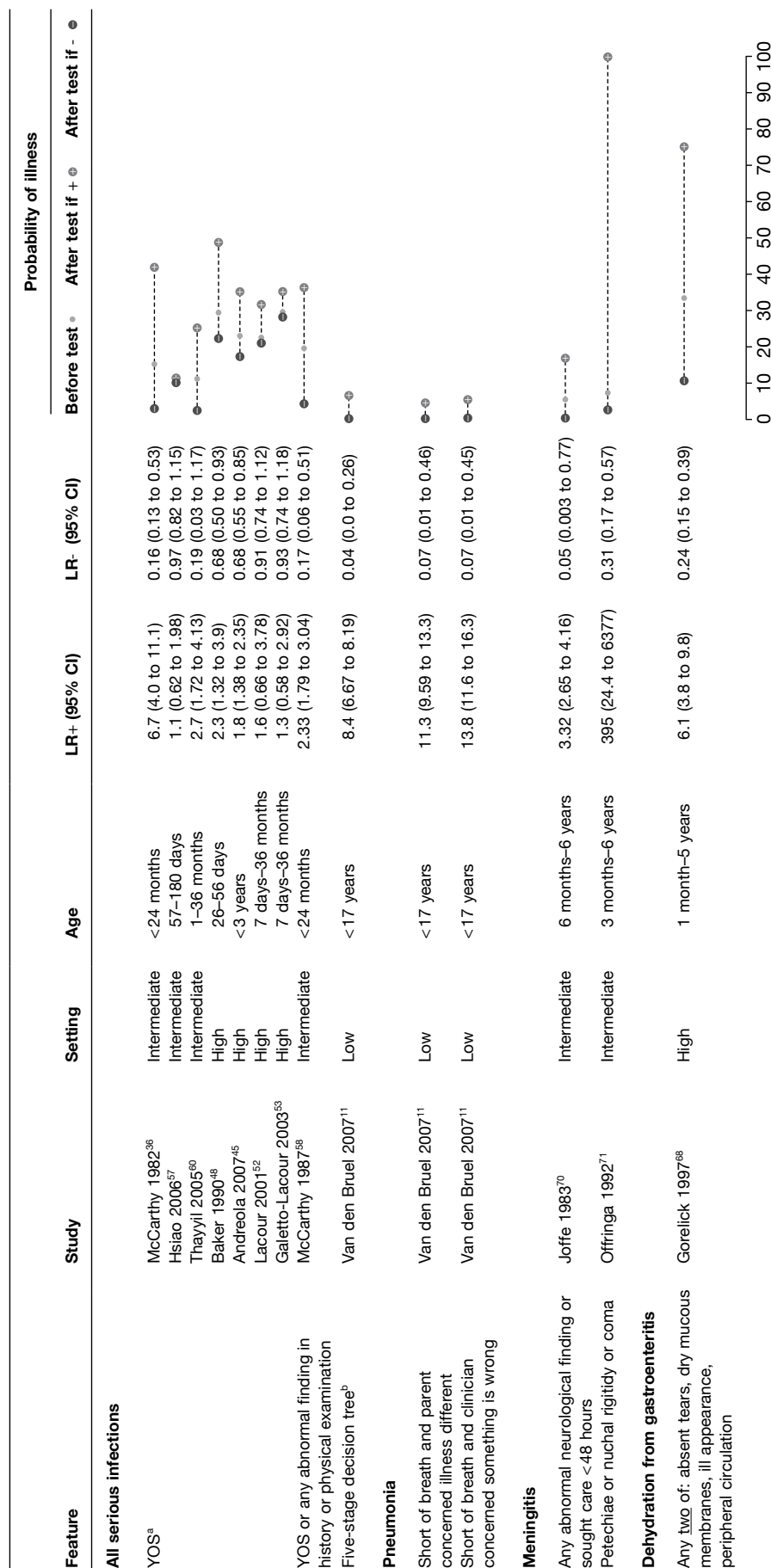


FIGURE 7 Clinical prediction rules with the potential to rule in or rule out serious infection (LR+ > 0.5 or LR- < 0.2 in at least one study). Setting: low, low prevalence of serious illness (<5%); intermediate, intermediate prevalence of serious illness (5–20%); high, high prevalence of serious illness (>20%). a, Cut-off used: >8, study 48; >9, study 45; >10, studies 36, 52, 53, 57, 60. b, If yes to any of five sequential questions: (1) clinician instinct that something is wrong; (2) dyspnoea; (3) temperature >39.5°C; (4) diarrhoea; (5) age 15–29 months.

TABLE 3 Presenting features of limited help in ruling in (LR+ >5) or ruling out (LR- <0.2) any serious infection

Feature	Study	Setting	LR+	LR-
Global assessment				
No obvious source of fever	Hsiao 2006 ⁵⁷	Int	3.04	0.87
Prediction rule ^a	Bleeker 2007 ⁵⁰	High	2.1	0.38
NICE traffic light system ^b	Thompson 2009 ²⁹	High	1.20	0.50
Manchester triage system	Thompson 2009 ²⁹	High	1.35	0.43
Prediction rule ^c	Thompson 2009 ²⁹	High	1.31	0.52
Child behaviour				
Child no longer smiles	Van den Bruel 2007 ¹¹	Low	4.24	0.64
Child is irritable	Van den Bruel 2007, ¹¹ Nademi 2001 ⁵⁹	Low and high	1.33–2.34	0.57–0.86
Child is somnolent	Van den Bruel 2007 ¹¹	Low	2.25	0.81
Child is reactive ^d	Berger 1996 ⁴⁹	High	1.33–1.97	0.56–0.79
Respiratory signs				
Changed breathing pattern	Van den Bruel 2007 ¹¹	Low	4.43	0.67
Cough	Van den Bruel 2007 ¹¹	Low	1.30	0.73
Signs of URTI	Van den Bruel 2007, ¹¹ Trautner 2006 ⁶¹	Low and int	0.46–0.99	1.01–2.21
Gastrointestinal signs				
Diarrhoea	Van den Bruel 2007, ¹¹ Berger 1996, ⁴⁹ Trautner 2006 ⁶¹	Low, int and high	0.99–2.91	0.69–1.00
Vomiting	Van den Bruel 2007, ¹¹ Bleeker 2007, ⁵⁰ Nademi 2001, ⁵⁹ Trautner 2006 ⁶¹	Low, int and high	0.83–1.60	0.69–1.10
Signs of dehydration ^e	Van den Bruel 2007, ¹¹ Bleeker 2007 ⁵⁰	Low and high	1.07–2.49	0.98
Poor feeding	Van den Bruel 2007, ¹¹ Nademi 2001 ⁵⁹	Low and high	1.37–1.54	0.51–0.83
Other signs				
Age	Berger 1996, ⁴⁹ Hsiao 2006, ⁵⁷ Trautner 2006 ⁶¹	Int and high	0.98–2.49	0.77–1.01
Underlying condition	Trautner 2006 ⁶¹	Int	2.42	0.76
Duration of fever/illness	Van den Bruel 2007, ¹¹ Andreola 2007, ⁴⁵ Berger 1996, ⁴⁹ Bleeker 2007, ⁵⁰ Trautner 2006 ⁶¹	Low, int and high	0.76–2.18	0.74–1.53
Abnormal skin colour	Van den Bruel 2007, ¹¹ Berger 1996, ⁴⁹ Bleeker 2007 ⁵⁰	Low and high	1.59–1.95	0.61–0.97
Tummy ache	Van den Bruel 2007 ¹¹	Low	0.41	1.15
Headache	Van den Bruel 2007 ¹¹	Low	0.23	1.20
Tachycardia ^f	Thompson 2009 ²⁹	High	1.49–2.05	0.65–0.85

APLS, advanced paediatric life support; URTI, upper respiratory tract infection.

a Duration of fever (days), history of vomiting, ill clinical appearance, chest-wall retractions ± tachypnoea, poor peripheral circulation.

b One or more red or amber feature.

c One or more of temperature ≥ 39 °C, oxygen saturation ≤ 94%, tachycardia and tachypnoea.

d Moving limbs, reaching for objects, looking around the room; in isolation or in combination.

e Other than skin inelasticity.

f APLS age-specific cut-offs or heart rate > 90th centile.

Setting: low, low prevalence of serious illness (<5%); int, intermediate prevalence of serious illness (5–20%); high, high prevalence of serious illness (>20%).

TABLE 4 Presenting features of limited help in ruling in (LR+ <5) or ruling out (LR- <0.2) specific infections

Feature	Study	Setting	LR+	LR-
<i>Bacteraemia</i>				
Child is irritable	Crocker 1985 ⁶²	Int	1.48	0.61
Child is lethargic	Crocker 1985 ⁶²	Int	0.64	1.10
Functional status ^a	Waskerwitz 1981 ⁶⁷	Int	1.21–2.57	0.26–0.55
Age (various cut-offs)	Teele 1975 ⁶⁶	Low	0.33–1.83	0.66–1.13
Referral status	Haddon 1999 ⁶³	Low	1.74	0.79
<i>Meningitis</i>				
Child is irritable	Oostenbrink 2001 ⁷²	High	0.76	1.05
Vomiting	Offringa 1992 ⁷¹	Int	2.53	0.64
Duration of fever/illness	Offringa 1992 ⁷¹	Int	1.43	0.81
Sought care in previous 48 hours	Joffe 1983, ⁷⁰ Offringa 1992 ⁷¹	Int	2.28–2.92	0.64–0.73
Paresis or paralysis	Offringa 1992 ⁷¹	Int	3.48	0.76
<i>Meningococcal infection</i>				
Cough	Nielsen 2001 ⁷⁵	Int	0.41	1.35
Vomiting	Nielsen 2001 ⁷⁵	Int	1.08	0.94
<i>Pneumonia</i>				
Grunting	Mahabee-Gittens 2005 ⁷³	Int	0.56	1.02
Wheezing	Mahabee-Gittens 2005 ⁷³	Int	1.25	0.95
Duration	Mahabee-Gittens 2005 ⁷³	Int	1.03	0.93
<i>Dehydration from gastroenteritis</i>				
Abnormal respirations	Gorelick 1997 ⁶⁸	High	3.10	0.66
Tachycardia	Gorelick 1997 ⁶⁸	High	2.18	0.68
Abnormal radial pulse	Gorelick 1997 ⁶⁸	High	3.10	0.66
Sunken eyes	Gorelick 1997 ⁶⁸	High	3.71	0.47
Dry mucous membranes	Gorelick 1997 ⁶⁸	High	3.62	0.26
Decreased urine output	Gorelick 1997 ⁶⁸	High	1.82	0.27

a With or without clinician impression of bacteraemia.

Setting: low, low prevalence of serious illness (<5%); int, intermediate prevalence of serious illness (5–20%); high, high prevalence of serious illness (>20%).

Chapter 6

Results of the systematic review of laboratory tests for serious infections

A total of 14 studies^{45–47,49–55,57,59–61} identified in the systematic review reported the diagnostic value of laboratory tests for serious infections. (Published in Van den Bruel *et al.* Diagnostic value of laboratory tests in identifying serious infections in febrile children: a systematic review. *BMJ* 2011;**342**:D3082.)

Diagnostic value of laboratory tests for composite outcome of serious infection

Inflammatory markers

Seven studies identified in the systematic review^{45,49,52–54,57,60} provided data on the diagnostic value of six inflammatory markers: procalcitonin (PCT), C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), interleukin-6, interleukin-8 and interleukin-1 receptor antagonist (Figure 8). None of the studies was performed in low-prevalence settings. The three studies^{45,54,60} which reported the results of PCT all used the same cut-off (0.5 ng/ml) and demonstrated a LR+ of 1.75–2.96, with a LR– of 0.08–0.35. The five studies of CRP^{45,49,54,57,60} had cut-offs ranging from 9.8 to 50 mg/l, and provided a LR+ of 2.53–3.79 and a LR– of 0.25–0.61. Pooling of CRP results was possible and using bivariate meta-analysis of CRP yielded a pooled LR+ of 3.15 (95% CI 2.67 to 3.71) and a pooled LR– of 0.33 (95% CI 0.22 to 0.49) across all cut-offs.

The summary receiver operator characteristics curve plots of CRP and PCT (Figure 9) confirm that the markers have similar diagnostic accuracy as the shapes of the curves are very similar and the CIs are overlapping. We plotted the LRs of CRP and PCT identified in the systematic review by cut-off value and disease prevalence to identify the optimal cut-off points to apply. To rule in serious infection, cut-off levels of 2 ng/ml for PCT or 80 mg/l for CRP both provide good diagnostic value (a LR+ of ≥ 4) but poor rule-out value. To rule out effectively, cut-offs of 0.5 ng/ml for PCT or 20 mg/l for CRP may be a better choice (a LR– of ≤ 0.2).

The performance of ESR in a single study⁴⁹ (cut-off of 50 mm/hour) provided LR+ 2.49 and LR– 0.34. The two studies investigating the three interleukins demonstrated somewhat inferior diagnostic values to CRP or PCT.^{52,53}

White blood cell counts

A total of nine studies^{45,49,54,57,59–61} provided data on either total WBC (seven studies^{45,49,54,57,59–61}), absolute neutrophil count (two studies^{45,61}), band count (three studies^{49,52,53}) or left shift (single study⁵⁴); all were from intermediate- or high-prevalence settings (Figure 10). All four WBC indices demonstrated little diagnostic value in ruling out serious infection: the minimum LR– is 0.61 with the 95% CI in most studies crossing 1.0. The LR+ were also weaker than the inflammatory markers, with values ranging from 0.87 to 3.05. There was little evidence to determine whether or not any of the WBC markers offered superior diagnostic value. The results of one study⁶¹ appeared to show paradoxical results for WBC: the likelihood of disease was lower in children with a WBC $\geq 15,000/\text{mm}^3$.

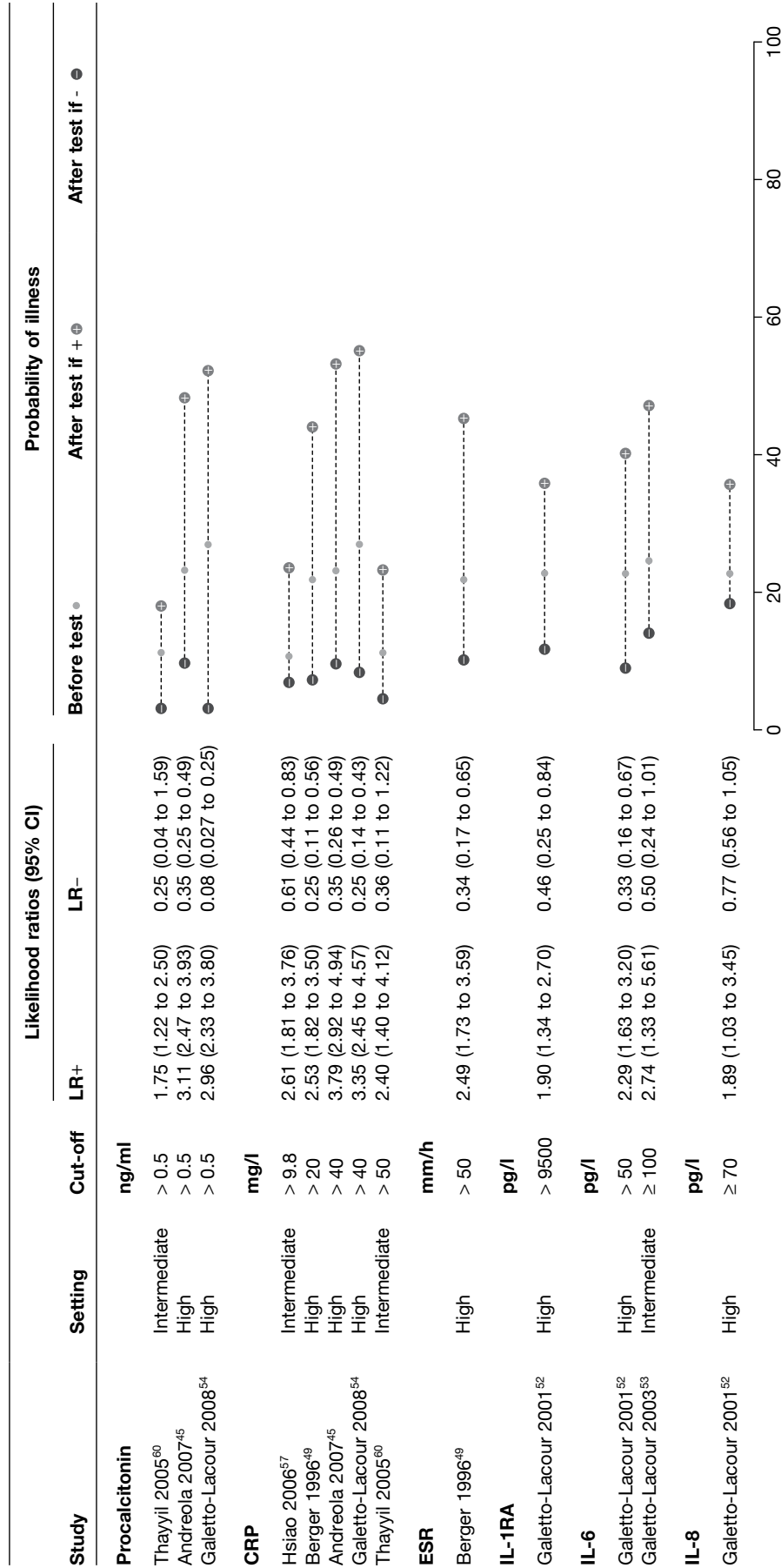


FIGURE 8 Diagnostic value of inflammatory markers for serious infection. IL-1RA, interleukin-1 receptor antagonist; IL-6, interleukin-6; IL-8, interleukin-8. Setting: low, low prevalence of serious illness (<5%); intermediate, intermediate prevalence of serious illness (5–20%); high, high prevalence of serious illness (>20%).

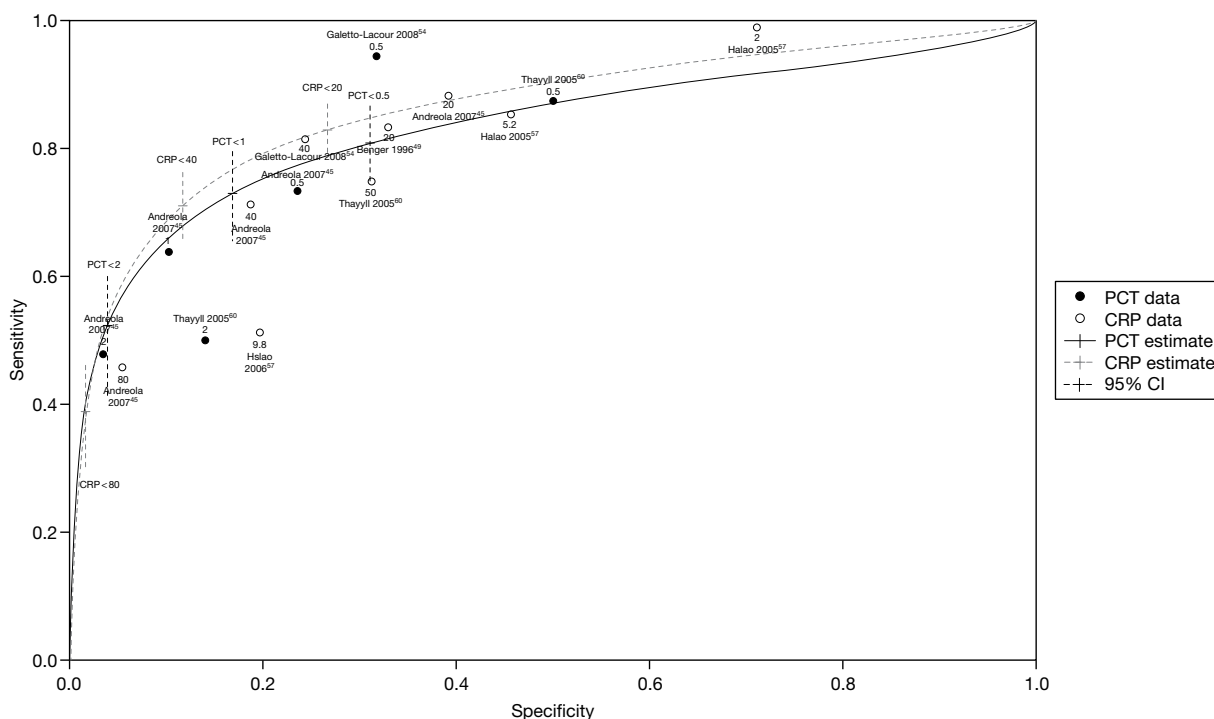


FIGURE 9 Summary receiver operator characteristic curves for CRP and PCT.

Added value of combinations of blood tests and clinical features

We identified eight studies^{46,47,50–55,60} which reported the diagnostic value of prediction rules consisting of a single blood test or combinations of blood tests, with or without clinical features (Figure 11). The clinical features, laboratory tests and cut-offs used are provided in Appendix 3.

Bleeker *et al.*⁵⁰ determined the added value of laboratory markers in patients testing high or low risk for a prediction rule of clinical symptoms. The clinical rule included duration of fever, history of vomiting, ill appearance, chest-wall retractions and/or tachypnoea, and decreased capillary refill. Their results showed that in children testing high risk on their clinical prediction rule, a combination of WBC, CRP and urinalysis lowered the probability from 42% to 15% when negative, but the probability was not increased substantially when positive (54%), indicating a greater ability to rule out than rule in serious infection. In contrast, the use of WBC, CRP and urine analysis in children who tested low risk on the clinical prediction rule lowered the probability of serious infection (from 12% to 4%) if negative and increased the probability to 31% if positive.

The laboratory clinical prediction rule reported by Thayyil *et al.*⁶⁰ provided the highest LR+ (10.67) based on the results of PCT, CRP and WBC. However, this offered little rule-out value with a LR– of 0.52. The best performing prediction rule by Galetto-Lacour *et al.*⁵⁴ reported a LR+ of 5 and a LR– of 0.07 and involved urine dipstick as well as measuring a score based on the levels of PCT and CRP. This rule had good rule-in and rule-out potential, a high score increasing the likelihood of a serious infection from 27% to 64%, and a below threshold score decreasing it to 2%. The inclusion criteria for both the studies of Thayyil *et al.*⁶⁰ and Galetto-Lacour *et al.*⁵⁴ included the presence of fever.

We also found that combinations of inflammatory markers offered little additional diagnostic value to the individual tests, with all the CIs of the combinations largely overlapping with those

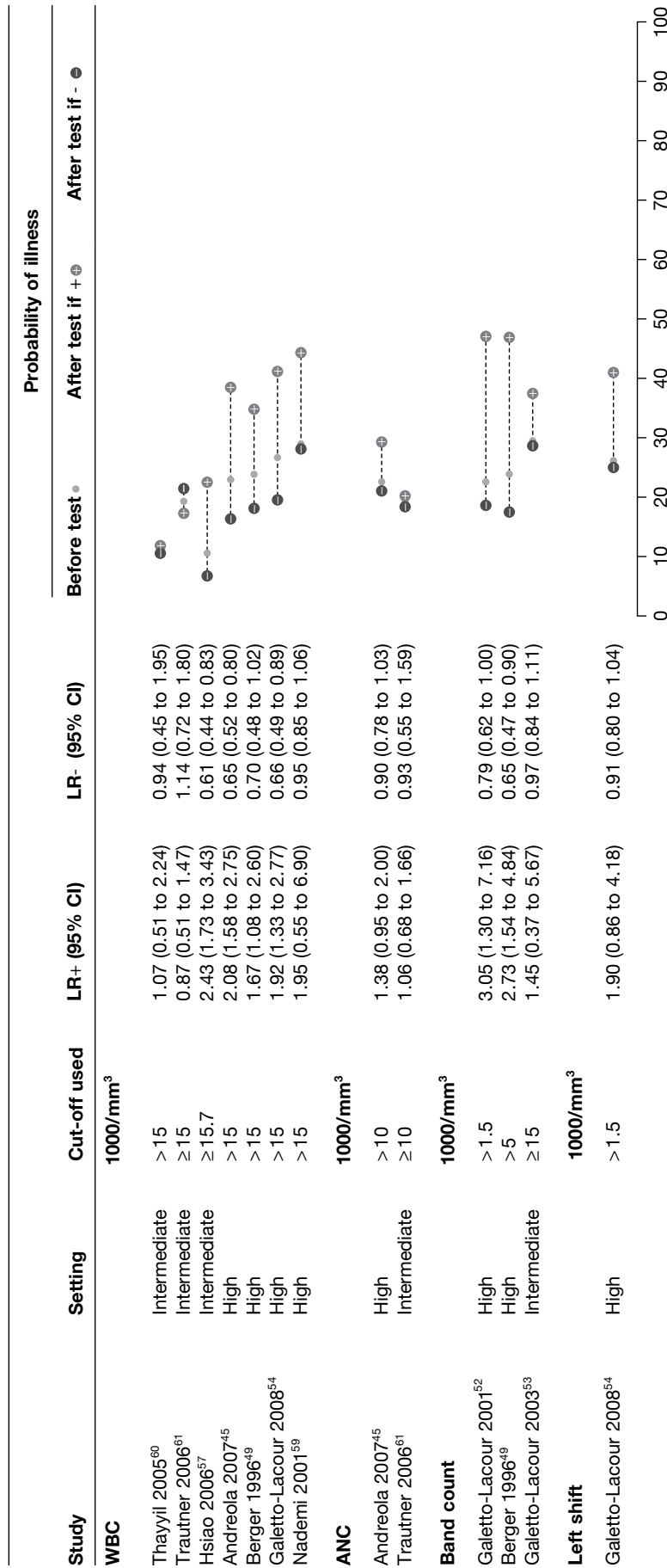


FIGURE 10 Diagnostic value of WBCs for serious infection. ANC, absolute neutrophil count. Setting: low, low prevalence of serious illness (<5%); intermediate, intermediate prevalence of serious illness (5–20%); high, high prevalence of serious illness (>20%).

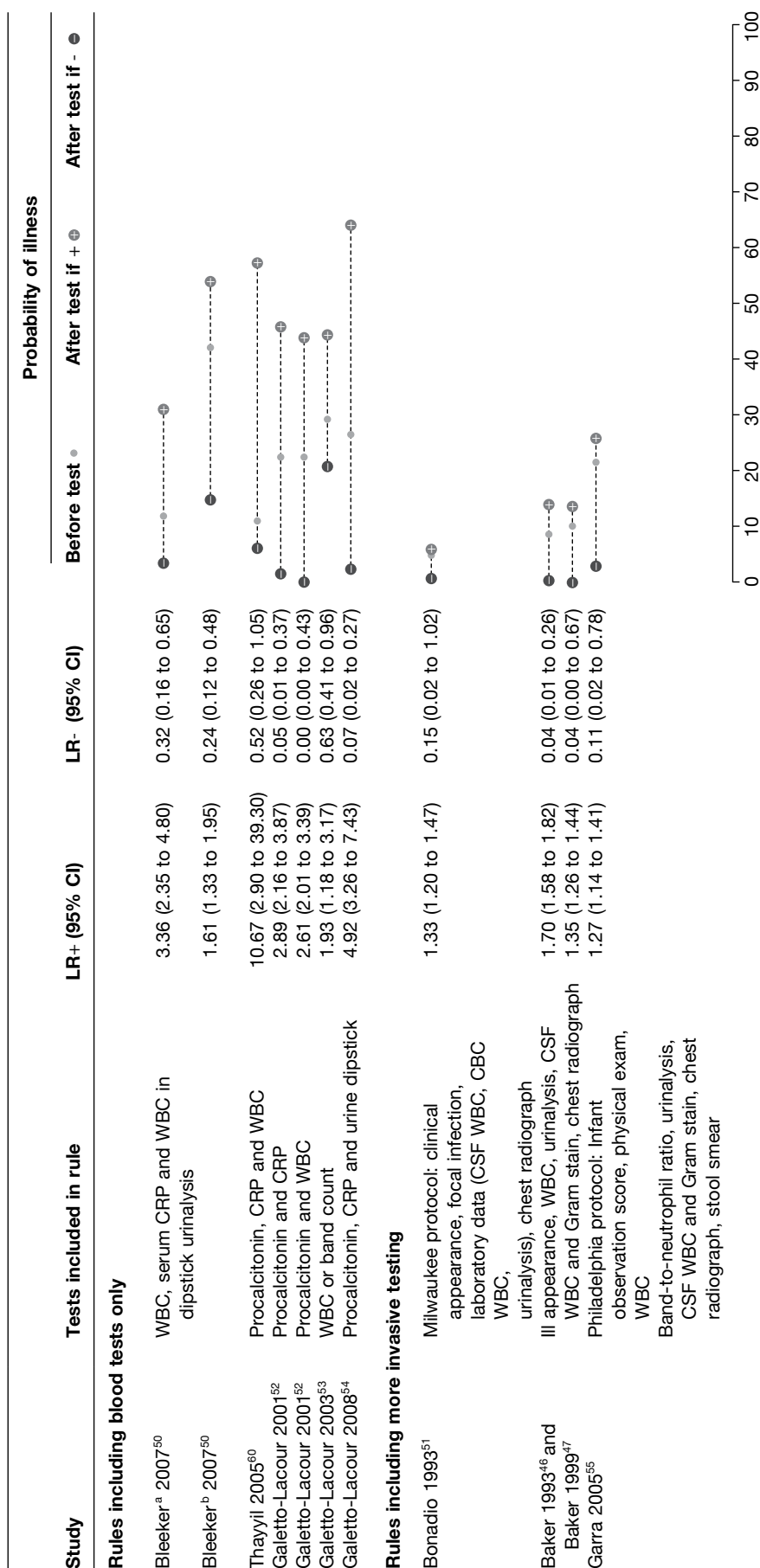


FIGURE 11 Prediction rules combining CRP, WBC and PCT with clinical features. a, In patients negative on a clinical prediction rule; b, in patients positive on a clinical prediction rule. ANC, absolute neutrophil count. Full details of cut-off points applied are given in Annex 3 (Van Den Bruel *et al.* Diagnostic value of laboratory tests in identifying serious infections in febrile children: a systematic review. *BMJ* 2011;**342**:D3082).

of the individual tests (Figure 12). In general, combinations in which all tests had to be positive increased the ability to rule in serious infection, whereas combinations in which at least one test had to be positive increased the ability to rule out serious infection.

Although clinical prediction rules which involve a series of clinical features as well as more invasive investigations (i.e. cerebrospinal fluid parameters and chest radiography) are not within the scope of this review, the results of the clinical prediction rules identified are presented in Appendix 3. These rules were derived for use in infants < 3 months of age presenting to the ED with fever. The dumbbell plots show they provide little diagnostic value in ruling in serious infection (LR+ 1.27–1.70), but provide LR– ranging from 0.04 to 0.15 (pooled LR– 0.06; 95% CI 0.018 to 0.19), indicating good rule-out value.

Diagnostic value of laboratory tests for meningitis and/or bacteraemia

Only one study evaluating CRP for the diagnosis of meningitis and/or bacteraemia was available that showed CRP is able to rule out meningococcal infection if normal (LR– 0.05).⁷⁶ Two studies on meningococcal infection^{75,76} evaluated the value of coagulation markers, i.e. activated partial thromboplastin time (aPTT), international normalised ratio (INR) and platelets, and found good diagnostic value at ruling in the disease (LR+ 2.05–13.08), except for platelets (LR+ 3.20). Normal coagulation markers do not exclude the possibility of meningococcal infection.

Six studies evaluated the value of WBC, all with bacteraemia as outcome.^{62–64,66,67,76} None of the studies reported a LR+ > 3.0 or a LR– < 0.3, suggesting that WBCs were of very little use in diagnosing or excluding bacteraemia. Indeed, after pooling these studies, the summary sensitivity of abnormal WBC was 62.71% (95% CI 52.60% to 71.81%), summary specificity was 69.27% (95% CI 62.71% to 75.13%), summary was LR+ 2.04 (95% CI 1.51 to 2.75) and summary LR– was 0.54 (95% CI 0.40 to 0.73). Absolute neutrophil count was reported in two studies, one on bacteraemia⁶² and one on meningococcal infection.⁷⁶ Both reported low LR (LR+ 1.65–2.57 and LR– 0.40–0.60).

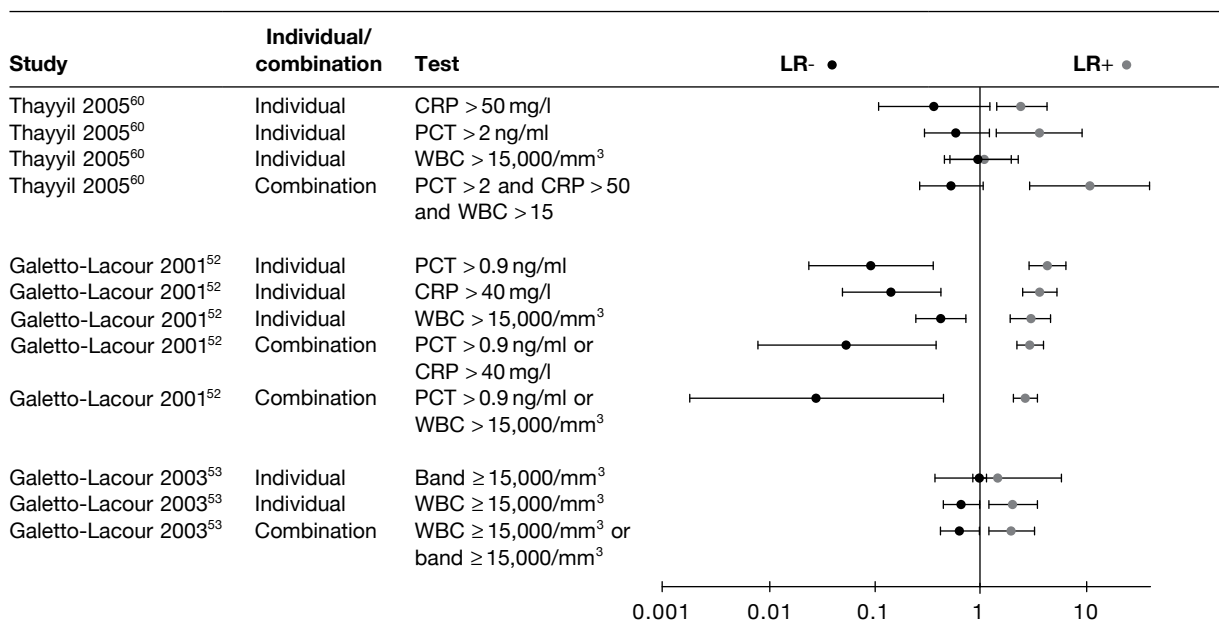


FIGURE 12 Likelihood ratios of individual and combinations of inflammatory markers and WBCs.

Five studies reported prediction rules for meningitis and bacteraemia which combined clinical features with inflammatory markers.^{62,66,67,72,75} The rules by Teele *et al.*,⁶⁶ Crocker *et al.*⁶² and Waskerwitz⁶⁷ did not increase or decrease the probability of bacteraemia significantly (LR+ 1.21–2.92). Two prediction rules, those of Nielsen *et al.*⁷⁵ and Oostenbrink *et al.*,⁷² combined clinical features with CRP, and both were able to rule out meningococcal disease and meningitis respectively (LR– 0.01–0.05). The Nielsen rule also increased the likelihood of meningococcal infection at higher cut-offs (LR+ 17.33). The Nielsen rule consists of the presence and characteristics of haemorrhages, general condition nuchal rigidity, band count and CRP. The Oostenbrink rule consists of duration of the main complaint, vomiting, meningeal irritation, cyanosis, petechiae, disturbed consciousness and CRP.

Chapter 7

Methods used for validation of prediction rules

We aimed to perform a multiple external validation analysis of the clinical prediction rules identified from the systematic review using existing data sets on populations of children attending ED or primary care. The details of the clinical prediction rules which were validated and the variables included in each prediction rule have been presented in previous chapters. This chapter describes the methods of the validation of the prediction rules.

Identification of data sets

We identified existing data on children attending ED, urgent-access primary care or primary care from two sources. Firstly, we identified data sets included in the systematic review. We approached authors of data sets published within the last 10 years and invited them to contribute data and expertise to the multiple external validation study. Secondly, we contacted content experts in this clinical area and asked them to identify other existing data sets. The criteria we used to include data sets were modified from those used as inclusion criteria for the systematic review (*Table 5*).

Identification of clinical prediction rules

The systematic review identified clinical prediction rules which have been applied to a composite outcome of serious infection, pneumonia, meningitis and dehydration from gastroenteritis

TABLE 5 Criteria for inclusion of data sets in validation data sets

Characteristic	Inclusion	Exclusion
Design	Studies which had recorded clinical features, laboratory tests Prospective or retrospective cohort study design	Unclear methods
Participants	Age between 1 month and 18 years Studies including children spanning this age range included if they reported age (or age could be calculated)	Children with pre-existing immune suppression (such as HIV infection or neutropenia due to chemotherapy) Outwith age range
Setting	Ambulatory care (defined as general or family practice, paediatric outpatient clinics, paediatric assessment units or EDs) Developed countries, defined using the United Nations list, which included Europe, Canada, USA, Australia, New Zealand and Japan	Studies conducted in developing countries
Outcome	Serious infection, defined as sepsis (including bacteraemia), meningitis, pneumonia, osteomyelitis, cellulitis, gastroenteritis with severe dehydration, complicated urinary tract infection (positive urine culture and systemic effects such as fever) and viral respiratory tract infections complicated by hypoxia (e.g. bronchiolitis)	Diagnosis other than serious infection
Data availability	Agreement to share data	

HIV, human immunodeficiency virus.

(see *Appendix 2*): the YOS, a five-stage decision tree, two prediction rules for pneumonia, two prediction rules for meningitis and a rule for dehydration from gastroenteritis.

To determine which data sets could be used to validate the prediction rules, we constructed a matrix of all variables recorded in the available data sets. Consensus through group discussion was reached on which clinical prediction rules we could validate based on the available data sets.

Exploring heterogeneity

Given that the analysis was performed on an individual patient data level, it was essential to tabulate heterogeneity and mark up similarities between studies. This heterogeneity may be related to patient characteristics, setting, geographical characteristics, test characteristics or study design.

Characteristics of included data sets

Key characteristics of each of the data sets were extracted, including inclusion and exclusion criteria, setting, age of participants and individual variables (see *Table 6*). The variables used in each data set were translated to English if necessary, and clarifications for variables were obtained from the researcher responsible for each data set. Coding of all key variables was checked not only on statistical grounds but also on clinical grounds, with the responsible researcher (i.e. meaning of the variable label 'unknown' in relation to missing observations). A merged data set was then created and checked for internal and external validity. The translation, the synopsis, the recoding and the data checking were carried out by one member of the study team (JV) and the results of each step were discussed with every responsible researcher when necessary.

Where variables were not entirely comparable, proxies were identified from other variables contained in the data sets and fitted to perform the analysis. If too few variables were present in one data set to allow cross-validation, that data set was excluded from the specific analysis (e.g. fewer than three variables present for a composite clinical prediction rule of six variables was deemed unsuitable). In addition to approximations on the variables, approximations to sum scores had to be used as well. For example, the YOS assigns points at three levels for each variable, depending on the severity of abnormality of that variable. For example, a normal reaction to parent stimulation is given one point, a moderately abnormal reaction is given three points and a severely abnormal reaction given five points. Some data sets had only dichotomous results for the variables included in the score, for which one point was awarded to a normal variable and five to an abnormal variable. Consequentially, the total sum score differs from the original sum score, and different cut-offs for defining an abnormal YOS were explored (i.e. cut-offs of 10 and 8).

Statistical analysis

The accuracy of the clinical prediction rules identified in the systematic review was assessed in each of the data sets in which this was possible. The outcome used to validate the prediction rule was as similar as possible to that of the original prediction rule derivation. For each prediction rule, we used dumbbell plots to display results of sensitivity, specificity, LR+, LR- and pre- and post-test probabilities. This analysis was performed for all children in each data set, as well as the subgroup analyses of only children who had been referred from primary care to the ED, and for children who had not been referred. When a data set had been used to derive a prediction rule, we did not use this data set to validate that same prediction rule. Because there were considerable

missing data on key variables (see *Appendix 6*), several specialised methods (i.e. multiple imputation techniques by the Markov chain Monte Carlo method or by chained equation) were evaluated but none was feasible, because the condition of observations assumed to be missing at random was almost certainly violated, and also because in some cases the problem was that variables were missing completely. All analyses were performed with Stata (version 11).

Chapter 8

Results of external validation of clinical prediction rules

Description of included data sets

We identified seven data sets (*Table 6*) for the multiple external validation.^{11,28,29,60,72,79} Two had been collected in the UK,^{29,79} four were from the Netherlands and one was from Belgium,¹¹ providing data on a total of 11,045 children. Two data sets had been collected from primary care,¹¹ the remainder from ED settings. Four of the data sets were included in the systematic review (Van den Bruel *et al.*,¹¹ Thompson *et al.*,²⁹ Bleeker *et al.*⁵⁰ and Oostenbrink *et al.*⁷²). Inclusion criteria for the data sets were children with fever (three), acute illness (two), or acute infection (one) and referred children with meningeal signs (one). Six of the data sets excluded children with various types of comorbidity, and in one case children who required immediate resuscitation. The mean age of children ranged from 0.94 to 5.0 years, and the prevalence of serious infection ranged from 0.78% to 44.71%.

Clinical predictor variables included in data sets

The variables recorded in each data set varied (see *Appendix 4*). Most data sets included basic demographic characteristics such as age, duration and severity of illness, or referral status. Presenting symptoms were recorded in almost all data sets, apart from two gastrointestinal symptoms. Temperature was recorded in all data sets, followed by heart rate (five data sets), respiratory rate (five data sets), capillary refill time (five data sets) and oxygen saturations (four data sets). Investigations in all data sets were performed at the discretion of the clinical team and none had performed a complete set of investigations on all children.

A matrix was created to determine which of the data sets could be used to validate the clinical prediction rules. Consensus was reached for four clinical prediction rules: YOS, a five-stage decision tree, a pneumonia rule and a meningitis rule. *Appendix 5* presents the variables used for external validation of the clinical prediction rules, either original ones or approximations of the predictor variables.

For the YOS, three data sets had variables which were identical to the original Yale scoring (Berger *et al.*, Brent *et al.*⁷⁹ and Thompson *et al.*²⁹). Two data sets (Van den Bruel *et al.*¹¹ and Oostenbrink *et al.*⁷²) required recoding of dichotomous variables to the YOS, in which cases we assigned a score of 1 for a normal value and 5 for an abnormal value. We considered using the intermediate value of 3 instead of 5 to indicate a normal value, but this did not provide any differences in distribution of the outcome variables. Different cut-offs (8 and 10) were used to validate the YOS, which was consistent with the cut-offs identified by the systematic review. Four data sets were available to validate the YOS using a cut-off of 10 (Van den Bruel *et al.*,¹¹ Berger *et al.*, Thompson *et al.*²⁹ and Brent *et al.*⁷⁹), and an additional data set when using a cut-off of 8 (Oostenbrink *et al.*⁷²).

TABLE 6 Characteristics of data sets used for external validation of prediction rules

Data set	Year	Setting	Country	n	Prevalence serious infection, % (95% CI)	Mean age, years (range)	Inclusion criteria	Exclusion criteria
Van den Bruel <i>et al.</i> ¹¹	2004	GP/AP/ED	BE	4102	0.78 (0.51 to 1.05)	5.00 (0.02–16.93)	Children ≤ 16 years with acute illness for a maximum of 5 days	Traumatic or neurological illness, intoxication, psychiatric or behavioural problems without somatic cause or an exacerbation of a chronic condition. No repeated inclusion of same infant within 5 days. Exclusion of physicians if the assumption of consecutive inclusion was probably violated
Oostenbrink <i>et al.</i> ⁷²	2001	ED	NL	593	44.35 (40.34 to 48.36)	3.66 (0.08–16.08)	Children aged 1 month to 16 years, meningeal signs at GP, paediatrician or self-referred with neck pain	Comorbidity, VP-drain
Roukema <i>et al.</i> ²⁸	2005	ED	NL	1772	12.97 (11.40 to 14.55)	2.88 (0.09–15.69)	All children with fever (T > 38 °C) at ED, without meningeal irritation	Chronic disease, immunodeficiency
Bleeker <i>et al.</i> ⁵⁰	2006	ED	NL	595	23.53 (20.11 to 26.95)	0.94 (0.00–2.99)	Children with fever T > 38 °C at ED, no clear focus identified after evaluation GP of history by paediatrician	Chronic disease, immunodeficiency
Thompson <i>et al.</i> ²⁹	2005	PAU	UK	700	44.71 (41.02 to 48.41)	4.62 (0.00–16.00)	Children age 3 months to 16 years with suspected acute infection	Children with diseases liable to cause repeated serious bacterial infection, and infections resulting from penetrating trauma
Brent <i>et al.</i> ⁷⁹	2001	ED	UK	2777	13.43 (12.16 to 14.70)	3.32 (0.18–18.39)	All children presenting with a medical problem to the paediatric emergency care unit whatever their age	Children who required immediate resuscitation. Comorbidity and chronic illness
Berger <i>et al.</i>	2005	GP	NL	506	6.52 (4.36 to 8.68)	2.16 (0.27–5.88)	Children aged 3 months to 6 years, contacting a GP co-operative after hours with fever as the presenting complaint	Language barriers, no repeated inclusion within the last 2 weeks

AP, ambulatory paediatric care; BE, Belgium; GP, general practice; NL, Netherlands; PAU, paediatric assessment unit; T, temperature; VP, ventriculo-peritoneal.

One of the data sets had been used to derive the five-stage decision tree (Van den Bruel *et al.*¹¹), and an additional six data sets were suitable for validation (Berger *et al.*, Roukema *et al.*,²⁸ Brent *et al.*,⁷⁹ Bleeker *et al.*,⁵⁰ Thompson *et al.*²⁹ and Oostenbrink *et al.*⁷²). One variable in the derivation study ('physician's gut feeling that something is wrong') was not recorded in any of the validation data sets. We, therefore, used a proxy variable of 'clinical impression' in the validation data sets. In addition, we compared the diagnostic characteristics of the rule in the derivation data set using both the original and the proxy variable. Four data sets had all variables present (with approximations, such as nasal flaring for dyspnoea); two data sets (Bleeker *et al.*⁵⁰ and Oostenbrink *et al.*⁷²) had up to two missing variables (diarrhoea and/or dyspnoea).

One data set was used to derive the pneumonia rule (Van den Bruel *et al.*¹¹) and five were available to validate it (Berger *et al.*, Oostenbrink *et al.*,⁷² Roukema *et al.*,²⁸ Thompson *et al.*²⁹ and Brent *et al.*⁷⁹). As with the five-stage decision tree, this rule used clinical impression in the validation data sets as a proxy for the physician's gut feeling that something is wrong. The same variables and approximations for clinical impression and dyspnoea were used as for the five-stage decision tree.

Four data sets were available to validate the meningitis rule (Van den Bruel *et al.*,¹¹ Thompson *et al.*,²⁹ Brent *et al.*⁷⁹ and Oostenbrink *et al.*⁷²). The variable petechiae was available in all data sets, nuchal rigidity in one data set (Oostenbrink *et al.*⁷²) with the proxies meningeal irritation, neck stiffness and AVPU (alert, responds to voice, pain, unresponsive) score of 3 in three of the data sets, whereas coma was approximated with unconsciousness and AVPU score of 4 in all four data sets.

Results of external validation

The results of the validation of the four clinical prediction rules for all children, those who were not referred and those who were referred are provided in *Figures 13–15*, respectively.

A normal YOS score, i.e. below a threshold of 8 or 10, provided little or no rule-out value, i.e. none of the LR_s– were smaller than 0.46 (*Figure 13*). In one low-prevalence setting,¹¹ the score was moderately to highly specific with a LR_s+ of 3.81–7.49, depending on the score cut-off selected. In two studies (intermediate and high prevalence) it provided a marginally useful rule-in value (LR_s+ 3.35–3.63).^{72,79} Subgroup analysis of the children who had not been referred (*Figure 14*) indicated similar results, again with a marginally useful LR_s+ (3.16–3.81) depending on the cut-off selected, and again with no useful LR_s–. In children who had been referred (*Figure 15*), only two data sets were available to validate results, and again indicated a marginally useful LR_s+ in one study (LR_s+ 3.42),⁷⁹ and no useful LR_s–.

We first compared the performance of the five-stage decision tree in the data set in which it had been derived (Van den Bruel *et al.*¹¹), substituting one variable used to derive this rule (physician's gut feeling that something is wrong) with a proxy variable (clinical impression). Using this proxy variable, both the LR_s+ and LR_s– were slightly lower, although the CIs of the latter were overlapping. Using the five-stage decision tree with the proxy variable of clinical impression, we found that the rule had no rule-in value in any of the validation data sets, regardless of whether all children, referred children or non-referred children were considered. However, we found a marginally useful rule-out value in four data sets (LR_s– 0.13–0.35),^{28,50,79} but with wide CIs. When the rule was validated on the subgroups of children who had not been referred (see *Figure 14*) and had been referred (see *Figure 15*), results were similar: marginally useful LR_s– of 0.20–0.24 in non-referred children in two data sets,²⁸ and LR_s– of 0.35 in referred children in one data set.⁵⁰

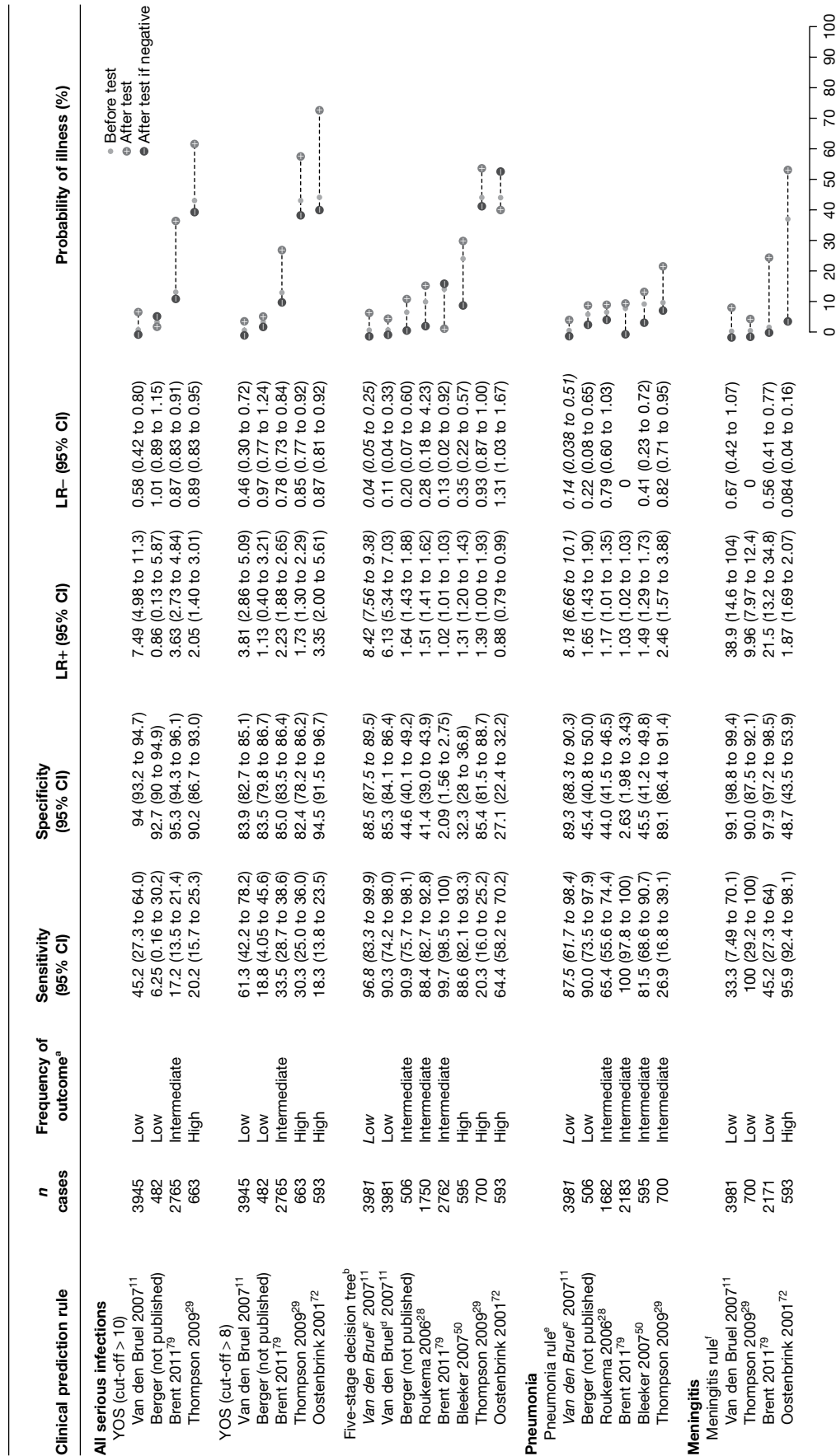


FIGURE 13 Results of multiple external validation of clinical prediction rules with potential to rule out serious infection in all children. a, Setting: low frequency (<5%), intermediate frequency (5–20%) and high frequency (>20%) of the selected outcome measure for serious infection in all children. b, If yes to any of the five sequential questions: (1) clinical instinct that something is wrong, (2) dyspnoea, (3) temperature >39.5°C, (4) diarrhoea and (5) age 15–29 months. c, Italic denotes a derivation study. d, 'clinical instinct that something is wrong' replaced by 'clinical impression'. e, If yes to any of (1) shortness of breath and (2) clinicians concern. f, If yes to any of (1) petechiae, (2) nuchal rigidity and (3) coma.

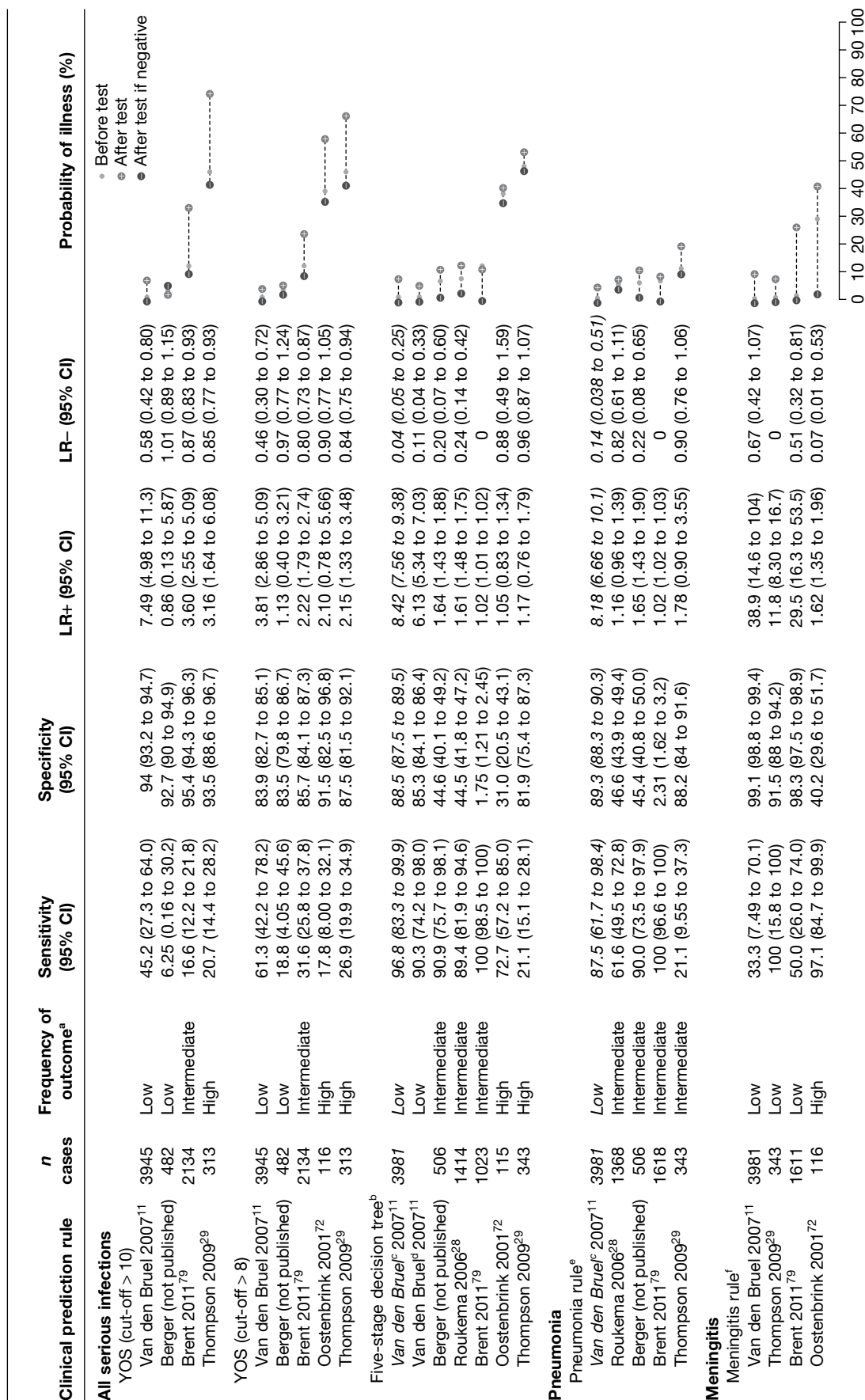


FIGURE 14 Results of multiple external validation of clinical prediction rules in non-referred children. a, Setting: low frequency (<5%), intermediate frequency (5–20%) and high frequency (>20%) of the selected outcome measure for serious infection in all children. b, If yes to any of the five sequential questions: (1) clinical instinct that something is wrong, (2) dyspnoea, (3) temperature >39.5°C, (4) diarrhoea and (5) age 15–29 months. c, *Italic* denotes a derivation study. d, 'clinical instinct that something is wrong' replaced by 'clinical impression'. e, if yes to any of (1) shortness of breath and (2) clinicians concern. f, if yes to any of (1) petechiae, (2) nuchal rigidity and (3) coma.

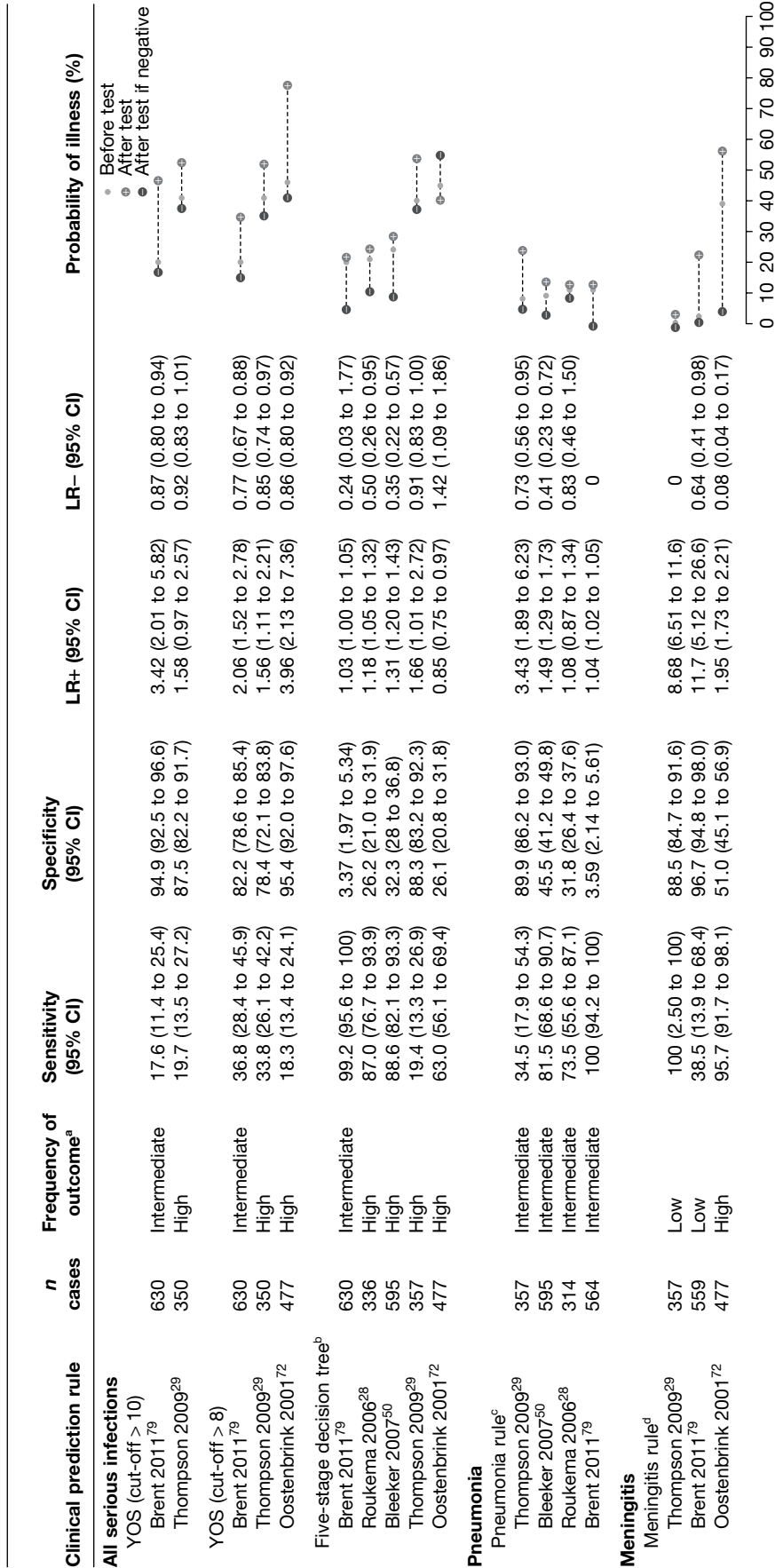


FIGURE 15 Results of multiple external validation of clinical prediction rules in referred children. a, Setting: low frequency (<5%), intermediate frequency (5–20%) and high frequency (>20%) of the selected outcome measure for serious infection in all children. b, If yes to any of the five sequential questions: (1) clinical instinct that something is wrong, (2) dyspnoea, (3) temperature > 39.5 °C, (4) diarrhoea and (5) age 15–29 months. c, If yes to any of (1) shortness of breath and (2) clinicians concern. d, If yes to any of (1) petechiae, (2) nuchal rigidity and (3) coma.

None of the five data sets^{28,29,50,79} used to validate the pneumonia rule demonstrated clinically useful LR+, ranging from 1.03 (95% CI 1.02 to 1.03) to 2.46 (95% CI 1.57 to 3.88); but in one data set (Berger *et al.*) the LR- was 0.22, but with a wide CI (95% CI 0.08 to 0.65) (see *Figure 13*). However, when only referred children were considered (see *Figure 15*), we found a marginally useful LR+ of 3.43 (95% CI 1.89 to 6.23) in one data set (Thompson *et al.*²⁹). Subgroup analysis of non-referred children (see *Figure 14*) did not demonstrate useful LRs+ or LRs-.

Finally, validation of the meningitis rule demonstrated highly clinically useful LR+ in three data sets from low- intermediate- and high-prevalence settings, with LRs+ ranging from 9.96 (95% CI 7.97 to 12.4) to 38.9 (95% CI 14.6 to 104), albeit with wide CIs (see *Figure 13*).^{11,29,79} However, in a fourth data set (high prevalence), the LR+ was only 1.87 (95% CI 1.69 to 2.07).⁷² These results did not change substantially when referred or non-referred subgroups were analysed (see *Figures 14* and *15*). The meningitis rule did provide a useful LR- in one data set from a high-prevalence setting (Oostenbrink *et al.*⁷²), with a LR- of 0.084 (95% CI 0.04 to 0.16). However, clinically useful LRs- were not found in any of the other data sets, regardless of referral status (see *Figures 14* and *15*).

Chapter 9

Discussion

Studies identified by systematic review

This systematic review of the existing literature on the diagnostic value of clinical features and laboratory tests in children presenting to primary care or EDs included a total of 36 studies. The majority of the studies were performed in the USA or Canada (18), with a further five from the UK, 11 from other European countries, and one from Australia. All but one of these studies were based in ED settings – we identified only a single study from primary care, performed in general practice in Belgium. As a result, the generalisability of our findings for primary-care settings in the UK may be limited, but we believe that the findings are generalisable to ED settings. The quality of the included studies was modest. However, the clinical context of this systematic review limited the likelihood that studies could meet ideal diagnostic study criteria. For example, verification of disease status for serious infections in all children would ideally involve an extensive diagnostic work-up in all children presenting with possible infection, which is neither feasible nor ethical. Blind interpretation of the reference standard from the index test was found in only four studies, which again reflects the fact that many clinical features are also used in the definition of the outcome.

Prevalence of serious infection varied widely in the included studies, from 0.78% in the primary-care study¹¹ to 55.3% in one study²⁹ from a paediatric assessment unit. The implication of this is that the pre- and post-test probability of the same clinical features and predictors may vary markedly between different studies. We therefore used various methods to reflect this when summarising the results. In particular, we developed ‘dumbbell’ plots, which provided a novel way of presenting the sensitivity, specificity, LR+, LR-, prevalence, and pre- and post-test probabilities in a simple visual format.

Serious infection is not a discrete clinical condition; the majority (20/35) of studies used a composite outcome of serious infection which included sepsis, bacteraemia, meningitis, pneumonia and urinary tract infection (and in some cases additional infections such as cellulitis, osteomyelitis and abscess). Other studies used outcomes of single clinical entities such as meningitis, gastroenteritis causing dehydration, pneumonia and meningococcal infection. In the majority of studies, children were included based on the presence of fever and, in some cases, clinical features. For example, the two studies which used meningococcal infection as an outcome included only children with haemorrhagic or non-blanching rash.

Predictive value of clinical features for serious infection

We identified a total of 30 studies^{11,29,36,45,48–50,52,53,56–76} in the systematic review which provided the diagnostic value of clinical features for serious infections as a composite outcome or for a specific serious infection. We grouped clinical features into those related to overall or global assessment, circulatory and respiratory features, meningeal irritation, temperature and common clinical symptoms. Overall and global assessment of severity of illness were useful diagnostic features for ruling in serious infection in the single study from a low-prevalence setting.¹¹ The diagnostic value of these overall or behavioural features was not consistently found in the studies

from intermediate- or high-prevalence settings. The definition of overall assessment varied, but included parental concern that the illness was different from previous illnesses, clinician's gut feeling that something was wrong, as well as changes in the child's behaviour such as a change in crying pattern, drowsiness, moaning or inconsolability. Lack of consistent definition of the components that constitute overall or global assessment limits the value of this feature for clinical practice. It is likely that it includes multiple components of the clinical assessment which are internalised (by parent or clinician) in the course of parenting or clinical training. GPs' gut feeling has also been described as the most discriminating test result in patients presenting with chest pain.⁸⁰ The usefulness of overall assessment is consistent with the results of a survey of GPs in the UK, which highlighted overall assessment as the most important part of the clinical assessment of pre-school children.¹⁵ In that study, GPs reported that playing, smiling and moving around the room were the most important aspects of their overall assessment. It is unclear why overall assessment would be less valuable in higher prevalence settings; one reason could be the referral paradox, whereby the presence of children who have an abnormal overall assessment of serious illness is inflated in settings where more children are referred because of the presence of these features. Additional reasons may be the immediate availability of a battery of technological tests and the fact that initial assessment of unwell children in these settings tends to be performed by junior clinicians. Finally, we do not know the extent to which these features can be learned or taught.

The most useful circulatory predictors of serious infection were the presence of cyanosis and poor peripheral circulation. However, objective definitions of these were lacking, and clinical features that are commonly used to assess poor circulation (e.g. capillary refill time and pale skin colour) did not provide useful LRs in three studies of low and high prevalence.^{11,49,50} Rapid breathing appeared to provide useful LR+ only in the low-prevalence study,¹¹ it was weaker in the four other studies that assessed rapid breathing.^{29,50,73,74} The presence of meningeal irritation, petechial rash, decreased consciousness and seizures was consistently identified as a red flag for bacterial meningitis across all settings. However, we know from other studies that these features occur late in the pre-hospital course of bacterial meningitis, and thus may have red flag value only later in the course of this illness.⁹

The presence of fever provided diagnostic value in settings where the prevalence of serious infection was low or intermediate. Using cut-offs of both 40°C and 38.5–38.9°C, three studies showed substantial increased post-test probability, suggesting value as a rule-in or red flag feature. However, in the five studies with higher prevalence, temperature provided no rule-in ability, i.e. the post-test probability of serious infection was similar to the pre-test probability irrespective of whether it was above or below the cut-off threshold. Not surprisingly, symptoms that are common in children, such as cough, headache, tummy ache, vomiting, diarrhoea, poor feeding and coryza, had little diagnostic value. The NICE traffic light system and the Manchester Triage System score were of little value in a single study from a high-prevalence setting.²⁹ The authors are currently collecting data from other high-prevalence settings to validate these prediction rules.

Predictive value of laboratory tests rules for serious infection

The systematic review identified 14 studies^{45–47,49–55,57,59–61} which reported the diagnostic value of laboratory tests for the diagnosis of serious infections. However none of the studies was performed in primary-care settings. CRP is used far more widely in hospital clinical practice and EDs in the UK than PCT, reflecting the fact that this is an older test, performed in highly automated analysers. PCT is not routinely used in the UK in children, as it is currently more costly and more difficult to perform in many laboratories. Our findings indicate that both CRP

and PCT provide comparable diagnostic accuracy. Pooling of the five studies^{45,49,54,57,60} which had assessed CRP for a composite outcome of serious infection yielded a LR+ of 3.15 (95% CI 2.67 to 3.71) and a LR- of 0.33 (95% CI 0.22 to 0.49). As all five studies were performed in intermediate- or high-prevalence settings, these moderate LR values provide relatively large changes in post-test probability. Our findings are similar to an earlier review of the value of CRP in children with fever which used slightly different methods, but found LR+ 3.64 and LR- 0.29 compared with LR+ 3.15 and LR- 0.33 in our study.¹⁹ We were not able to pool the results of the three studies that had assessed PCT,^{45,54,60} which demonstrated LR+ ranging from 1.75 to 2.96 and LR- of 0.08 to 0.35. The optimal cut-off point to use for CRP and PCT depends on whether the clinical goal is to rule in or rule out serious infection. Based on our results, cut-off levels of 2 ng/ml for PCT and 80 mg/l for CRP provide good diagnostic value (a LR+ of ≥ 4) to rule-in serious infection. To rule out effectively, cut-off levels of 0.5 ng/ml for PCT or 20 mg/l for CRP may be a better choice (providing a LR- of ≤ 0.2).

White blood cell indicators (total WBC, absolute neutrophil count, band count or left shift) provided substantially lower diagnostic value than CRP or PCT for ruling in serious infection, and no value for ruling out. This also confirms the findings of a previous review which showed that WBCs did not provide additional diagnostic value over and above CRP. The performance of ESR in a single study⁴⁹ (cut-off of 50 mm/hour) provided LR+ 2.49 and LR- 0.34. The two studies^{52,53} investigating the three interleukins demonstrated somewhat inferior diagnostic values to CRP or PCT.

We found no evidence to determine the roles of any of the inflammatory markers at different time points in the course of a serious infection. As PCT levels rise earlier in the course of bacterial infection than CRP, it is possible that PCT would have a comparatively greater diagnostic value in the early course of serious bacterial infections than CRP or WBCs. The use of combinations of inflammatory markers provides some evidence to support this hypothesis. In general, a combination of inflammatory markers in which all tests have to be positive increases the ability to rule in, whereas a combination in which at least one has to be positive increases the ability to rule out. The best performing prediction rule by Galetto-Lacour *et al.*⁵⁴ reported a LR+ of 5 and a LR- of 0.07 and involved urine dipstick as well as measuring a score based on the levels of PCT and CRP in febrile children. This rule had good rule-in and rule-out potential, a high score substantially increasing the likelihood of a serious infection from 27% to 64%, and a below-threshold score decreasing it to 2%. Using composite outcomes of serious infection to assess inflammatory markers may also mask important differences in the rule-in or rule-out values of these markers in different types of serious infection. However, we found no studies which allowed us to assess the value of individual or combinations of inflammatory markers for individual infections in ED or ambulatory settings.

In clinical practice, laboratory tests are usually requested based on the results of a clinical assessment, rather than in isolation. This is implicit in the inclusion criteria for the studies which assessed laboratory tests, given that fever (or other clinical features, or indeed referral) was used as an inclusion criterion. However, only a single study assessed the results of laboratory markers in patients testing positive or negative on a symptoms-only prediction rule.⁵⁰ This study found that in children testing positive on the symptoms-only prediction rule, a combination of WBC, CRP and urinalysis lowered the probability of serious infection from 42% to 15% when negative, but did not raise the probability substantially when positive (54%), indicating a greater ability to rule-out than to rule-in serious infection. In contrast, in children who tested negative on the symptoms-only prediction rule, the use of WBC, CRP and urinalysis moderately lowered the probability of serious infection (12% to 4%) if all were negative and moderately increased the probability to 31% if all were positive.

Predictive value of clinical prediction rules for serious infection

Given that most of the clinical prediction rules identified had limited evidence of validation, we attempted to use existing data sets to perform cross-validation. We identified a total of six existing data sets^{11,28,29,50,72,79} including 11,045 children presenting to primary care or EDs with suspected acute infection. Some of the data sets were used to derive the prediction rules we had identified in the systematic review, and thus could not be used to validate the same rule.

The best performing clinical prediction rule ('five-stage rule') for ruling out serious infection identified in the systematic review was derived in a low-prevalence setting and was based on the presence of five features: clinician's impression something is wrong, dyspnoea, temperature, diarrhoea and age.¹¹ We did not identify any published studies which had validated this rule. Our attempts to validate this rule produced highly variable results; it had no rule-in value in any of the validation data sets, regardless of whether or not children were referred, but had marginally useful rule-out value in four data sets (LR⁻ 0.13–0.35) with wide CIs.^{28,50,79} Although this rule appears to offer some diagnostic value in primary care, the lack of consistency within and between settings limits the adoption of this prediction rule in routine care.

The YOS was developed in the early 1980s, in a tertiary US hospital, and assesses the presence of a series of clinical features in children < 3 years of age grouped as follows: colour, breathing difficulty, response to social overture, hydration, level of activity and state variation. We had sufficient data to calculate a summary sensitivity of 32.5% (95% CI 21.7% to 45.5%) and a summary specificity of 78.9% (95% CI 73.9% to 83.1%). When we attempted to validate the YOS, we found little or no rule-out value, i.e. none of the LR^{s-} was < 0.46. However, in both low- and high-prevalence settings,^{11,72,79} the YOS provided moderate to high specificity and a LR⁺ of 3.81–7.49 depending on the cut-off selected. The YOS is the best known and most studied clinical prediction rule, and the overall findings and heterogeneity were disappointing. The methodology used to derive clinical prediction rules has advanced considerably since the YOS was developed. Indeed, some key methodological limitations of the YOS were lack of weighting for items (all are assigned equal weight), crude methods used to identify and combine clinical predictors, lack of information on interobserver interpretation of items and the scoring system used, as well as limited clinical generalisability as it was developed in US tertiary paediatrics hospital and was limited to children < 3 years of age.

The systematic review identified two prediction rules which had a LR⁻ sufficiently low to rule out pneumonia: both involved the absence of shortness of breath and absence of either parental or a clinician's concern. These were derived in only a single low-prevalence study¹¹ and, therefore despite good LR^{s+} (11.3 and 13.8), were not sufficient on their own to significantly raise post-test probabilities in this setting. We did not identify any published prediction rules for pneumonia which had been derived or validated in intermediate or higher prevalence settings. When we attempted to validate the pneumonia prediction rules in the new data sets, none of them demonstrated clinically useful LR⁺, but in one low-prevalence setting (Berger *et al.*) the LR⁻ was 0.22, suggesting a useful rule-out value.

We identified two prediction rules for bacterial meningitis in the systematic review, both of which had been derived in intermediate settings.^{70,71} One provided rule-out value (no abnormal neurological finding, and the parent had not sought care within 48 hours), whereas the other provided rule-in ability (any of petechia, nuchal rigidity or coma present). Clinically useful LR^{s+} were found when we validated the meningitis rule in three data sets from low-prevalence settings^{11,29,79} (LR⁺ 9.96–38.9), but not in the data set⁷² from a high-prevalence setting (LR 1.87).

In contrast to the derivation studies, we could not confirm useful LR⁻ in the low-prevalence studies, but we did find an extremely small LR⁻ (0.084) from one high-prevalence study,⁷² suggesting rule-out value.

Finally, the systematic review identified a single rule for dehydration from gastroenteritis, which provided a modest LR⁺ and LR⁻ from a single high-prevalence study.⁶⁸ We did not find any published validation studies for this rule, and were not able to validate it in the data sets identified.

Limitations of the systematic review

We searched the literature using standard systematic review methods. However, given that the search filters and strategies for diagnostic systematic reviews are less well developed than those for reviews of interventions, we used several additional methods to attempt to identify all relevant studies, such as searches of relevant guidelines, reference lists and contact with content experts.

The inclusion criteria for our review were explicit. We excluded studies of hospitalised children and those from developing (or transitional) countries. Studies from developing countries have a far higher incidence of serious infection in children, and a higher prevalence of comorbidities than is found in the UK or elsewhere in Europe. In addition, our review focused on serious infection, rather than other illnesses, so we may have missed potentially useful studies which addressed other non-infectious serious illnesses in children. For example, we excluded studies of the Baby Check score, which was devised in the UK in order to help parents and clinicians detect all serious illness in infants, and which has been shown in a small series of 87 children from UK general practice to have a sensitivity of 100% and a specificity 67%.⁸¹ In addition, we eliminated studies which used a case-control design as these are recognised to substantially overestimate diagnostic odds ratios.⁸²

The studies included in the systematic review were highly heterogeneous in terms of setting, prevalence of serious infection, definition of outcome quality, age range and inclusion of clinical predictors. We found only a single study based on primary care, and although this was a large well-conducted study, we are cautious in basing our conclusions for primary care on the results of a single study.

In the analysis and summaries of the predictive value of clinical features, we selected a LR⁺ of 5.0 and a LR⁻ of 0.2 as cut-off values to indicate clinically useful LRs. We recognise that this is arbitrary, and where possible have commented on the LRs that are close to these cut-offs. We also recognise that the effect of a set LR will obviously be dependent on the pre-test probability. Therefore, in presenting the results we include plots of pre- and post-test probabilities, so that the impact of the LR can be assessed.

Limitations of validation of existing clinical prediction rules

The steps which are necessary for multivariable diagnostic research have been described by Toll *et al.*⁸³ Because good calibration (agreement between predicted probabilities and observed outcome frequencies) and good discrimination (ability to distinguish between patients with and without the outcome) in a derivation study are no guarantee of prediction in new patients, it is important to conduct validation studies of clinical prediction rules.^{84,85} Indeed, most prediction rules demonstrate worse performance when validated in new populations.^{84,86} Therefore, additional

aims of this project were not only to identify published validation studies, but also to identify data sets which could be used to validate the clinical prediction rules that we identified in the systematic review.

The YOS was the only clinical prediction rule identified in the systematic review for which validation studies had been published. We therefore attempted to validate this and the other prediction rules that we identified by using existing data sets of children presenting to primary and ED settings with acute infection or illness. These seven data sets provided data on 11,045 children from three European countries presenting to either primary care or EDs with suspected acute infection. However, our ability to use these data sets to validate existing rules was limited for several reasons. In some cases, a data set had been used to derive a prediction rule and therefore could not be used to then validate it. In other cases, we were limited by heterogeneity between the derivation and validation populations, and issues related to external generalisability.

Heterogeneity between the derivation and validation populations may occur when definitions of predictors, outcome variables and measurement methods are different, and especially if measurement or interpretation is subjective. The seven data sets used for external validation required intensive work to translate variables into English, understand the exact meaning of variables (including cultural differences in measuring and interpreting physical signs and symptoms), and assess which clinical predictors were shared across data sets and how outcomes were defined. Even across these data sets, we were limited not only by the number of shared predictors, but also by how they were defined. The results from the multiple external validation, presented in *Chapter 8*, demonstrate the degree of approximation which had to be used to validate the original clinical prediction rule. A further cause of heterogeneity is when case mix differs between derivation and validation populations, especially if the outcome measure is predicted by a variable responsible for a difference in case mix. Case mix tends to be related to the setting in which a study is performed. In both our systematic literature review and our validation study, we included studies performed in a variety of ambulatory settings. Prevalence of serious disease was used as a proxy to account for the effect of setting. Furthermore, clinicians working in a similar setting in different countries may differ in the types of patients who present to them, the proportion of referred or self-referred patients and their role in dealing with emergencies. Differences in case mix can lead not only to different prevalence of serious diseases, but also to different measures of diagnostic accuracy. Sensitivity and specificity of the same test will change when used in different settings⁸⁷ in directions that are not always predictable. To some extent, these changes are influenced by the presence/absence of referral and the reasons why a referral decision has been made. Finally, heterogeneity may have been caused when there were fewer cases in the validation study as a result of random variation, which makes the populations appear different.

Two components for external validity need to be assessed to ensure generalisability: reproducibility and transportability. Reproducibility should be assessed in the derivation study by statistical techniques such as bootstrapping to evaluate the degree of overfitting of a model. This is not part of this report and is extensively described elsewhere. Transportability can be caused by underfitting, which occurs when important independent predictors of the outcome are left out of the prediction rule. Five types of transportability are described in the literature: historical, geographic, methodological, spectrum and interval. In our validation studies, historical transportability was not an issue because all data sets were performed in more or less the same time period. Geographical transportability could have influenced the validation, although all validation studies were conducted in Europe, North America or Australia. Methodological transportability could be an important issue in these validation studies, as it may well be the most common problem when systems fail to generalise. Spectrum transportability was difficult as well, because the spectrum (i.e. being more or less advanced in the disease process) could be

influenced by differences in referral status and settings. As described by Justice *et al.*,⁸⁶ in reality the calibration of most clinical prediction rules will be compromised when these systems are tested in a sample of patients with very different levels of disease severity, as was the case in our validation studies, owing to calibration to the overall outcome prevalence in the development study. Follow-up period transportability probably did not influence our validation results, because there were no variables or outcomes related to longer follow-up periods.

Chapter 10

Conclusions

Overall clinical implications

The 35 studies identified in this systematic review were highly heterogeneous, and we are therefore limited in drawing firm conclusions about our findings. Indeed, our findings illustrate the diagnostic gap between the predictive value achievable by considering the clinical features and the threshold of risk of serious infection.⁷⁸ This gap is currently filled by clinical ‘gut feeling’ and diagnostic safety-netting, which are still not well defined in primary or ED settings.⁸⁸ Moreover, teaching ‘gut feeling’ to trainee clinicians is problematic, particularly given the decrease in opportunities in clinical training to encounter children with serious infections. New video-based interactive educational resources, such as Spotting the Sick Child (www.spottingthesickchild.com), may help improve the ability of clinical staff to recognise children with acute infection.⁸⁶

A single abnormal clinical finding is insufficient on its own to substantially raise or lower the risk of serious infection. We did identify several clinical features that were highly specific for serious infection, and can be considered ‘red flags’:

- parental concern that the illness is different from previous illnesses
- clinician’s gut feeling that something is wrong
- change in the child’s crying pattern
- drowsiness
- moaning or inconsolability
- fever (depending on cut-off selected and type of setting)
- cyanosis
- poor peripheral circulation
- rapid breathing and shortness of breath
- crackles on auscultation
- diminished breath sounds
- meningeal irritation
- petechial rash
- decreased consciousness
- seizures
- loss of consciousness.

When present, these features should prompt a more thorough or repeat assessment of the child. However, as these red flags will occur infrequently (owing to their low sensitivity) even in children with a serious infection, they provide limited value as screening tests. It may be equally useful to identify a set of frequently occurring clinical features, which when all are absent would effectively ‘rule out’ most serious infections – these might be termed ‘green flags’. However, existing evidence is not sufficient to be confident in defining these.

Clinical prediction rules should be ideally suited to the complicated task of identifying children with serious infection. Although we identified two rules for overall serious infection and five for specific conditions, only one (YOS score) had any published validation studies. By using existing

data sets to validate these rules we were able to draw some clinical conclusions without the need to collect new data from children. Firstly, clinical prediction rules offer different diagnostic value, depending particularly (but not only) on the prevalence of serious infection. Secondly, the five-stage decision tree and YOS score provide moderate specificity and offer some rule-in value in primary and ED settings, and one rule for meningitis provided high specificity and rule-in value. We are aware of strategies to adjust prediction rules to other settings, but methodological analysis of this type was outwith the scope of this project.⁸⁹⁻⁹¹

Measuring CRP or PCT is a helpful step in the diagnostic work-up of children in an ED setting where the prevalence of serious disease is relatively high. We did not find a difference in the diagnostic performance of CRP or PCT, but these were superior to WBCs. The LR+ and LR- of CRP and PCT are not very high, confirming the importance of assessing results in the light of clinical findings. Moreover, cut-off values need to be selected depending on whether these will be used as rule-in or rule-out tests, which may vary depending on setting in particular. There is insufficient evidence to confirm the utility of CRP or PCT in primary care and out-of-hours assessment centres.

Research implications

The authors have several recommendations for further research in this area. These recommendations are based on the analysis and findings of the systematic review and external validation of prediction rules carried out for this report.

1. *The need for research in primary-care and low-prevalence emergency department settings*

We identified a major mismatch between the clinical settings where the majority of children with acute infections seek help (i.e. primary care) and the number of studies performed in that setting (a single study). Thus, there is a pressing need for more studies conducted in primary-care or low-prevalence ED settings to identify clinical predictors of serious infection, or for hospital referral/admission. Given the relative infrequency of serious infections, such studies need to include large cohorts of children, and therefore call for national networks or international collaboration.

Research in primary-care and low-prevalence ED settings is needed to determine the diagnostic role of vital signs either alone or in combination as predictors of serious illness in children. Vital signs are currently underused in many clinical settings, which the authors believe is due to difficulties in accurate measurement and interpretation in children.

Given the difficulties in achieving perfect diagnostic accuracy for serious infections in primary-care or ED settings, research is needed on 'safety-netting' strategies in these settings. Safety-netting involves providing information and empowering parents to know when and how to return to seek medical care when their child deteriorates. Although most clinicians use safety-netting, and most parents use a variety of information sources to help inform their care-seeking behaviour, there has been surprisingly little research in to what constitutes effective safety-netting, how it can best be implemented in frontline NHS services and, particularly, the needs of parents and caregivers.

Finally, there are currently no studies based on primary care which evaluate the diagnostic value of inflammatory markers in children. Some markers are now available as point-of-care tests using fingerstick droplets rather than large aliquots of blood, suggesting that this type of study would now be feasible in such settings. The use of such markers, especially during urgent home visits

or urgent access out of hours, largely depends on the devices being user friendly and fast, with limited weight and volume to carry. Currently available devices still need improvements with respect to these characteristics.

2. Understanding the value of repeated testing using inflammatory markers

Given that acute infections are dynamic illnesses, the authors speculate that repeat testing for inflammatory markers as well as other markers of severity such as vital signs could improve their diagnostic value. In addition, it would also be clinically useful to determine whether or not different infections have different inflammatory marker profiles. This type of study would be more likely to be feasible in ED-type settings where children can be evaluated repeatedly over a period of time.

3. Collaborative studies

Research on the prediction of serious infection in children is limited by the low incidence of serious infection and the difficulties in conducting high-quality research in busy clinical settings. As the research conducted for this report demonstrates, collaborations of primary and ED researchers at the national or international level are essential in order to provide sufficient sample sizes and explore generalisability across settings. Given the difficulties and costs involved in conducting primary research on children in busy clinical settings, such collaborations also minimise the burden on children, their parents and clinical teams. Further research in this area should involve appropriate collaborations where possible.

4. Improvements to the design of future studies

While performing the systematic review, we noted a series of problems in the methodology used to conduct the primary studies. In addition, we found very limited evidence for the validation of the clinical prediction rules for serious infection, and none for their potential impact on clinical practice.

We have made the following suggestions to improve the design of future studies in this area:

- (a) We observed substantial heterogeneity in the temperature and age thresholds used to recruit children. Primary studies which use such age and temperature thresholds limit the external validity of study findings because they may not apply to children outwith these thresholds. Indeed, owing to vaccination programmes, the risk of serious bacterial infections in young children now approximates the risk in older ones. Infants and younger children, however, are still at greater risk for complications of infections. We consider that it is more important to include age as a predictor itself, rather than selecting the population within certain age limits. Using certain temperature thresholds at initial consultation as inclusion criteria seems illogical, as it makes assumptions: for example, that children who have lower levels of fever at consultation have low (or no) risk of serious infection, or that children with serious infections respond differently to antipyretics than children with minor infections. We do not believe that these assumptions hold currently true.⁹² To allow generalisability of results to the broader clinical diagnostic dilemma of febrile children, inclusion should be based on parental reported raised temperature⁹³ or a measured temperature above normal values.⁹⁴
- (b) In contrast to some known widely used prediction rules for many acute conditions such as ankle injuries⁹⁵ or head injuries,⁹⁶ the diagnostic outcomes in studies of children presenting with fever are far more heterogeneous. Although a diagnosis is preferably confirmed by a reference standard (culture),²⁵ this is rarely achievable in primary-care and ED settings. As most diagnostic studies in children are performed in acute clinical settings, we may need a pragmatic outcome assessment, such as a composite reference combining multiple test results, following children over a period of time (for example assessing admission rates at initial contact or reattendance in the days after initial contact), or normalisation of

temperature or other symptoms.⁹⁷⁻⁹⁹ For primary care, hospitalisation for > 24 hours may be usable as a primary outcome measure. Specific diagnostic categories can be added as secondary outcomes.

- (c) The definition of 'serious' illness may depend on setting, and future studies should consider carefully the outcomes of interest relevant to settings. In GP settings, identifying children who need referral to secondary care for further assessment or therapy is usually more important than confirming a particular diagnosis. In contrast, in ED and inpatient hospital settings, clinicians evaluate the child for diagnostic aetiology, have more access to diagnostic testing, and place greater emphasis on 'ruling in' diagnostic outcomes in order to start targeted treatment.
- (d) Clinical prediction rules should be ideally suited to difficult clinical situations such as identifying serious infection in children. As we have noted, we found little evidence to support many of the existing prediction rules, but that does not imply that identifying evidence for these or novel prediction rules should be abandoned. Research on prediction rules in this area should consider the following: (1) the prior risk of the outcome, mainly depending on the type of setting, e.g. ED versus primary care; (2) the type of decision to be made, e.g. decision on referral, additional diagnostic tests, treatment or watchful waiting; and (3) the specific diagnosis the prediction rule is targeting. The diagnosis influences which risk estimate for serious bacterial infections is acceptable to start or to withhold additional diagnostics or treatment. For example, missing meningitis increases morbidity and mortality and therefore a very low risk estimate (<0.01%) might be acceptable.^{100,101} For pneumonia, on the other hand, the risk estimate for starting treatment or 'wait and see' is accepted to be higher than for meningitis.

Acknowledgements

University of Oxford: Ruth Davis (Programme Manager, Madox programme); Nigel Coad (University Hospitals Coventry and Warwickshire NHS Trust); Coventry-based research nurses: Diane Mcleod, Roddy MacFaul (Pinderfields Hospital), Nia Roberts (Information Specialist), Jason Oke (Statistician) and Dawn Fraser. We also acknowledge all parents, children and clinicians who have been involved in the primary research studies. The Department of Primary Health Care is part of the National Institute of Health Research National School for Primary Care Research.

Katholieke Universiteit Leuven – official grant of the Fonds Wetenschappelijk Onderzoek Vlaanderen (FWO G.0232.04N) and an unconditional grant of Eurogenerics. We would also like to thank all participating physicians, parents and children.

Rotterdam – the emergency care nurses at Erasmus MC–Sophia Pediatric Emergency Care Department.

Leicester – Professor Terence Stephenson (Professor of Child Health, Institute of Child Health, professor at University of Nottingham during data collection), Dr Stephanie Smith (consultant paediatrician, University hospital Nottingham), Roddy MacFaul (Pinderfields Hospital), all parents and staff in the ED and on the paediatrics wards at University Hospital Nottingham. Dr Samiran Ray who contributed to the collection of the data set collected from Nottingham. Well Child Charity who funded Dr Lakhanpaul during the period of the data collection from Nottingham. Professor Jacqueline Collier Norwich University who advised on data collection and interpretation on data set collected in Nottingham.

Contribution of authors

M Thompson conceived the study, collected and contributed original data, performed the literature search and analyses, drafted the manuscript and was responsible for submitting the final report.

A Van den Bruel conceived the study, performed the literature search and analyses, collected and contributed original data and drafted the manuscript.

J Verbakel performed data analysis and interpretation and drafted the manuscript.

M Lakhanpaul collected and contributed original data, assisted with the drafting of the manuscript and performed data analysis and interpretation.

T Haj-Hassan performed the literature search and analyses and assisted with the drafting of the manuscript.

R Stevens conducted analyses, interpreted results and contributed to the drafting of the manuscript.

H Moll collected and contributed original data, interpreted results and contributed to the drafting of the manuscript.

F Buntinx assisted with the drafting of the manuscript and supervised J Verbakel and his analyses.

M Berger collected and contributed original data, interpreted results and contributed to the drafting of the manuscript.

B Aertgeerts commented on drafts of the manuscript and supervised JV and his analyses.

R Oostenbrink collected and contributed original data, performed data analysis and interpretation and contributed to the drafting of the manuscript.

D Mant assisted with the interpretation of the analyses and the drafting of the manuscript.

All authors had full access to the data.

Publications

Van den Bruel A, Thompson MJ, Haj-Hassan T, Stevens R, Moll H, Lakhanpaul M, Mant D. Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review. *BMJ* 2011;**8**:342:d3082.

Van den Bruel A, Haj-Hassan T, Thompson M, Buntinx F, Mant D. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet* 2010;**375**:834–45.

Oostenbrink R, Thompson M, Steyerberg EW, ERNIE members. Barriers to translating diagnostic research in febrile children to clinical practice: a systematic review. *Arch Dis Child* 2012;in press.

References

1. Armon K, Stephenson T, Gabriel V, MacFaul R, Eccleston P, Werneke U, *et al.* Determining the common medical presenting problems to an accident and emergency department. *Arch Dis Child* 2001;**84**:390–2.
2. Fleming DM, Smith GE, Charlton JR, Charlton J, Nicoll A. Impact of infections on primary care – greater than expected. *Commun Dis Public Health* 2002;**5**:7–12.
3. Avery AJ, Groom L, Boot D, Earwicker S, Carlisle R. What problems do patients present with outside normal general practice surgery hours? A prospective study of the use of general practice and accident and emergency services. *J Public Health Med* 1999;**21**:88–94.
4. Brogan C, Pickard D, Gray A, Fairman S, Hill A. The use of out of hours health services: a cross sectional survey. *BMJ* 1998;**316**:524–7.
5. Munro J, Nicholl J, O’Cathain A, Knowles E. Impact of NHS direct on demand for immediate care: observational study. *BMJ* 2000;**321**:150–3.
6. Kai J. What worries parents when their preschool children are acutely ill, and why: a qualitative study. *BMJ* 1996;**313**:983–6.
7. MacFaul R. *The health of children and young people*. London: Office for National Statistics; 2004.
8. Van den Bruel A, Bartholomeeusen S, Aertgeerts B, Truyers C, Buntinx F. Serious infections in children: an incidence study in family practice. *BMC Fam Pract* 2006;**7**:23.
9. 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**:397–403.
10. Pearson GA, editor. *Why children die: a pilot Sstudy 2006; England (South West, North East and West Midlands), Wales and Northern Ireland*. London: CEMACH; 2008.
11. 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.
12. Bleeker SE, Moons KG, Derksen-Lubsen G, Grobbee DE, Moll HA. Predicting serious bacterial infection in young children with fever without apparent source. *Acta Paediatr* 2001;**90**:1226–32.
13. Kempe A, Bunik M, Ellis J, Magid D, Hegarty T, Dickinson LM, *et al.* How safe is triage by an after-hours telephone call center? *Pediatrics* 2006;**118**:457–63.
14. Giesen P, Ferwerda R, Tijssen R, Mookink H, Drijver R, van den Bosch W, *et al.* Safety of telephone triage in general practitioner cooperatives: do triage nurses correctly estimate urgency? *Qual Saf Health Care* 2007;**16**:181–4.
15. Thompson M, Mayon-White R, Harnden A, Perera R, McLeod D, Mant D. Using vital signs to assess children with acute infections: a survey of current practice. *Br J Gen Pract* 2008;**58**:236–41.
16. Stiell I, Wells G, Laupacis A, Brison R, Verbeek R, Vandemheen K, *et al.* Multicentre trial to introduce the Ottawa ankle rules for use of radiography in acute ankle injuries. Multicentre Ankle Rule Study Group. *BMJ* 1995;**311**:594–7.
17. Reilly BM, Evans AT. Translating clinical research into clinical practice: impact of using prediction rules to make decisions. *Ann Intern Med* 2006;**144**:201–9.

18. Van den Bruel A, Cleemput I, Aertgeerts B, Ramaekers D, Buntinx F. The evaluation of diagnostic tests: evidence on technical and diagnostic accuracy, impact on patient outcome and cost-effectiveness is needed. *J Clin Epidemiol* 2007;**60**:1116–22.
19. Sanders S, Barnett A, Correa-Velez I, Coulthard M, Doust J. Systematic review of the diagnostic accuracy of C-reactive protein to detect bacterial infection in nonhospitalized infants and children with fever. *J Pediatr* 2008;**153**:570–4.
20. Falk G, Fahey T. C-reactive protein and community-acquired pneumonia in ambulatory care: systematic review of diagnostic accuracy studies. *Fam Pract* 2009;**26**:10–21.
21. van der Meer V, Neven AK, van den Broek PJ, Assendelft WJ. Diagnostic value of C reactive protein in infections of the lower respiratory tract: systematic review. *BMJ* 2005;**331**:26.
22. Simon L, Gauvin F, Amre DK, Saint-Louis P, Lacroix J. Serum procalcitonin and C-reactive protein levels as markers of bacterial infection: a systematic review and meta-analysis. *Clin Infect Dis* 2004;**39**:206–17.
23. Flood RG, Badik J, Aronoff SC. The utility of serum C-reactive protein in differentiating bacterial from nonbacterial pneumonia in children: a meta-analysis of 1230 children. *Pediatr Infect Dis J* 2008;**27**:95–9.
24. Mantadakis E, Plessa E, Vouloumanou EK, Karageorgopoulos DE, Chatzimichael A, Falagas ME. Serum procalcitonin for prediction of renal parenchymal involvement in children with urinary tract infections: a meta-analysis of prospective clinical studies. *J Pediatr* 2009;**155**:875–81.
25. Oostenbrink R, Moons KG, Bleeker SE, Moll HA, Grobbee DE. Diagnostic research on routine care data: prospects and problems. *J Clin Epidemiol* 2003;**56**:501–6.
26. Mackway-Jones K. *Emergency triage: Manchester Triage Group*. London: BMJ Publishing Group; 1996.
27. Cooke MW, Jinks S. Does the Manchester triage system detect the critically ill? *J Accid Emerg Med* 1999;**16**:179–81.
28. Roukema J, Steyerberg EW, van Meurs A, Ruige M, van der Lei J, Moll HA. Validity of the Manchester Triage System in paediatric emergency care. *Emerg Med J* 2006;**23**:906–10.
29. Thompson M, Coad N, Harnden A, Mayon-White R, Perera R, Mant D. How well do vital signs identify children with serious infections in paediatric emergency care? *Arch Dis Child* 2009;**94**:888–93.
30. Baumann MR, Strout TD. Evaluation of the Emergency Severity Index (version 3) triage algorithm in pediatric patients. *Acad Emerg Med* 2005;**12**:219–24.
31. Tanabe P, Gimbel R, Yarnold PR, Adams JG. The Emergency Severity Index (version 3) 5-level triage system scores predict ED resource consumption. *J Emerg Nurs* 2004;**30**:22–9.
32. Gouin S, Gravel J, Amre DK, Bergeron S. Evaluation of the Paediatric Canadian Triage and Acuity Scale in a pediatric ED. *Am J Emerg Med* 2005;**23**:243–7.
33. Chamberlain JM, Patel KM, Pollack MM. The Pediatric Risk of Hospital Admission score: a second-generation severity-of-illness score for pediatric emergency patients. *Pediatrics* 2005;**115**:388–95.
34. Gorelick MH, Lee C, Cronan K, Kost S, Palmer K. Pediatric emergency assessment tool (PEAT): a risk-adjustment measure for pediatric emergency patients. *Acad Emerg Med* 2001;**8**:156–62.

35. Nelson KG. An index of severity for acute paediatric illness. *Am J Public Health* 1980;**70**:804–7.
36. 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.
37. Hewson P, Poulakis Z, Jarman F, Kerr J, McMaster D, Goodge J, *et al.* Clinical markers of serious illness in young infants: a multicentre follow-up study. *J Paediatr Child Health* 2000;**36**:221–5.
38. Hewson PH, Gollan RA. A simple hospital triaging system for infants with acute illness. *J Paediatr Child Health* 1995;**31**:29–32.
39. National Institute for Clinical Excellence. *Feversh illness in children – assessment and initial management in children younger than 5 years*. London: NICE; 2007.
40. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. *JAMA* 2000;**284**:79–84.
41. National Institute for Health and Clinical Excellence. *Urinary tract infection in children: diagnosis, treatment and long-term management*. London: NICE; 2007.
42. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:25.
43. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 1994;**271**:703–7.
44. Dukic V, Gatsonis C. Meta-analysis of diagnostic test accuracy assessment studies with varying number of thresholds. *Biometrics* 2003;**59**:936–46.
45. Andreola B, Bressan S, Callegaro S, Liverani A, Plebani M, Da Dalt L. Procalcitonin and C-reactive protein as diagnostic markers of severe bacterial infections in febrile infants and children in the emergency department. *Pediatr Infect Dis J* 2007;**26**:672–7.
46. Baker MD, Bell LM, Avner JR. Outpatient management without antibiotics of fever in selected infants. *N Engl J Med* 1993;**329**:1437–41.
47. Baker MD, Bell LM, Avner JR. The efficacy of routine outpatient management without antibiotics of fever in selected infants. *Pediatrics* 1999;**103**:627–31.
48. Baker MD, Avner JR, Bell LM. Failure of infant observation scales in detecting serious illness in febrile, 4- to 8-week-old infants. *Pediatrics* 1990;**85**:1040–3.
49. Berger RM, Berger MY, van Steensel-Moll HA, Dzoljic-Danilovic G, Derksen-Lubsen G. A predictive model to estimate the risk of serious bacterial infections in febrile infants. *Eur J Pediatr* 1996;**155**:468–73.
50. Bleeker SE, Derksen-Lubsen G, Grobbee DE, Donders AR, Moons KG, Moll HA. Validating and updating a prediction rule for serious bacterial infection in patients with fever without source. *Acta Paediatr* 2007;**96**:100–4.
51. 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**:111–14.
52. Lacour AG, Gervaix A, Zamora SA, Vadas L, Lombard PR, Dayer JM, *et al.* Procalcitonin, IL-6, IL-8, IL-1 receptor antagonist and C-reactive protein as identifiers of serious

- bacterial infections in children with fever without localising signs. *Eur J Pediatr* 2001;**160**:95–100.
53. Galetto-Lacour A, Zamora SA, Gervais A. Bedside procalcitonin and C-reactive protein tests in children with fever without localizing signs of infection seen in a referral center. *Pediatrics* 2003;**112**:1054–60.
 54. Lacour AG, Zamora SA, Gervais A. A score identifying serious bacterial infections in children with fever without source. *Pediatr Infect Dis J* 2008;**27**:654–6.
 55. Garra G, Cunningham SJ, Crain EF. Reappraisal of criteria used to predict serious bacterial illness in febrile infants less than 8 weeks of age. *Acad Emerg Med* 2005;**12**:921–5.
 56. Grupo de Trabajo sobre el Nino Febril de la Sociedad Espanola de Urgencias de P. The young febrile child. Results of a multicenter survey. *An Esp Pediatr* 2001;**55**:5–10.
 57. Hsiao AL, Chen L, Baker MD. Incidence and predictors of serious bacterial infections among 57- to 180-day-old infants. *Pediatrics* 2006;**117**:1695–701.
 58. McCarthy PL, Lembo RM, Fink HD, Baron MA, Cicchetti DV. Observation, history, and physical examination in diagnosis of serious illnesses in febrile children less than or equal to 24 months. *J Pediatr* 1987;**110**:26–30.
 59. Nademi Z, Clark J, Richards CG, Walshaw D, Cant AJ. The causes of fever in children attending hospital in the north of England. *J Infect* 2001;**43**:221–5.
 60. Thayyil S, Shenoy M, Hamaluba M, Gupta A, Frater J, Verber IG. Is procalcitonin useful in early diagnosis of serious bacterial infections in children? *Acta Paediatr* 2005;**94**:155–8.
 61. Trautner BW, Caviness AC, Gerlacher GR, Demmler G, Macias CG. Prospective evaluation of the risk of serious bacterial infection in children who present to the emergency department with hyperpyrexia (temperature of 106 degrees F or higher). *Pediatrics* 2006;**118**:34–40.
 62. Crocker PJ, Quick G, McCombs W. Occult bacteremia in the emergency department: diagnostic criteria for the young febrile child. *Ann Emerg Med* 1985;**14**:1172–7.
 63. Haddon RA, Barnett PL, Grimwood K, Hogg GG. Bacteraemia in febrile children presenting to a paediatric emergency department. *Med J Aust* 1999;**170**:475–8.
 64. Jaffe DM, Fleisher GR. Temperature and total white blood cell count as indicators of bacteremia. *Pediatrics* 1991;**87**:670–4.
 65. Osman O, Brown D, Beattie T, Midgley P. Management of febrile children in a paediatric emergency department. *Health Bull* 2002;**60**:33–9.
 66. Teele DW, Pelton SI, Grant MJ, Herskowitz J, Rosen DJ, Allen CE, *et al*. Bacteremia in febrile children under 2 years of age: results of cultures of blood of 600 consecutive febrile children seen in a “walk-in” clinic. *J Pediatr* 1975;**87**:227–30.
 67. Waskerwitz S, Berkelhamer JE. Outpatient bacteremia: clinical findings in children under two years with initial temperatures of 39.5 degrees C or higher. *J Pediatr* 1981;**99**:231–3.
 68. Gorelick MH, Shaw KN, Murphy KO. Validity and reliability of clinical signs in the diagnosis of dehydration in children. *Pediatrics* 1997;**99**:E6.
 69. Shavit I, Brant R, Nijssen-Jordan C, Galbraith R, Johnson DW. A novel imaging technique to measure capillary-refill time: improving diagnostic accuracy for dehydration in young children with gastroenteritis. *Pediatrics* 2006;**118**:2402–8.
 70. Joffe A, McCormick M, DeAngelis C. Which children with febrile seizures need lumbar puncture? A decision analysis approach. *Am J Dis Child* 1983;**137**:1153–6.

71. Offringa M, Beishuizen A, Derksen-Lubsen G, Lubsen J. Seizures and fever: can we rule out meningitis on clinical grounds alone? *Clin Pediatr* 1992;**31**:514–22.
72. Oostenbrink R, Moons KG, Donders AR, Grobbee DE, Moll HA. Prediction of bacterial meningitis in children with meningeal signs: reduction of lumbar punctures. *Acta Paediatr* 2001;**90**:611–17.
73. Mahabee-Gittens EM, Grupp-Phelan J, Brody AS, Donnelly LF, Bracey SE, Duma EM, *et al.* Identifying children with pneumonia in the emergency department. *Clin Pediatr* 2005;**44**:427–35.
74. Taylor JA, Del Beccaro M, Done S, Winters W. Establishing clinically relevant standards for tachypnea in febrile children younger than 2 years. *Arch Pediatr Adolesc Med* 1995;**149**:283–7.
75. Nielsen HE, Andersen EA, Andersen J, Böttiger B, Christiansen KM, Daugbjerg P, *et al.* Diagnostic assessment of haemorrhagic rash and fever. *Arch Dis Child* 2001;**85**:160–5.
76. Wells LC, Smith JC, Weston VC, Collier J, Rutter N. The child with a non-blanching rash: how likely is meningococcal disease? *Arch Dis Child* 2001;**85**:218–22.
77. Bachur RG, Harper MB. Predictive model for serious bacterial infections among infants younger than 3 months of age. *Pediatrics* 2001;**108**:311–16.
78. Van den Bruel A, Haj-Hassan T, Thompson M, Buntinx F, Mant D; European Research Network on Recognising Serious Infection investigators. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet* 2001;**375**:834–45.
79. 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**:361–7.
80. Buntinx F, Knockaert D, Bruyninckx R, de Blaey N, Aerts M, Knottnerus JA, *et al.* Chest pain in general practice or in the hospital emergency department: is it the same? *Fam Pract* 2001;**18**:586–9.
81. Morley CJ, Thornton AJ, Green SJ, Cole TJ. Field trials of the Baby Check score card in general practice. *Arch Dis Child* 1991;**66**:111–14.
82. Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061–6.
83. Toll DB, Janssen KJ, Vergouwe Y, Moons KG. Validation, updating and impact of clinical prediction rules: a review. *J Clin Epidemiol* 2008;**61**:1085–94.
84. Van den Bruel A, Aertgeerts B, Buntinx F. Results of diagnostic accuracy studies are not always validated. *J Clin Epidemiol* 2006;**59**:559–66.
85. Buntinx F, Aertgeerts B, Aerts M, Bruyninckx R, Knottnerus JA, Van den Bruel A. Multivariate analysis in diagnostic studies. What are the possibilities? In Knottnerus JA and Buntinx F, editors. *The evidence based of clinical diagnosis*. London: Wiley-Blackwell; 2008.
86. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med* 1999;**130**:515–24.
87. Knottnerus JA, Leffers P. The influence of referral patterns on the characteristics of diagnostic tests. *J Clin Epidemiol* 1992;**45**:1143–54.
88. Almond S, Mant D, Thompson M. Diagnostic safety-netting. *Br J Gen Pract* 2009;**59**:872–4; discussion 874.

89. Bleeker SE, Moll HA, Steyerberg EW, Donders AR, Derksen-Lubsen G, Grobbee DE. External validation is necessary in prediction research: a clinical example. *J Clin Epidemiol* 2003;**56**:826–32.
90. Moons KG, Harrell FE, Steyerberg EW. Should scoring rules be based on odds ratios or regression coefficients? *J Clin Epidemiol* 2002;**55**:1054–5.
91. Steyerberg EW, Bleeker SE, Moll HA, Grobbee DE, Moons KG. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *J Clin Epidemiol* 2003;**56**:441–7.
92. Bonadio WA. Evaluation and management of serious bacterial infections in the febrile young infant. *Pediatr Infect Dis J* 1990;**9**:905–12.
93. Graneto JW, Soglin DE. Maternal screening of childhood fever by palpation. *Pediatr Emerg Care* 1996;**12**:183–4.
94. National Institute for Health and Clinical Excellence. *Feverish illness in children: a quick reference guide*. London: NICE; 2007.
95. Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. *Ann Emerg Med* 1992;**21**:384–90.
96. Kuppermann N, Holmes JF, Dayan PS, Hoyle JD Jr, Atabaki SM, Holubkov R, *et al*. Identification of children at very low risk of clinically-important brain injuries after head trauma: a prospective cohort study. *Lancet* 2009;**374**:1160–70.
97. Glasziou P, Irwig L, Deeks JJ. When should a new test become the current reference standard? *Ann Intern Med* 2008;**149**:816–22.
98. 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.
99. Reitsma JB, Rutjes AW, Khan KS, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol* 2009;**62**:797–806.
100. Oostenbrink R, Moll HA, Moons KG, Grobbee DE. Predictive model for childhood meningitis. *Pediatr Infect Dis J* 2004;**23**:1070–1.
101. Nigrovic LE, Kuppermann N, Macias CG, Cannavino CR, Moro-Sutherland DM, Schremmer RD, *et al*. Clinical prediction rule for identifying children with cerebrospinal fluid pleocytosis at very low risk of bacterial meningitis. *JAMA* 2007;**297**:52–60.

Appendix 1

Search terms used for systematic review

MEDLINE	EMBASE	CINAHL	DARE
1 "Arthritis, Infectious"[Mesh] OR "Bone Diseases, Infectious"[Mesh] OR "Community-Acquired Infections"[Mesh] OR "Respiratory Tract Infections"[Mesh] OR "Sepsis"[Mesh] OR "Skin Diseases, Infectious"[Mesh] OR "Soft Tissue Infections"[Mesh] OR "Urinary Tract Infections"[Mesh] OR "Meningitis"[Mesh] OR meningitis OR serious infections OR "Gastroenteritis"[Mesh]	'infectious arthritis'/exp OR 'hematogenous osteomyelitis'/ exp OR 'communicable disease'/exp OR 'respiratory tract infection'/exp OR 'sepsis'/ exp OR 'skin infection'/exp OR 'soft tissue infection'/exp OR 'urinary tract infection'/ exp OR 'meningitis'/exp OR 'gastroenteritis'/exp OR serious AND infections	"Arthritis, Infectious"[Mesh] OR "Bone Diseases, Infectious"[Mesh] OR "Community-Acquired Infections"[Mesh] OR "Respiratory Tract Infections"[Mesh] OR "Sepsis"[Mesh] OR "Skin Diseases, Infectious"[Mesh] OR "Soft Tissue Infections"[Mesh] OR "Urinary Tract Infections"[Mesh] OR "Meningitis"[Mesh] OR meningitis OR serious infections OR "Gastroenteritis"[Mesh]	exp Arthritis, Infectious/OR exp Bone Diseases, Infectious/ OR exp Community-Acquired Infections/OR exp Respiratory Tract Infections/OR exp SEPSIS/OR exp Skin Diseases, Infectious/OR exp Soft Tissue Infections/OR exp Urinary Tract Infections/OR exp Meningitis/ OR exp GASTROENTERITIS/OR serious infections.mp. [mp=title, subject heading word, abstract, instrumentation]
2 "Signs and Symptoms"[MeSH] OR signs and symptoms OR "Fever"[MeSH] OR fever OR fast breathing OR tachypnoea OR respiratory rate OR yale observation scale OR yale score OR yale scale OR Nelson score OR Nelson scale OR young infant observation scale OR "Tachycardia"[Mesh] OR fast heart rate OR capillary refill time	'physical disease by body function'/exp OR (signs AND symptoms) OR 'fever'/exp OR fever OR (fast AND breathing) OR tachypnoea OR (respiratory AND rate) OR (yale AND observation AND scale) OR (yale AND score) OR (yale AND scale) OR (nelson AND score) OR (nelson AND scale) OR (young AND infant AND observation AND scale) OR 'tachycardia'/ exp OR (fast AND heart AND rate) OR (capillary AND refill AND time)	("Signs and Symptoms"[MeSH] OR signs and symptoms OR "Fever"[MeSH] OR fever OR fast breathing OR tachypnoea OR respiratory rate OR yale observation scale OR yale score OR yale scale OR Nelson score OR Nelson scale OR young infant observation scale OR "Tachycardia"[Mesh] OR fast heart rate OR capillary refill time)	(signs and symptoms).mp. [mp=title, subject heading word, abstract, instrumentation] OR exp FEVER/OR exp Respiratory Rate/OR tachypnoea.mp. OR fast breathing.mp. OR yale observation scale.mp. OR yale score.mp. OR yale scale.mp. OR nelson score.mp. OR nelson scale.mp. OR young infant observation scale.mp. OR exp TACHYCARDIA/OR fast heart rate.mp. OR capillary refill time. mp.
3 "Laboratory Techniques and Procedures"[Mesh]	'laboratory diagnosis'/exp	("Laboratory Techniques and Procedures"[Mesh])	exp Diagnosis, Laboratory/
4 "infant"[MeSH Terms] OR "child"[MeSH Terms] OR "adolescent"[MeSH Terms] OR paediatric [All fields] OR pediatric [All fields] OR "pediatrics" [MeSH term] OR p*ediatric* OR child* OR infant* OR bab* OR neonat* OR newborn* OR toddler*	'infant'/exp OR 'preschool child'/exp OR 'school child'/ exp OR 'toddler'/exp OR 'adolescent'/exp OR 'pediatrics'/ exp OR p*ediatric* OR child* OR infant* OR bab* OR neonat* OR newborn* OR toddler*	"infant"[MeSH Terms] OR "child"[MeSH Terms] OR "adolescent"[MeSH Terms]	exp INFANT/OR exp CHILD/OR exp Adolescence/OR paediatric. mp. OR exp Pediatrics/OR child\$.mp. OR infant\$.mp. OR newborn\$.mp. OR bab\$.mp. OR neonat\$.mp. OR toddler\$.mp. OR exp Child, Preschool/
5 "Ambulatory Care"[Mesh] OR "Family Practice"[Mesh] OR general practice OR GP OR "Physicians, Family"[Mesh] OR "Primary Health Care"[Mesh] OR "Emergency Service, Hospital"[Mesh] OR primary care	'ambulatory care'/exp OR 'general practice'/exp OR (general AND practice) OR gp OR 'general practitioner'/ exp OR (family AND physician) OR 'primary medical care'/ exp OR (primary AND care) OR 'emergency ward'/exp	"Ambulatory Care"[Mesh] OR "Family Practice"[Mesh] OR general practice OR GP OR "Physicians, Family"[Mesh] OR "Primary Health Care"[Mesh] OR "Emergency Service, Hospital"[Mesh] OR primary care	exp Ambulatory Care/OR exp Family Practice/OR general practice.mp. OR GP.mp. OR exp Physicians, Family/OR exp Primary Health Care/OR exp Emergency Service/OR primary care.mp.
1 AND (2 OR 3) AND 4 AND 5	1 AND (2 OR 3) AND 4 AND 5	1 AND (2 OR 3) AND 4 AND 5	1 AND (2 OR 3) AND 4 AND 5

CINAHL, Cumulative Index to Nursing and Allied Health Literature; DARE, Database of Abstracts of Reviews of Effects.

Appendix 2

Details of the clinical prediction rules identified in the systematic review

Name of clinical prediction rule	Clinical features						Derivation study
All serious infections							
YOS	Quality of cry	Reaction to parents' stimulation	State variation	Colour	Hydration	Response to social overtures	McCarthy <i>et al.</i> ³⁶
Values	Strong OR not crying 1	Cries briefly 1	Stays awake 1	Pink 1	Skin normal 1	Smiles OR alerts 1	
	Whimpering 3	Cries off and on 3	Awakes with stimulation 3	Pale extremities 3	Dry mouth 3	Brief smile OR alerts briefly 3	
	Weak 5	Continual cry 5	Falls to sleep 5	Pale OR cyanotic 5	Skin doughy 5	No smile OR face anxious 5	
	Calculate the sum of all six feature values (cut-offs used in literature: 8, 9 or 10)						
Five-stage decision tree	Clinician instinct that something is wrong	Dyspnoea	Temperature > 39.95 °C	Diarrhoea	Age 15–25 months		Van den Bruel <i>et al.</i> ¹¹
Values	No 0	No or unknown 0	< 39.95 °C 0	No or unknown 0	No or unknown 0		
	Yes or unknown 1	Yes 1	> 39.95 °C 1	Yes 1	Yes 1		
	If yes to any of these five features						
Pneumonia							
Pneumonia rule no. 1	Parental concern illness is different			Shortness of breath			Van den Bruel <i>et al.</i> ¹¹
Values	If yes to any of these two features						
Pneumonia rule no. 2	Clinician concern illness is different			Shortness of breath			Van den Bruel <i>et al.</i> ¹¹
Values	If yes to any of these two features						
Meningitis							
Meningitis rule no. 1	Any abnormal neurological finding			Sought care < 48 hours			Offringa <i>et al.</i> ⁷¹
Values	If yes to any of these two features						
Meningitis rule no. 2	Petechiae		Nuchal rigidity		Coma		Joffe <i>et al.</i> ⁷⁰
Values	If yes to any of these three features						
Gastroenteritis with dehydration							
Gastroenteritis rule no. 1	Absent tears	Dry mucous membranes		Ill appearance	Poor peripheral circulation		Gorelick <i>et al.</i> ⁶⁸
Values	If yes to any two of these four features						

Appendix 3

Details of clinical features and cut-off values in prediction rules included in *Chapter 6*

Study	Tests included in prediction rule	Cut-off	Score
Baker 1993 ⁴⁶	Infant Observation Scale or	> 10	Positive
	WBC or	≥ 15,000/mm ³	Positive
	Spun urine specimen with or positive on bright-field microscopy or	≥ 10 WBC/HPF	Positive
	CSF WBC or	≥ 8/mm ³	Positive
	Positive Gram stain or Infiltrate on chest radiograph		Positive
Baker 1999 ⁴⁷	Ill appearance or		Positive
	WBC or	≥ 15,000/mm ³	Positive
	Spun urine specimen with ≥ 10 WBC/HPF or positive on bright-field microscopy or		Positive
	CSF WBC or	≥ 8/mm ³	Positive
	Positive Gram stain or Infiltrate on chest radiograph		Positive
Bleeker 2007 ⁵⁰	WBC (× 10 ⁹ /l)	< 10	0
		10–19	2
		20–29	4
		30–39	6
		≥ 40	8
	Serum CRP (mg/l)	0–99	First integer
		≥ 100	First and second integers Maximum = 16 points
≥ 70 WBC/μl in dipstick urinalysis	No	0	
	Yes	9	
Bonadio 1993 ⁵¹	Milwaukee protocol		Positive if not fulfilling the protocol
	Physical examination with normal clinical appearance and no sign of focal infection		
	Normal laboratory data:		
	CSF WBC count	< 10/ml	
	CBC WBC count	< 15,000/ml	
	Urinalysis WBC/HPF	≤ 5–10	
	And no bacteria/HPF		
	Dipstick for leucocyte esterase and nitrite	Negative	
	Chest radiograph if performed	No infiltrate	
	Reliable caretaker		
No allergy to beta-lactam antibiotics			
Private paediatrician contact			
Galletto-Lacour 2001 ⁵²	PCT or	>0.9 ng/ml	Positive
	CRP	>40 mg/l	Positive

Study	Tests included in prediction rule	Cut-off	Score
Galetto-Lacour 2001 ⁵²	PCT or	>0.9 ng/ml	Positive
	WBC	>15,000/mm ³	Positive
Galetto-Lacour 2008 ⁵⁴	PCT	<0.5	0
		≥0.5	2
		≥2	4
	CRP	<40	0
		40–99	2
		≥100	4
	Urine dipstick	Negative	0
		Positive	1
Garra 2005 ⁵⁵	Philadelphia protocol		Positive if not fulfilling the protocol
	Infant observation score ≤10		
	No recognisable bacterial infection on examination		
	WBC	<15,000/mm ³	
	Band-to-neutrophil ratio <0.2		
	Urine: WBC	<10/mm ³	
	And few bacteria per HPF		
	CSF: WBC	<8/mm ³	
And a negative Gram stain			
Chest radiograph	No infiltrate		
Stool smear negative for blood and few or no WBC (for infants with diarrhoea)			
Thayil 2005 ⁶⁰	PCT, and	>2	Positive
	CRP, and	>50	
	WBC	>15	

CBC, complete blood count; HPF, high-power field.

Appendix 4

Characteristics of variables included in the data sets used to validate clinical prediction rules

Data set	Van den Bruel <i>et al.</i> ¹¹	Oostenbrink <i>et al.</i> ⁷²	Roukema <i>et al.</i> ²⁸	Bleeker <i>et al.</i> ⁵⁰	Thompson <i>et al.</i> ²⁹	Brent <i>et al.</i> ⁷⁹	Berger <i>et al.</i>
Country	BE	NL	NL	NL	UK	UK	NL
Setting	GP/AP/ED	ED	ED	ED	PAU	ED	GP
Data set characteristics							
Overall number	3981	593	1750	595	700	2777	506
General characteristics							
Age	+	+	+	+	+	+	+
Gender	+	+	+	+	+	+	+
Past medical history							
History of severe chronic disease	+	-	-	-	+	+	+
Presenting illness							
Duration	+	+	+	+	+	+	+
Referred or primary presentation	+	+	+	+	+	+	+
Overall appearance of severity	+	+	+	+	+	+	+
Presenting symptoms/signs reported by parents or during physical examination							
Irritable	+	+	-	+	+	+	+
Neurological (headache, drowsy, impaired consciousness)	+	+	+	+	+	+	+
Different cry/vocalisation	+	+	-	+	+	-	+
Upper respiratory signs	+	+	+	+	+	+	+
Lower respiratory signs	+	+	+	+	+	+	+
Nausea/vomiting	+	+	+	+	+	+	+
Diarrhoea	+	-	+	-	+	+	+
Abdominal pain	+	-	-	+	+	+	+
Meningeal irritation	+	+	+	+	+	+	+
Rash	+	+	+	-	+	+	+
Colour	+	+	-	+	+	+	+
Impaired circulation signs	+	+	+	+	+	+	+
Vital signs							
Temperature	+	+	+	+	+	+	+
Heart rate	-	+	+	-	+	+	+
Respiratory rate	+	-	+	-	+	+	+
Capillary refill	-	+	+	-	+	+	+
Saturations	-	+	+	-	+	+	-
Investigations							

Data set	Van den Bruel <i>et al.</i> ¹¹	Oostenbrink <i>et al.</i> ⁷²	Roukema <i>et al.</i> ²⁸	Bleeker <i>et al.</i> ⁵⁰	Thompson <i>et al.</i> ²⁹	Brent <i>et al.</i> ⁷⁹	Berger <i>et al.</i>
Urine dip	–	+	+	+	+	+	+
Chest radiograph	–	–	+	–	+	+	–
Full blood count	–	+	+	+	+	+	–
CRP	–	+	+	+	+	+	+
Cultures (urine, blood, other)	–	+	+	–	+	+	–
Outcomes							
Admitted/sent home	+	+	+	+	+	+	+
Diagnosis	+	+	+	+	+	+	+
Treatment data	–	+	+	–	+	+	+
Number with serious infections (% of total)	31 (0.78%)	263 (44.35%)	227 (12.97%)	140 (23.53%)	313 (44.71%)	373 (13.43%)	35 (6.52%)

+, variable present; –, variable absent; AP, ambulatory paediatric care; BE, Belgium; NL, the Netherlands; PAU, paediatric assessment unit.

Appendix 5

Data sets identified which allowed multiple external validation of clinical prediction rules

Clinical prediction rule	Cases	Variables
All serious infections		
<i>YOS (cut-off > 10)</i>		
Van den Bruel <i>et al.</i> ¹¹	3945	obs_cry; obs_unconsolable; obs_drowsy; pe_cyanosis; pe_dehydration; obs_laughing
Berger <i>et al.</i>	482	Yale_cry; Yale_response; Yale_state; Yale_colour; Yale_hydration; Yale_social
Brent <i>et al.</i> ⁷⁹	2765	Yale_response; Yale_state; Yale_colour; Yale_hydration; Yale_social
Thompson <i>et al.</i> ²⁹	663	Yale_response; Yale_state; Yale_colour; Yale_hydration; Yale_respiratory; Yale_activity
<i>YOS (cut-off > 8)</i>		
Van den Bruel <i>et al.</i> ¹¹	3945	obs_cry; obs_unconsolable; obs_drowsy; pe_cyanosis; pe_dehydration; obs_laughing
Berger <i>et al.</i>	482	Yale_cry; Yale_response; Yale_state; Yale_colour; Yale_hydration; Yale_social
Brent <i>et al.</i> ⁷⁹	2765	Yale_response; Yale_state; Yale_colour; Yale_hydration; Yale_social
Oostenbrink <i>et al.</i> ⁷²	593	pe_cry or pe_diaperpain; pe_drowsy or pe_unconscious; pe_cyanosis or pe_skin
Thompson <i>et al.</i> ²⁹	663	Yale_response; Yale_state; Yale_colour; Yale_hydration; Yale_respiratory; Yale_activity
<i>Five-stage decision tree^a</i>		
^b Van den Bruel <i>et al.</i> ¹¹	3981	pe_something_is_wrong; pe_dyspnoea; temperature; anamn_diarrhoea; age
	3981	Child_seriously_ill?; pe_dyspnoea; temperature; pc_diarrhoea; age
Berger <i>et al.</i>	506	clinical_impression; pe_dyspnoea or pc_dyspnoea; temperature; pe_diarrhoea or pc_diarrhoea; age
Roukema <i>et al.</i> ²⁸	1750	how_ill_physician; pe_flairing or pc_flairing; temperature; diarrhoea; age
Thompson <i>et al.</i> ²⁹	700	toxic_appearance_nurse; pc_breathing_difficulty; temperature; pc_diarrhoea; age
Brent <i>et al.</i> ⁷⁹	2762	patient_status; respiratory_distress; temperature; pc_diarrhoea_vomitting; age
Bleeker <i>et al.</i> ⁵⁰	595	pe_clinical_impression; pe_respiratory_distress; temperature; age
Oostenbrink <i>et al.</i> ⁷²	593	pe_clinical_impression; pe_temperature; age
Pneumonia		
<i>Pneumonia rule^c</i>		
^b Van den Bruel <i>et al.</i> ¹¹	3981	Child_seriously_ill?; pe_dyspnoea
Berger <i>et al.</i>	506	clinical_impression; pe_dyspnoea or pc_dyspnoea
Brent <i>et al.</i> ⁷⁹	2183	patient_status; respiratory_distress
Bleeker <i>et al.</i> ⁵⁰	595	pe_clinical_impression; pe_respiratory_distress
Thompson <i>et al.</i> ²⁹	700	toxic_appearance_nurse; pc_breathing_difficulty
Roukema <i>et al.</i> ²⁸	1682	howill_physician; pe_flairing or pc_flairing

Clinical prediction rule	Cases	Variables
Meningitis		
<i>Meningitis rule^d</i>		
Van den Bruel <i>et al.</i> ¹¹	3981	pe_petechiae; pe.meningeal irritation; obs_unconscious
Thompson <i>et al.</i> ²⁹	700	nonblanching_rash_nurse; neckstiff_nurse; unconscious_nurse
Brent <i>et al.</i> ⁷⁹	2171	petechial_rash; AVPU_score=3: AVPU_score=4
Oostenbrink <i>et al.</i> ⁷²	593	pe_petechiae; pe_tripod_sign or pe_nuchal rigidity or pe_meningeal_irritation or pe_kernig or pe_brudzinski; pe_unconscious

obs, observation by parents; pe, physical examination; pc, presenting complaint.

a If yes to any of five sequential questions: (1) clinical instinct that something is wrong, (2) dyspnoea, (3) temperature > 39.5 °C, (4) diarrhoea or (5) age 15–29 months.

b *Italic denotes a derivation study.*

c If yes to any of (1) shortness of breath or (2) clinician's concern.

d If yes to any of (1) petechiae, (2) nuchal rigidity or (3) coma.

Appendix 6

Percentage of missing values of variables included in the clinical prediction rules

Clinical prediction rule	n cases	Prevalence ^a	Variables (% missing)					Response to social overtures
			Quality of cry	Reaction to parent stimulation	State variation	Colour	Hydration	
All serious infections								
<i>YOS (cut-off > 10)</i>								
Van den Bruel <i>et al.</i> ¹¹	3945	Low	1.70%	1.08%	1.03%	0.95%	1.68%	1.73%
Berger <i>et al.</i>	482	Low	4.74%	4.74%	4.94%	5.73%	6.52%	5.73%
Brent <i>et al.</i> ⁷⁹	2765	Intermediate	13.61%	–	0.65%	0.86%	0.47%	1.26%
Thompson <i>et al.</i> ²⁹	663	High	5.28%	52.71%	4.29%	10.00%	5.71%	53.71%
<i>YOS (cut-off > 8)</i>								
Van den Bruel <i>et al.</i> ¹¹	3945	Low	1.70%	1.08%	1.03%	0.95%	1.68%	1.73%
Berger <i>et al.</i>	482	Low	4.74%	4.74%	4.94%	5.73%	6.52%	5.73%
Brent <i>et al.</i> ⁷⁹	2765	Intermediate	13.61%	–	0.65%	0.86%	0.47%	1.26%
Thompson <i>et al.</i> ²⁹	663	High	28.67%	–	0.00%	0.00%	–	–
Oostenbrink <i>et al.</i> ⁷²	593	High	5.28%	52.71%	4.29%	10.00%	5.71%	53.71%
Something is wrong (1) or child seriously ill? (2)								
<i>Five-stage decision tree^b</i>								
^c Van den Bruel <i>et al.</i> ¹¹	3981	Low	2.06% (1)		0.73%	17.63%	0.78%	2.34%
	3981	Low	5.85% (2)		0.73%	17.63%	0.78%	2.34%
Berger <i>et al.</i>	506	Intermediate	5.53% (2)		38.14%	0.98%	0.59%	16.60%
Roukema <i>et al.</i> ²⁸	1750	Intermediate	42.40% (2)		47.49%	58.80%	47.60%	0.00%
Brent <i>et al.</i> ⁷⁹	2762	Intermediate	54.59% (2)		13.61%	4.00%	4.29%	15.34%
Bleeker <i>et al.</i> ⁵⁰	595	High	6.89% (2)		9.75%	1.18%	–	0.00%
Thompson <i>et al.</i> ²⁹	700	High	0.00% (2)		61.14%	1.43%	88.43%	0.00%
Oostenbrink <i>et al.</i> ⁷²	593	High	47.89% (2)		–	3.04%	–	0.00%
Clinician's concern something is wrong								
<i>Pneumonia</i>								
<i>Pneumonia rule^d</i>								
^c Van den Bruel <i>et al.</i> ¹¹	3981	Low	5.85%		0.73%			
Berger <i>et al.</i>	506	Low	5.53%		38.14%			
Roukema <i>et al.</i> ²⁸	1682	Intermediate	42.40%		47.49%			
Brent <i>et al.</i> ⁷⁹	2183	Intermediate	54.59%		13.61%			
Bleeker <i>et al.</i> ⁵⁰	595	Intermediate	6.89%		9.75%			
Thompson <i>et al.</i> ²⁹	700	Intermediate	0.00%		61.14%			

Clinical prediction rule	<i>n</i> cases	Prevalence ^a	Petechiae	Nuchal rigidity	Coma
<i>Meningitis</i>					
<i>Meningitis rule^c</i>					
Van den Bruel <i>et al.</i> ¹¹	3981	Low	0.68%	0.75%	1.00%
Thompson <i>et al.</i> ²⁹	700	Low	0.00%	0.00%	0.00%
Brent <i>et al.</i> ⁷⁹	2171	Low	0.00%	0.76%	0.76%
Oostenbrink <i>et al.</i> ⁷²	593	High	0.00%	0.00%	0.00%

a Setting: low prevalence of serious infection (< 5%); intermediate prevalence of serious infection (5–20%); high prevalence of serious infection (> 20%).

b If yes to any of five sequential questions: (1) clinical instinct that something is wrong, (2) dyspnoea, (3) temperature > 39.5 °C, (4) diarrhoea or (5) age 15–29 months.

c *Italic denotes a derivation study.*

d If yes to any of (1) shortness of breath or (2) clinician's concern.

e If yes to any of (1) petechiae, (2) nuchal rigidity or (3) coma.

Appendix 7

Research protocol

1. Project title: Systematic review and validation of clinical prediction rules for identifying children with serious infections in emergency departments and urgent-access primary care

2. How the project has changed since the outline proposal was submitted

The Board had several comments on the outline proposal which we have addressed.

a) We have increased the emergency care expertise on our team

Although our team already included Dr Henrietta Moll, Emergency Department Paediatrician at Sophia Children's Hospital in Rotterdam, we have also included Dr Shelly Segal who is the Lead Paediatrician in the Accident and Emergency Department at the John Radcliffe Hospital in Oxford, and Dr Monica Lakhanpaul who is a Consultant Paediatrician and who brings extensive experience as the project lead of the recent NICE guideline on the management of the febrile child and who is the clinical lead on the development of the nurse-led urgent care service in Leicester.

b) We have been more explicit about eligibility criteria

The Board wanted to see a strong justification of the patient eligibility criteria. We agree that for the proposed systematic review, more details of the inclusion and exclusion criteria are needed. We have included these in the full proposal.

c) We have reduced the scope of the study to focus on the systematic review element

The main concern of the Board was the feasibility of including the cross validation of prediction rules as well as the systematic review. Particular concern was expressed about our ability to secure access to the databases planned within the time frame. However we have already secured access to the five key datasets of which we are already aware and this will allow us to conduct an individual patient data based meta-analysis on at least this sub-set of studies.

The Board also had several comments noted on June 17, 2008 to which we responded on July 8, 2008 and which are incorporated into this final project description. Our responses are summarised as follows:

- (a) The applicability of the clinical prediction tool needs to be clearly defined in terms of who will be able and likely to use it.

We feel that the management of children presenting to emergency and urgent care settings with infections presents an ideal opportunity for application of a clinical prediction rule. In general clinical prediction rules are most likely to be helpful in situations where 'decision making is complex, the clinical stakes are high, or there are opportunities to achieve cost savings without compromising patient care' (McGinn, JAMA 2000). The clinical prediction tool that we will develop and cross-validate will incorporate components of the history,

vital signs and basic examination findings. We feel that this prediction rule will therefore be most applicable to front line clinicians, such as GPs, paramedics, practice nurses, A&E triage nurses and nurse practitioners, and A&E junior medical staff. The advantage of the methods planned is that we will be able to validate this rule in multiple clinical settings with varying prevalence of serious infection, and thus the prediction rule will be applicable in many different acute care settings in the NHS. If the rule is found to be robust we will disseminate the findings widely via appropriate peer reviewed publications and by contact with the relevant professional bodies. In addition, we anticipate that it will complement the current NICE guideline on the assessment of feverish children (Feverish illness in children, NICE May 2007) by formally testing, simplifying and quantifying the accuracy of many of the clinical predictors used in that guideline. As with any clinical prediction rule, the impact of the rule will need to be evaluated once it has been implemented.

- (b) The amount of time allocated to staff involved in the project appears too low in some cases, this should be reviewed along with the project costs.

We agree that the time allocated particularly to senior staff on this project was far too low in our application and would like to revise this, with the permission of the Board. We propose the following changes to staff hours over the course of the 1 year project: 1) increase in Dr Thompson's (Lead applicant) hours from 132 hours to 330 hours, 2) Professors Mant and Glasziou from 11 to 44 hours, 3) Dr Lakhanpaul's hours from 36 to 82 hours. The costs allocated to the other staffing costs both at Oxford (Dr Perera & Research Assistant), Oxford Radcliff Trust (Dr Segal), Leuven (Belgium) (Professor Buntinx, Drs Aertgert and Van en Bruel) and Rotterdam (Dr Moll) and Maastricht (Dr Dinant) have not been altered. These changes are outlined in the accompanying spreadsheet and will increase the overall budget to £125,657 (see attached spreadsheet).

- (c) The distribution of end points available for analysis should be described.

In all datasets we have the main outcomes recorded of need for admission to hospital, and number of children with serious infection. The definition of serious infection will be standardised across all datasets, but will include clinical conditions such as meningitis, UTI, bacterial gastroenteritis, pneumonia, sepsis.

- (d) The core items eligible for inclusion across the five datasets should be stated.

We have provisionally examined the core items from each of the datasets, from which we will identify predictors of serious infection. This shows that all datasets include details of the general characteristics of the children such as age, gender, as well as the setting and whether referred or not referred. The completeness of the presenting clinical features, i.e. symptoms and signs varies between datasets. All the datasets include the core vital signs heart rate, respiratory rate, and temperature, and some also include oxygen saturations and capillary refill time. Five of the datasets include a large number of clinical features identified from parental history or initial triage/examination. Two of the datasets include fewer clinical features. The number of investigations performed on children varies with type of clinical setting, and we are likely only to consider results of white cell count or C-reactive protein (CRP) as predictors. We will also use the systematic review that we will be undertaking to assist us in deciding which predictors have been most useful in previous studies in this area.

- (e) The project should consider alternatives to splitting the dataset randomly in half.

We agree that splitting the data may not be the best method for validating the prediction rule. The two issues that we will aim to address are: a) over-optimistic estimates and b) transferability of the prediction rule across settings. To achieve this we will use k-fold cross-validation to obtain more realistic estimates and calibration using other datasets to test the transferability of the model. By validating the clinical prediction rules on patients in broader settings (and thus different disease prevalence and spectrum) from those used to derive the rule, we will be able to demonstrate the generalizability or external validity of the rule (McGinn *et al.*, JAMA 2000). We anticipate that this process will require model revision and/or shrinkage methods (Steyerberg EW *et al.* Statist Med 2004).

- (f) Service user involvement.

We did not specify the level of service user involvement in the application as we did not feel that it was particularly relevant to this type of study. However, we agree that it would be useful to involve parents/carers input in assessing the likely impact of this rule in the real world setting, and to ensure that the predictors we identify (e.g. vital sign measurements, possibly blood tests) are acceptable to most parents/carers. We will therefore assemble a group of parents who have had personal experience with children in emergency care or urgent access primary care and obtain their input on the final prediction rules.

3. Planned investigation

Research objectives

The overall aim of this research is to systematically identify simple clinical decision rules which can allow children with self-limiting illness to be safely discharged from emergency and urgent primary care settings while not missing any cases of serious infection. We propose to undertake a systematic review of the literature on prediction rules for triaging children with acute illness in emergency and urgent care settings.

The specific objectives of the systematic review are:

1. To identify the clinical features and decision rules which have already been shown to have predictive value for identifying (or excluding) children with severe infection.
2. To identify and compare the best performing prediction rules from the literature.
3. To explore the added value of including laboratory tests and vital signs to prediction rules based on clinical history and observation.

Clinical prediction rules are a simple pragmatic technology that can be used by clinical staff to assist them in assessment and clinical management. A widely implemented example which has been shown to reduce both resource use and missed diagnoses in A&E is the Ottawa Ankle Rule for ordering an X-ray.¹ The marginal NHS cost of implementing a clinical prediction rule depends primarily on the cost of any additional staff time or investigations required. The prediction rules that we propose validating have very low marginal cost because the main components are an integral part of the standard clinical assessment of children that clinicians use in routine NHS practice (i.e. medical history, presenting complaints, vital signs and examination findings).

The main economic benefit to the NHS is the potential to reduce the need for urgent hospital admission by reliably identifying the vast majority of children who can safely be discharged home or to lower acuity care (e.g. GP follow up). However, more effective triage using a formal prediction rule will also improve the care of children with serious infections (e.g. by signalling the need for 999 transfer to A+E, for urgent diagnostic or therapeutic intervention, or for urgent paediatrician review), thus optimising use and effectiveness of emergency services. If shown to discriminate effectively, such a prediction rule would be used at several levels of the emergency medical system in the UK, including paramedics, walk-in or out of hours surgeries, paediatric assessment units, as well as A&E Departments.

Existing research

Acute illness is one of the most common problems encountered in children attending emergency departments as well as by urgent-access primary care services in the UK.

Between 27–47% of patients who present to A&E departments in the UK do so for medical illness, rather than trauma.² For children, the most common medical reasons for attending A&E are breathing difficulty (31%), febrile illness (20%), diarrhoea/vomiting (16%), abdominal pain (6%), seizure (5%), or rash (5%).²

Children under 5 years of age also constitute a substantial part of the workload of urgent-access primary care services. Indeed, the patient group which presents most commonly to out-of-hours assessment clinics is children with acute infections.^{3,4} Similarly, acute illness in children is also a major component of the work of NHS Direct, where 22% of all telephone calls are related to children under 5 years of age.⁵

One of the key tasks in both hospital emergency departments and urgent-access primary care clinics is therefore to distinguish children who may have serious infections or complications of infections (e.g. meningitis, bacteraemia, hypoxia from bronchiolitis, dehydration from gastroenteritis) from the vast majority with self-limiting or minor infections who can safely be managed as outpatients or referred to primary care services. This task is challenging. With increasing A&E attendance rates in the UK, hospital admission of children is becoming more common despite a falling incidence of serious infection. At the same time, approximately half of children with meningococcal disease are still missed at first consultation with a doctor, which results in poorer health outcome.⁶ If the simple clinical decision rules we plan to assess are shown to be effective, they are likely to be welcomed and widely adopted.

There are several triage systems currently in use in emergency departments in the UK. The Manchester Triage System assigns the patient to one of five categories based on the maximum time that they can wait for full assessment.^{7,8} It provides only modest sensitivity (63%) to detect emergency or very urgent cases and is a generic instrument to deal with emergencies including trauma.⁹ Other triage systems used internationally include the Emergency Severity Index, the Paediatric Canadian Triage and Acuity Scale, Paediatric Risk of Admission Score, and the Paediatric Emergency Assessment Tool.^{10–14} A number of more specific ‘scoring systems’ for children presenting to emergency departments with medical illness have been developed. None have shown sufficient ability to rule out serious infection in children to be widely adopted in an NHS context.^{15–18}

The National Institute for Health and Clinical Excellence (NICE) guideline for the management of feverish illness in children under 5 years of age was published in 2007.¹⁹ It is an important starting point for us because its recommendations are based on a literature review utilising stakeholders to identify key documents. However, we are aware of important recent studies which were not included and no attempt was made to explore the data at an individual patient level.

Conducting individual patient level meta-analysis is important to provide evidence to underpin several of the NICE recommendations – for example the recommendation that ‘Healthcare professionals should measure and record temperature, heart rate, respiratory rate and capillary refill time as part of the routine assessment of a child with fever.’¹⁹

Research methods

The proposed project will involve a systematic review of the literature on clinical predictors of serious infection in children, including systematic review and standard meta-analysis where appropriate of all studies and individual-patient data meta-analysis of identified studies on unselected populations (the studies most likely to provide reliable predictive values for triage in urgent access primary care and emergency care settings in the UK).

In our preparatory work developing this protocol (to assess the size of the task and the feasibility of individual patient data meta-analysis) we have identified four studies published in the past 20 years in unselected populations of children presenting to urgent-access primary care (from Belgium²⁰) and children presenting to emergency department (from the Netherlands^{21,22}) but have been made aware of three substantial but as yet unpublished datasets from the UK. The investigators of all seven studies (total sample 11,328) have agreed to supply us with their individual patient data and support this analysis (letters available).

Literature search

The literature will be searched in MEDLINE, EMBASE, DARE and CINAHL. The search strategy will consist of a combination of terms on serious infections, terms referring to ‘signs and symptoms’, laboratory tests, children, ambulatory care and infections, using both MeSH terms and free text words if appropriate. In addition, the reference lists of the articles thus retrieved will be checked. A search for any unpublished material will consist of contacting known researchers in the field.

The research team have already performed a provisional literature search using the proposed search strategy: 892 articles were identified from MEDLINE, 718 from EMBASE, 7 from DARE and 86 from CINAHL. After duplicates had been discarded, the total number of citations was 1578, as detailed below:

MEDLINE

5	Infections	“Arthritis, Infectious”[Mesh] OR “Bone Diseases, Infectious”[Mesh] OR “Community-Acquired Infections”[Mesh] OR “Respiratory Tract Infections”[Mesh] OR “Sepsis”[Mesh] OR “Skin Diseases, Infectious”[Mesh] OR “Soft Tissue Infections”[Mesh] OR “Urinary Tract Infections”[Mesh] OR “Meningitis”[Mesh] OR meningitis OR serious infections OR “Gastroenteritis”[Mesh]	586,606
7	Signs and symptoms	“Signs and Symptoms”[MeSH] OR signs and symptoms OR “Fever”[MeSH] OR fever OR fast breathing OR tachypnoea OR respiratory rate OR yale observation scale OR yale score OR yale scale OR Nelson score OR Nelson scale OR young infant observation scale OR “Tachycardia”[Mesh] OR fast heart rate OR capillary refill time	1,324,204
9	Laboratory tests	“Laboratory Techniques and Procedures”[Mesh]	1,158,471
12	Child	“infant”[MeSH Terms] OR “child”[MeSH Terms] OR “adolescent”[MeSH Terms] OR paediatric [All fields] OR pediatric [All fields] OR “pediatrics” [MeSH term] OR p*ediatric* OR child* OR infant* OR bab* OR neonat* OR newborn* OR toddler*	2,584,184
14bis	Ambulatory care	“Ambulatory Care”[Mesh] OR “Family Practice”[Mesh] OR general practice OR GP OR “Physicians, Family”[Mesh] OR “Primary Health Care”[Mesh] OR “Emergency Service, Hospital”[Mesh] OR primary care	208,882
24	Combination	5 AND (7 OR 9) AND 12 AND 14bis	892

EMBASE

E1	Infections	'infectious arthritis'/exp OR 'hematogenous osteomyelitis'/exp OR 'communicable disease'/exp OR 'respiratory tract infection'/exp OR 'sepsis'/exp OR 'skin infection'/exp OR 'soft tissue infection'/exp OR 'urinary tract infection'/exp OR 'meningitis'/exp OR 'gastroenteritis'/exp OR serious AND infections	73,777
E2	Signs and symptoms	'physical disease by body function'/exp OR (signs AND symptoms) OR 'fever'/exp OR fever OR (fast AND breathing) OR tachypnoea OR (respiratory AND rate) OR (yale AND observation AND scale) OR (yale AND score) OR (yale AND scale) OR (nelson AND score) OR (nelson AND scale) OR (young AND infant AND observation AND scale) OR 'tachycardia'/exp OR (fast AND heart AND rate) OR (capillary AND refill AND time)	4,255,612
E3	Laboratory tests	'laboratory diagnosis'/exp	91,178
	Children	'infant'/exp OR 'preschool child'/exp OR 'school child'/exp OR 'toddler'/exp OR 'adolescent'/exp OR 'pediatrics'/exp OR <i>p</i> *ediatric* OR child* OR infant* OR bab* OR neonat* OR newborn* OR toddler*	2,578,259
E4	Ambulatory care	'ambulatory care'/exp OR 'general practice'/exp OR (general AND practice) OR gp OR 'general practitioner'/exp OR (family AND physician) OR 'primary medical care'/exp OR (primary AND care) OR 'emergency ward'/exp	362,300
E5	Combination	E1 AND (E2 OR E3) AND E4	718

DARE

D1	Infections	"Arthritis, Infectious"[Mesh] OR "Bone Diseases, Infectious"[Mesh] OR "Community-Acquired Infections"[Mesh] OR "Respiratory Tract Infections"[Mesh] OR "Sepsis"[Mesh] OR "Skin Diseases, Infectious"[Mesh] OR "Soft Tissue Infections"[Mesh] OR "Urinary Tract Infections"[Mesh] OR "Meningitis"[Mesh] OR meningitis OR serious infections OR "Gastroenteritis"[Mesh]	254
D2	Signs and symptoms OR laboratory tests	("Signs and Symptoms"[MeSH] OR signs and symptoms OR "Fever"[MeSH] OR fever OR fast breathing OR tachypnoea OR respiratory rate OR yale observation scale OR yale score OR yale scale OR Nelson score OR Nelson scale OR young infant observation scale OR "Tachycardia"[Mesh] OR fast heart rate OR capillary refill time)	495
D3	Laboratory tests	("Laboratory Techniques and Procedures"[Mesh])	17
D4	Child	"infant"[MeSH Terms] OR "child"[MeSH Terms] OR "adolescent"[MeSH Terms] OR paediatric [All fields] OR pediatric [All fields] OR "pediatrics" [MeSH term] OR <i>p</i> *ediatric* OR child* OR infant* OR bab* OR neonat* OR newborn* OR toddler*	0
D5	Child	"infant"[MeSH Terms] OR "child"[MeSH Terms] OR "adolescent"[MeSH Terms] OR paediatric [All fields] OR pediatric [All fields] OR "pediatrics" [MeSH term] OR child* OR infant* OR bab* OR neonat* OR newborn* OR toddler*	0
D6	Child	"infant"[MeSH Terms] OR "child"[MeSH Terms] OR "adolescent"[MeSH Terms] OR "pediatrics" [MeSH term] OR child* OR infant* OR bab* OR neonat* OR newborn* OR toddler*	0
D7	Child	"infant"[MeSH Terms] OR "child"[MeSH Terms] OR "adolescent"[MeSH Terms]	973
D8	Ambulatory care	"Ambulatory Care"[Mesh] OR "Family Practice"[Mesh] OR general practice OR GP OR "Physicians, Family"[Mesh] OR "Primary Health Care"[Mesh] OR "Emergency Service, Hospital"[Mesh] OR primary care	2346
D9	Combination	D1 AND (D2 OR D3) AND D7 AND D8	7

CINAHL

C1	Infections	exp Arthritis, Infectious/OR exp Bone Diseases, Infectious/OR exp Community-Acquired Infections/OR exp Respiratory Tract Infections/OR exp SEPSIS/OR exp Skin Diseases, Infectious/OR exp Soft Tissue Infections/OR exp Urinary Tract Infections/OR exp Meningitis/OR exp GASTROENTERITIS/OR serious infections.mp. [mp=title, subject heading word, abstract, instrumentation]	30,021
C2	Signs and symptoms	(signs and symptoms).mp. [mp=title, subject heading word, abstract, instrumentation] OR exp FEVER/OR exp Respiratory Rate/OR tachypnoea.mp. OR fast breathing.mp. OR yale observation scale.mp. OR yale score.mp. OR yale scale.mp. OR nelson score.mp. OR nelson scale.mp. OR young infant observation scale.mp. OR exp TACHYCARDIA/OR fast heart rate.mp. OR capillary refill time.mp.	8723
C3	Lab tests	exp Diagnosis, Laboratory/	38,138
C4	Child	exp INFANT/OR exp CHILD/OR exp Adolescence/OR paediatric.mp. OR exp Pediatrics/OR child\$.mp. OR infant\$.mp. OR newborn\$.mp. OR bab\$.mp. OR neonat\$.mp. OR toddler\$.mp. OR exp Child, Preschool/	233,317
D4	Ambulatory care	exp Ambulatory Care/OR exp Family Practice/OR general practice.mp. OR GP.mp. OR exp Physicians, Family/OR exp Primary Health Care/OR exp Emergency Service/OR primary care.mp.	42,377
D5	Combination	C1 AND (C2 OR C3) AND C4 AND C5	86

Screening of titles and abstracts

Titles and abstracts will be screened by two independent reviewers, with discrepancies resolved by a third independent reviewer. The principal inclusion criterion will be a study on the predictive value of potential indicators for the diagnosis of serious infections in children; we will include systematic reviews and meta-analyses as well as primary studies. Serious infections will be defined as sepsis, pneumonia, meningitis, urinary tract infection, bacterial gastro-enteritis, cellulitis requiring hospital intervention, osteomyelitis and bronchiolitis requiring hospitalisation. Diagnostic indicators will be defined as any symptom, sign, test or other potential discriminator (e.g. doctor or parent opinion) used to predict or rule out the presence of illness.

We have drawn up the provisional reviewer guidelines below, which will be expanded and refined if studies come to light that are not easily included/excluded by the guideline:

Characteristic	Include if	Exclude if
Design	<ul style="list-style-type: none"> i. Cross-sectional study of immediate diagnostic accuracy ii. Longitudinal study of predictive accuracy iii Systematic reviews of above studies 	<ul style="list-style-type: none"> i. Case series of <50 children ii. Letters without research results iii. Narrative study or comment only iv. Therapy evaluation
Population	<ul style="list-style-type: none"> i. Includes children age 1 month-18 years (separately delineable) ii. Otherwise healthy 	<ul style="list-style-type: none"> i. out of age range ii. pre-existing illness
Setting	<ul style="list-style-type: none"> i. General practice/family medicine ii. Other Ambulatory care iii. Paediatric assessment unit iv. Pre-admission Emergency Care 	<ul style="list-style-type: none"> i. Post-admission secondary care ii Outside Europe, North-America, Australia/NZ
Outcome (i.e. serious infection)	<ul style="list-style-type: none"> i. Hospitalisation with presumed or confirmed serious infection (e.g. LRTI or pneumonia, meningitis, sepsis; osteomyelitis; complications of gastrointestinal or respiratory infection). ii. Specific infections diagnosed in a community setting: Pneumonia (with x-ray confirmation); other LRTI with quantified hypoxia; UTI (with microbiological confirmation). 	Diagnosis other than serious infection
Diagnostic procedures	<ul style="list-style-type: none"> i. History and presenting symptoms – e.g. fever, cough, vomiting, pallor, crying pattern, lethargy, irritability ii. Observation scales (e.g. McCarthy, Baby Check, Young infant observation scale) or triage scores (e.g. Manchester triage score) iii. Physical examination – e.g. vital signs, meningeal signs, capillary refill time iv. Near-patient tests – e.g. urine dipstick; influenza or RSV testing; CRP v. Rapid laboratory tests – e.g. WBC; inflammatory markers; urine microscopy 	<ul style="list-style-type: none"> i. Imaging ii. Invasive testing iii. In-lab microbiology other urine culture and microscopy. <p>NB These procedures may be used for confirmation of outcome</p>

The selection process will be piloted on a sample of 20 articles, and interobserver agreement will be calculated for the entire sample.

Quality assessment

Selected articles will be assessed on quality by using the QUADAS instrument.^{23,24} The QUADAS instrument has 11 core items and 9 additional items. The use of QUADAS is currently endorsed by the Cochrane Collaboration in the new handbook of diagnostic systematic reviews. Some of the co-applicants were involved in the writing of this new handbook, which is due to be published shortly. Not all items will be applicable in our review.

	QUADAS	Applicable
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Yes
2	Is the reference standard likely to correctly classify the target condition?	Yes
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Yes
4	Did the whole sample or a random selection of the sample receive verification using a reference standard or diagnosis?	Yes
5	Did patients receive the same reference standard regardless of the index test used?	Yes
6	Was the reference standard independent of the index test?	Not always, e.g. sepsis
7	Were the index test results interpreted without knowledge of the results of the reference standard?	Yes
8	Were the reference standard results interpreted without knowledge of the results of the index test?	Not always
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Yes, although clinical data are index tests
10	Were uninterpretable/intermediate test results reported?	Yes
11	Were withdrawals from the study explained?	Yes
	Additional items	
12	If a cut-off has been used, was it established before the study was started?	Not always (e.g. ROC analysis)
13	Is the technology of the index test likely to have changed since the study was carried out?	No, unless for lab tests
14	Did the study provide a clear definition of what was considered to be a 'positive' result?	Yes
15	Was treatment started after the index test was carried out but before the reference standard was performed?	Yes
16	Were data on observer variation reported?	Less applicable
17	Were data on instrument variation reported?	Not applicable
18	Were data presented for appropriate patient subgroups?	Yes
19	Was an appropriate sample size included?	Yes
20	Were objectives pre-specified?	Yes

Data extraction

The following data will be extracted from the included articles:

1. Design features, prospective or retrospective, consecutive patient inclusion.
2. The setting: emergency department, ambulatory care, in hospital or other.
3. The age and other patient characteristics.
4. The outcome and how that was defined (reference standard).
5. The index test, with details and cut-off used.
6. The number of participants and the prevalence of the outcome.
7. The results from the study, in sensitivity, specificity, positive or negative predictive value, odds ratios, area under curves (AUC) or *p*-values. Confidence intervals (CI) will be extracted where possible. When sufficient data are reported, 2×2 tables will be extracted.

The data will be extracted in duplicate by two independent researchers. If possible, authors will be contacted to supplement missing data.

Summarising the data

A. Study level meta-analysis

Depending on the nature of the available data, a meta-analysis will be performed. Diagnostic accuracy studies will be pooled using the bivariate method.^{25–27} The bivariate approach preserves the two-dimensional nature of the original data. Pairs of sensitivity and specificity are jointly analyzed, incorporating any correlation that might exist between these two measures using a random effects approach. This method has been shown to be equivalent to the hierarchical summary receiver operating characteristic (ROC) model which is considered the gold standard for diagnostic meta-analyses, but the results of the bivariate method are easier to interpret in clinical terms.^{28,29}

Before deciding to pool any studies, heterogeneity both in terms of clinical heterogeneity (by detailed study of the methods section of the paper and of protocol articles if available) and statistical heterogeneity (by calculating I^2) will be assessed. If possible, pooling will take the natural order of tests into account – in clinical practice, history and clinical examination are done before requesting laboratory tests.

B. Individual patient data meta-analysis

As stated above, in our provisional work we have identified five studies (with an aggregate population of about 10,000 patients) on unselected populations and have formal agreement to use the crude data for IPD analysis. These datasets are detailed below:

- Coventry, UK: 700 children presenting to hospital paediatric assessment unit/A&E with suspected acute infection (Dr Thompson).
- Oxfordshire & Somerset, UK: 2000 children presenting to general practice and out of hours centres with acute infection (Dr. M. Thompson);
- Nottingham, UK: 1700 children presenting to A&E with suspected acute infection (Dr Lakhanpaul)
- Netherlands: 3 datasets – 595 children presenting to emergency department with fever without source; 400 children with meningeal signs; 1787 children presenting to emergency department with fever (Dr H. Moll)
- Belgium: 4000 children presenting to primary care with acute infection (Dr Van Den Bruel)

These datasets will be complemented by the studies identified in the systematic review. Authors will be asked to contribute data in whatever format they prefer in order to facilitate contribution to the study. If further studies on unselected populations are identified we will attempt to

include them, although we recognise that it is usually impossible to retrieve individual patient data for studies published more than 20 years ago, and sometimes difficult to get agreement for release of data for more recent studies. Moreover, we will need to address the applicability of studies performed prior to vaccination for *Haemophilus influenzae*, *Streptococcus pneumoniae* and *Neisseria meningitidis* C. We will explore bias and generalizability by comparing the test characteristics (e.g. sensitivity, specificity) generated by the IPD analysis with the results reported by any studies we identify on unselected populations for which full data is unavailable.

In conducting the IPD analysis we will take a two-stage approach, generating diagnostic algorithms using logistic regression (to generate odds ratios) in stage 1 and then assessing the predictive value and ROC characteristics of these algorithms in stage 2. In order to validate the prediction rule we will need to address two issues: a) over-optimistic estimates and b) transferability of the prediction rule across settings. To achieve this we will use k-fold cross-validation to obtain more realistic estimates and calibration using other datasets to test the transferability of the model. By validating the clinical prediction rules on patients in broader settings (and thus different disease prevalence and spectrum) from those used to derive the rule, we will be able to demonstrate the generalizability or external validity of the rule (McGinn *et al.*, JAMA 2000). We anticipate that this process will require model revision and/or shrinkage methods (Steyerberg EW *et al.*, Statist Med 2004).

As with the standard meta-analyses, to decide whether pooling of data for analysis is justified we will assess heterogeneity between studies using I^2 , which describes the percentage of variation between studies due to heterogeneity rather than chance. The range for I^2 lies between 0% (i.e., no observed heterogeneity) and 100%; we will pool if I^2 is lower than 25% ($p > 0.30$).

In stage 1 a two-level multilevel regression model will be fitted for the diagnostic variables of interest, with patients corresponding to level one units and individual study as level two units. This will generate odds ratios for the likelihood of the main outcome (serious infection). Study effects will be represented by fixed effects, whilst patient effects will be represented by random effects. The diagnostic factors included in the analysis will be used as covariates. We will use a binary dummy variable to identify each study within the regression analysis.

To reduce bias and to increase statistical efficiency, we will impute missing data using the linear regression method (multivariate analyses) available in Spss (version 12.0). Regression will be based on the correlation between individual variables with missing values and all other variables, as estimated from the complete set of data. We will impute missing values only within individual studies.

In stage 2 we will construct a number of diagnostic algorithms, using the odds ratios for individual diagnostic markers derived in stage 1, and calculate their sensitivity, predictive value when applied to the second half of the dataset. To conduct sensitivity analyses, we will also report these results applied to each dataset separately. Confidence intervals around these test characteristics will be reported with 95% confidence intervals based on the standard error of a proportion. Where appropriate we will develop two-level staged algorithms (e.g. undertaking a diagnostic test being dependent on presenting symptoms and signs) and present ROC curves where the algorithm includes a diagnostic test or marker generating a continuous variable (e.g. % oxygen saturation).

Subgroup analyses will be attempted. Subgroups will be based on patient age, i.e. children under the age of 1 year, children between 1 and 4 years, children between 5 and 12 years, and adolescents. Another subgroup is based on setting, reflecting increasing prevalence of serious infections: general practice – urgent access primary care – paediatric assessment unit

– emergency department. A final category will be based on outcome. Generating additional separate algorithms for sepsis/meningitis and pneumonia would be desirable, as the first outcome requires immediate action, and the second outcome is the most prevalent serious infection in children in primary and secondary care.

Research Governance

The University of Oxford will be the nominated sponsor for this study.

4. Project timetable and milestones

The following will be the key milestones for the study:

Milestone	Start date	Completion date
Perform literature search	Month 1	Month 1
Obtain data from 5 existing datasets	Month 1	Month 4
Screen titles and abstracts	Month 2	Month 2
Quality grading of included studies	Month 3	Month 4
Data extraction from included studies	Month 4	Month 6
Summarising data, meta-analysis	Month 6	Month 9
Individual patient data meta-analysis	Month 6	Month 10
Writing final report and submitting for publication	Month 10	Month 12

5. Expertise

The research team that has been assembled for this project brings together methodological expertise in systematic reviewing, diagnostic test systematic reviewing, individual patient data meta analysis, as well as considerable clinical expertise in both emergency departments and primary care settings. Moreover it draws on this expertise not only from the UK, but also from Belgium and the Netherlands.

Dr Thompson is a Clinical Lecturer in Primary Care and half-time Principal in General Practice who also works regularly in an out of hours GP surgery. He has performed several research studies examining clinical predictors of serious infections in primary care and paediatric assessment units. These have included prospective studies of predictive value of vital signs, severity of illness scores and inflammatory markers in children a paediatric assessment unit. He has also published on the early signs of meningococcal disease in children. His systematic review experience includes the treatment of common upper respiratory tract infections with steroids, and he is also currently a member of two National Institute for Health & Clinical Excellence (NICE) guideline development groups (prescribing antibiotics for upper respiratory tract infections, and diagnosis and treatment of meningitis). In addition to his clinical experience and expertise in diagnostic studies he will be responsible for coordinating the proposed study and will supervise the staff funded by this grant.

Professor David Mant is the head of the Department of Primary Health Care at the University of Oxford and has an international reputation in primary care research. His research has included numerous seminal studies on childhood infections and cardiovascular disease. He was the PI of the MRC-funded Oxford Childhood infection study from 2001–6, and sits on national committees such as the Standing Advisory Committee on Antibiotic Resistance (paediatric sub-group) 2005–7 and National Expert Panel on New and Emerging Infections (2003–7). He

will contribute extensive methodological input on the study design and meta-analysis, and will provide direct support to Dr Thompson.

Dr Glasziou has extensive expertise in conducting systematic reviews and individual patient data meta-analyses. He has published several textbooks on systematic reviewing, and authored numerous systematic reviews. He is a member of the Cochrane Collaborations diagnostics sub-group and currently Professor of Evidence Based Medicine at the University of Oxford. He is also a practising GP in Oxford. He will contribute extensively to the methodology of the systematic review and IPD meta-analysis.

Dr Van den Bruel has done research in the area of serious infections in children for the last 6 years. She has performed several studies, including one in which a clinical prediction rule for the exclusion of serious infections was developed. In addition, she has been working at the Belgian HTA agency for the last 4 years, where she is responsible for the evaluation of diagnostic tests and the methodology of systematic reviews. Previously she worked as a GP for seven years. In addition to sharing her dataset of children with acute infections, she will contribute expertise in systematic reviews of diagnostic studies.

Dr Moll is Head of the Paediatric Emergency Department of the Sophia's Children's Hospital –ErasmusMC in Rotterdam. Her research has focussed on emergency department triage, and the development and validation of prediction rules for acute paediatric infections. In particular she has performed studies on meningitis, fever/serious bacterial infections, RSV and pneumonia in the emergency department setting. She will contribute not only her dataset to this study, but also her clinical experience in emergency paediatrics, and research experience in diagnostic studies and validation of prediction rules.

Professor Buntinx has been working as a GP in Belgium for 32 years and as a researcher and Professor at the Departments of General Practice of the University of Maastricht (Netherlands) and Leuven (Belgium) since 1989. From the start of his research career, he has been focusing on the methodology and execution of diagnostic studies, including multivariate analyses and diagnostic meta-analysis. He has published some 175 papers in international peer reviewed journals and almost as many in Dutch language peer reviewed journals. He currently serves as the research director of the Department of General Practice in Leuven and as the founding president of the Belgian Centre of Evidence-based Medicine. With Prof. Knottnerus he also is the co-editor of the new edition of 'The evidence base of clinical diagnosis' (Blackwells, in press). In 2006, he was elected member of the Belgian Royal Academy of Medicine. He will contribute to the methodological input on diagnostic studies and systematic review of diagnostic studies.

Professor Bert Aertgeerts is a GP in Belgium. He has done research on screening for alcohol abuse and dependence in different settings, and has conducted several systematic diagnostic reviews on various clinical topics. He was also responsible for the European First Aid Manual, led by Stijn Vandeveld from the Red Cross Flanders. He is the director (2001) of the Centre of Evidence-Based Medicine (Belgian Branch of the Dutch Cochrane Collaboration) and is currently head of the department of General Practice at the Katholieke Universiteit Leuven. In 2006, he was elected member of the Belgian Royal Academy of medicine. He will contribute to the systematic review methodology, particularly in relation to diagnostic studies.

Professor Geert-Jan Dinant is vice-chair of the Department of General Practice at the University of Maastricht in the Netherlands. He has extensive research experience in performing diagnostic studies in primary care on pneumonia and osteoporosis and has authored textbooks on evidence based clinical diagnosis. He will bring methodological expertise in diagnostic studies to the research team.

Dr Shelly Segal is a Consultant Paediatrician with special interest in Infectious disease and is the Clinical lead in the Paediatric Emergency department at the John Radcliffe Hospital in Oxford. In addition to her paediatric emergency experience, she has performed several studies on the genetic susceptibility to infectious diseases in children, particularly invasive pneumococcal disease. She will contribute her paediatric emergency clinical experience to the research team.

Dr Monica Lakhanpaul is a Consultant Paediatrician and Co-Director for National Collaborating Centre for Women's Health and Children's Health of the Royal College of Paediatrics and Child Health which has been responsible for undertaking several NICE reviews, in particular the recent guideline on management of the Feverish child. She is also clinical lead for developing nurse-led urgent care services in Leicester. She has recently completed a prospective study of children attending A&E in Nottingham. In addition to sharing dataset for the IPD meta-analysis, she will contribute clinical paediatric experience, and experience of literature review to the project team.

Dr Rafael Perera is a University Lecturer in Statistics and Director of Research Methodologies at the Centre for Evidence-Based Medicine. He has extensive experience conducting systematic reviews and IPD meta-analysis and has published numerous systematic reviews. He will contribute extensively to the data analysis for this study.

6. Service users

The research team has extensive current experience as front-line service clinicians in the provision of clinical care to children in emergency and urgent primary care settings. Dr Segal is in charge of paediatric emergency medicine at the John Radcliffe Hospital in Oxford, Dr Moll is Head of Paediatric Emergency Department at Sophia's Children Hospital in Rotterdam, and Dr Lakhanpaul is a Community Paediatrician in Leicester. Professors Mant, Glasziou, Dinant, Aertgeerts, Buntinx, and Drs Thompson and Van den Bruel are all general practitioners who have worked in general practice in England, the Netherlands, and Belgium. Dr Thompson also works in an out of hours GP centre in Oxford. We will also gather the input of parents/carers input in order to assess the likely impact of this rule in the real world setting, and to ensure that the predictors we identify (e.g. vital sign measurements, possibly blood tests) are acceptable to most parents/carers. We will therefore assemble a group of parents who have had personal experience with children in emergency care or urgent access primary care and obtain their input on the final prediction rules.

7. Justification of support required

The main support required for this project are salary support for a data manager, salary support for a statistician, and reimbursement for meetings of the research team. The data manager will be responsible for the retrieval of articles identified in the systematic review, obtaining

and assembling the datasets required for the individual patient data meta analysis, as well as organising meetings of the research team. The data manager may be expected to undertake simple descriptive analysis of the datasets. Salary support has been requested for Dr Perera to undertake statistical work for this study. Dr Perera will be responsible for the summarising of the studies identified in the systematic review. He will also undertake the individual patient data meta-analysis. Dr Van den Bruel will be reimbursed for undertaking one part of the analysis, and will invoice the University of Oxford for a specified component of this work. In order to take advantage of the considerable expertise of the collaborators who have agreed to work on this project, reimbursement for attending four research team meetings during the study period will be provided to Professor Buntinx, Professor Aertgeerts, Professor Dinant, Dr Van den Bruel, Dr Moll, Dr Lakhanpaul and Dr Segal. Nominal salary support has been requested for Dr Lakhanpaul, Professor Mant, Professor Glasziou, Dr Segal and Dr Thompson. No salary support has been requested for Professor Buntinx, Professor Aertgeerts, Professor Dinant, or Dr Moll.

In order to minimise impact on the environment the Dutch and Belgian collaborators on this study will attend meetings in England by rail where possible rather than flying. This project also seeks to reduce the need and costs associated with further prospective studies of predictors of serious infection in children, by taking advantage of literature that has already been published on this topic, and by using individual patient data meta-analysis of studies that have already been carried out.

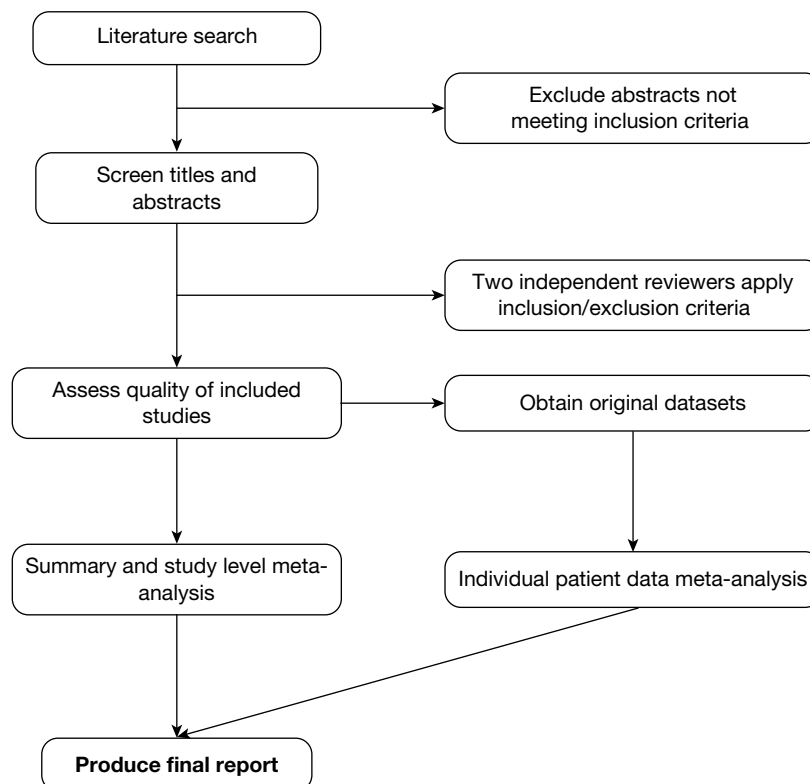
8. References

1. Stiell I, Wells G, Laupacis A, Brison R, Verbeek R, Vandemheen K, *et al.* Multicentre trial to introduce the Ottawa ankle rules for use of radiography in acute ankle injuries. Multicentre Ankle Rule Study Group. *BMJ* 1995;**311**(7005):594–7.
2. Armon K, Stephenson T, Gabriel V, MacFaul R, Eccleston P, Werneke U, *et al.* Determining the common medical presenting problems to an accident and emergency department. *Arch Dis Child* 2001;**84**(5):390–2.
3. Avery AJ, Groom L, Boot D, Earwicker S, Carlisle R. What problems do patients present with outside normal general practice surgery hours? A prospective study of the use of general practice and accident and emergency services. *J Public Health Med* 1999;**21**(1):88–94.
4. Brogan C, Pickard D, Gray A, Fairman S, Hill A. The use of out of hours health services: a cross sectional survey. *BMJ* 1998;**316**(7130):524–7.
5. Munro J, Nicholl J, O’Cathain A, Knowles E. Impact of NHS direct on demand for immediate care: observational study. *BMJ* 2000;**321**(7254):150–3.
6. 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.
7. Mackway-Jones K. *Emergency triage: Manchester Triage Group*: BMJ Publishing Group, 1996.
8. Cooke MW, Jinks S. Does the Manchester triage system detect the critically ill? *J Accid Emerg Med* 1999;**16**(3):179–81.
9. Roukema J, Steyerberg EW, van Meurs A, Ruige M, van der Lei J, Moll HA. Validity of the Manchester Triage System in paediatric emergency care. *Emerg Med J* 2006;**23**(12):906–10.

10. Baumann MR, Strout TD. Evaluation of the Emergency Severity Index (version 3) triage algorithm in pediatric patients. *Acad Emerg Med* 2005;**12**(3):219–24.
11. Tanabe P, Gimbel R, Yarnold PR, Adams JG. The Emergency Severity Index (version 3) 5-level triage system scores predict ED resource consumption. *J Emerg Nurs* 2004;**30**(1):22–9.
12. Gouin S, Gravel J, Amre DK, Bergeron S. Evaluation of the Paediatric Canadian Triage and Acuity Scale in a pediatric ED. *Am J Emerg Med* 2005;**23**(3):243–7.
13. Chamberlain JM, Patel KM, Pollack MM. The Pediatric Risk of Hospital Admission score: a second-generation severity-of-illness score for pediatric emergency patients. *Pediatrics* 2005;**115**(2):388–95.
14. Gorelick MH, Lee C, Cronan K, Kost S, Palmer K. Pediatric emergency assessment tool (PEAT): a risk-adjustment measure for pediatric emergency patients. *Acad Emerg Med* 2001;**8**(2):156–62.
15. Nelson KG. An index of severity for acute paediatric illness. *Am J Pub Health* 1980;**70**:804–7.
16. 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**(5):802–9.
17. Hewson P, Poulakis Z, Jarman F, Kerr J, McMaster D, Goodge J, *et al.* Clinical markers of serious illness in young infants: a multicentre follow-up study. *J Paediatr Child Health* 2000;**36**(3):221–5.
18. Hewson PH, Gollan RA. A simple hospital triaging system for infants with acute illness. *J Paediatr Child Health* 1995;**31**(1):29–32.
19. NICE. National Institute for Clinical Excellence: Feversh illness in children – assessment and initial management in children younger than 5 years, 2007.
20. 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**:538–546.
21. Bleeker SE, Derksen-Lubsen G, Grobbee DE, Donders AR, Moons KG, Moll HA. Validating and updating a prediction rule for serious bacterial infection in patients with fever without source. *Acta Paediatr* 2007;**96**(1):100–4.
22. Oostenbrink R, Moons KG, Derksen-Lubsen AG, Grobbee DE, Moll HA. A diagnostic decision rule for management of children with meningeal signs. *Eur J Epidemiol* 2004;**19**(2):109–16.
23. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:25.
24. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. *BMC Med Res Methodol* 2006;**6**:9.
25. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med* 2002;**21**(4):589–624.
26. Van Houwelingen HC, Zwinderman KH, Stijnen T. A bivariate approach to meta-analysis. *Stat Med* 1993;**12**(24):2273–84.

27. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol* 2005;**58**(10):982–90.
28. Rutter CM, Gatsonis CA. A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. *Stat Med* 2001;**20**(19):2865–84.
29. Harbord RM, Deeks JJ, Egger M, Whiting P, Sterne JA. A unification of models for meta-analysis of diagnostic accuracy studies. *Biostatistics* 2007;**8**(2):239–51.

9. Flow diagram



Health Technology Assessment programme

Director,
Professor Tom Walley, CBE,
 Director, NIHR HTA programme,
 Professor of Clinical Pharmacology,
 University of Liverpool

Deputy Director,
Professor Hywel Williams,
 Professor of Dermato-Epidemiology,
 Centre of Evidence-Based Dermatology,
 University of Nottingham

Prioritisation Group

Members

Chair,
Professor Tom Walley, CBE,
 Director, NIHR HTA
 programme, Professor of Clinical
 Pharmacology, University of
 Liverpool

Professor Imti Choonara,
 Professor in Child Health,
 Academic Division of Child
 Health, University of Nottingham
 Chair – Pharmaceuticals Panel

Dr Bob Coates,
 Consultant Advisor – Disease
 Prevention Panel

Dr Andrew Cook,
 Consultant Advisor – Intervention
 Procedures Panel

Dr Peter Davidson,
 Director of NETSCC, Health
 Technology Assessment

Dr Nick Hicks,
 Consultant Adviser – Diagnostic
 Technologies and Screening Panel,
 Consultant Advisor–Psychological
 and Community Therapies Panel

Ms Susan Hird,
 Consultant Advisor, External
 Devices and Physical Therapies
 Panel

Professor Sallie Lamb,
 Director, Warwick Clinical Trials
 Unit, Warwick Medical School,
 University of Warwick
 Chair – HTA Clinical Evaluation
 and Trials Board

Professor Jonathan Michaels,
 Professor of Vascular Surgery,
 Sheffield Vascular Institute,
 University of Sheffield
 Chair – Interventional Procedures
 Panel

Professor Ruairidh Milne,
 Director – External Relations

Dr John Pounsford,
 Consultant Physician, Directorate
 of Medical Services, North Bristol
 NHS Trust
 Chair – External Devices and
 Physical Therapies Panel

Dr Vaughan Thomas,
 Consultant Advisor –
 Pharmaceuticals Panel, Clinical
 Lead – Clinical Evaluation Trials
 Prioritisation Group

Professor Margaret Thorogood,
 Professor of Epidemiology, Health
 Sciences Research Institute,
 University of Warwick
 Chair – Disease Prevention Panel

Professor Lindsay Turnbull,
 Professor of Radiology, Centre for
 the MR Investigations, University
 of Hull
 Chair – Diagnostic Technologies
 and Screening Panel

Professor Scott Weich,
 Professor of Psychiatry, Health
 Sciences Research Institute,
 University of Warwick
 Chair – Psychological and
 Community Therapies Panel

Professor Hywel Williams,
 Director of Nottingham Clinical
 Trials Unit, Centre of Evidence-
 Based Dermatology, University of
 Nottingham
 Chair – HTA Commissioning
 Board
 Deputy HTA Programme Director

HTA Commissioning Board

Chair,
Professor Hywel Williams,
 Professor of Dermato-Epidemiology,
 Centre of Evidence-Based Dermatology,
 University of Nottingham

Deputy Chair,
Professor Jon Deeks,
 Department of Public Health and
 Epidemiology,
 University of Birmingham

Professor Tom Walley, CBE,
 Professor of Clinical Pharmacology,
 Director, NIHR HTA programme,
 University of Liverpool

Members

Professor Ann Ashburn,
 Professor of Rehabilitation and
 Head of Research, Southampton
 General Hospital

Professor Judith Bliss,
 Director of ICR-Clinical Trials
 and Statistics Unit, The Institute of
 Cancer Research

Professor Peter Brocklehurst,
 Professor of Women's Health,
 Institute for Women's Health,
 University College London

Professor David Fitzmaurice,
 Professor of Primary Care
 Research, Department of Primary
 Care Clinical Sciences, University
 of Birmingham

Professor John W Gregory,
 Professor in Paediatric
 Endocrinology, Department of
 Child Health, Wales School of
 Medicine, Cardiff University

Professor Steve Halligan,
 Professor of Gastrointestinal
 Radiology, University College
 Hospital, London

Professor Angela Harden,
 Professor of Community and
 Family Health, Institute for
 Health and Human Development,
 University of East London

Dr Martin J Landray,
 Reader in Epidemiology, Honorary
 Consultant Physician, Clinical
 Trial Service Unit, University of
 Oxford

Dr Joanne Lord,
 Reader, Health Economics
 Research Group, Brunel University

Professor Stephen Morris,
 Professor of Health Economics,
 University College London,
 Research Department of
 Epidemiology and Public Health,
 University College London

Professor Dion Morton,
 Professor of Surgery, Academic
 Department of Surgery, University
 of Birmingham

Professor Gail Mountain,
 Professor of Health Services
 Research, Rehabilitation and
 Assistive Technologies Group,
 University of Sheffield

Professor Irwin Nazareth,
 Professor of Primary Care and
 Head of Department, Department
 of Primary Care and Population
 Sciences, University College
 London

Professor E Andrea Nelson,
 Professor of Wound Healing and
 Director of Research, School of
 Healthcare, University of Leeds

Professor John David Norrie,
 Chair in Clinical Trials and
 Biostatistics, Robertson Centre for
 Biostatistics, University of Glasgow

Dr Rafael Perera,
 Lecturer in Medical Statistics,
 Department of Primary Health
 Care, University of Oxford

HTA Commissioning Board (continued)

Professor Barney Reeves,
Professorial Research Fellow
in Health Services Research,
Department of Clinical Science,
University of Bristol

Professor Peter Tyrer,
Professor of Community
Psychiatry, Centre for Mental
Health, Imperial College London

Professor Martin Underwood,
Professor of Primary Care
Research, Warwick Medical
School, University of Warwick

Professor Caroline Watkins,
Professor of Stroke and Older
People's Care, Chair of UK
Forum for Stroke Training, Stroke
Practice Research Unit, University
of Central Lancashire

Dr Duncan Young,
Senior Clinical Lecturer and
Consultant, Nuffield Department
of Anaesthetics, University of
Oxford

Observers

Dr Tom Foulks,
Medical Research Council

Dr Kay Pattison,
Senior NIHR Programme
Manager, Department of Health

HTA Clinical Evaluation and Trials Board

Chair,

Professor Sallie Lamb,
Director,
Warwick Clinical Trials Unit,
Warwick Medical School,
University of Warwick and Professor of
Rehabilitation,
Nuffield Department of Orthopaedic,
Rheumatology and Musculoskeletal Sciences,
University of Oxford

Deputy Chair,

Professor Jenny Hewison,
Professor of the Psychology of Health Care,
Leeds Institute of Health Sciences,
University of Leeds

Programme Director,

Professor Tom Walley, CBE,
Director, NIHR HTA programme,
Professor of Clinical Pharmacology,
University of Liverpool

Members

Professor Keith Abrams,
Professor of Medical Statistics,
Department of Health Sciences,
University of Leicester

Professor Martin Bland,
Professor of Health Statistics,
Department of Health Sciences,
University of York

Professor Jane Blazeby,
Professor of Surgery and
Consultant Upper GI Surgeon,
Department of Social Medicine,
University of Bristol

Professor Julia M Brown,
Director, Clinical Trials Research
Unit, University of Leeds

Professor Alistair Burns,
Professor of Old Age Psychiatry,
Psychiatry Research Group, School
of Community-Based Medicine,
The University of Manchester &
National Clinical Director for
Dementia, Department of Health

Dr Jennifer Burr,
Director, Centre for Healthcare
Randomised trials (CHART),
University of Aberdeen

Professor Linda Davies,
Professor of Health Economics,
Health Sciences Research Group,
University of Manchester

Professor Simon Gilbody,
Prof of Psych Medicine and Health
Services Research, Department of
Health Sciences, University of York

Professor Steven Goodacre,
Professor and Consultant in
Emergency Medicine, School of
Health and Related Research,
University of Sheffield

Professor Dyfrig Hughes,
Professor of Pharmacoeconomics,
Centre for Economics and Policy
in Health, Institute of Medical
and Social Care Research, Bangor
University

Professor Paul Jones,
Professor of Respiratory Medicine,
Department of Cardiac and
Vascular Science, St George's
Hospital Medical School,
University of London

Professor Khalid Khan,
Professor of Women's Health and
Clinical Epidemiology, Barts and
the London School of Medicine,
Queen Mary, University of London

Professor Richard J McManus,
Professor of Primary Care
Cardiovascular Research, Primary
Care Clinical Sciences Building,
University of Birmingham

Professor Helen Rodgers,
Professor of Stroke Care, Institute
for Ageing and Health, Newcastle
University

Professor Ken Stein,
Professor of Public Health,
Peninsula Technology Assessment
Group, Peninsula College
of Medicine and Dentistry,
Universities of Exeter and
Plymouth

Professor Jonathan Sterne,
Professor of Medical Statistics
and Epidemiology, Department
of Social Medicine, University of
Bristol

Mr Andy Vail,
Senior Lecturer, Health Sciences
Research Group, University of
Manchester

Professor Clare Wilkinson,
Professor of General Practice and
Director of Research North Wales
Clinical School, Department of
Primary Care and Public Health,
Cardiff University

Dr Ian B Wilkinson,
Senior Lecturer and Honorary
Consultant, Clinical Pharmacology
Unit, Department of Medicine,
University of Cambridge

Observers

Ms Kate Law,
Director of Clinical Trials,
Cancer Research UK

Dr Morven Roberts,
Clinical Trials Manager, Health
Services and Public Health
Services Board, Medical Research
Council

Diagnostic Technologies and Screening Panel

Members

<p>Chair, Professor Lindsay Wilson Turnbull, Scientific Director of the Centre for Magnetic Resonance Investigations and YCR Professor of Radiology, Hull Royal Infirmary</p> <p>Professor Judith E Adams, Consultant Radiologist, Manchester Royal Infirmary, Central Manchester & Manchester Children's University Hospitals NHS Trust, and Professor of Diagnostic Radiology, University of Manchester</p> <p>Mr Angus S Arunkalaivanan, Honorary Senior Lecturer, University of Birmingham and Consultant Urogynaecologist and Obstetrician, City Hospital, Birmingham</p> <p>Dr Diana Baralle, Consultant and Senior Lecturer in Clinical Genetics, University of Southampton</p>	<p>Dr Stephanie Dancer, Consultant Microbiologist, Hairmyres Hospital, East Kilbride</p> <p>Dr Diane Eccles, Professor of Cancer Genetics, Wessex Clinical Genetics Service, Princess Anne Hospital</p> <p>Dr Trevor Friedman, Consultant Liaison Psychiatrist, Brandon Unit, Leicester General Hospital</p> <p>Dr Ron Gray, Consultant, National Perinatal Epidemiology Unit, Institute of Health Sciences, University of Oxford</p> <p>Professor Paul D Griffiths, Professor of Radiology, Academic Unit of Radiology, University of Sheffield</p> <p>Mr Martin Hooper, Public contributor</p>	<p>Professor Anthony Robert Kendrick, Associate Dean for Clinical Research and Professor of Primary Medical Care, University of Southampton</p> <p>Dr Nicola Lennard, Senior Medical Officer, MHRA</p> <p>Dr Anne Mackie, Director of Programmes, UK National Screening Committee, London</p> <p>Mr David Mathew, Public contributor</p> <p>Dr Michael Millar, Consultant Senior Lecturer in Microbiology, Department of Pathology & Microbiology, Barts and The London NHS Trust, Royal London Hospital</p> <p>Mrs Una Rennard, Public contributor</p>	<p>Dr Stuart Smellie, Consultant in Clinical Pathology, Bishop Auckland General Hospital</p> <p>Ms Jane Smith, Consultant Ultrasound Practitioner, Leeds Teaching Hospital NHS Trust, Leeds</p> <p>Dr Allison Streetly, Programme Director, NHS Sickle Cell and Thalassaemia Screening Programme, King's College School of Medicine</p> <p>Dr Matthew Thompson, Senior Clinical Scientist and GP, Department of Primary Health Care, University of Oxford</p> <p>Dr Alan J Williams, Consultant Physician, General and Respiratory Medicine, The Royal Bournemouth Hospital</p>
--	--	---	---

Observers

<p>Dr Tim Elliott, Team Leader, Cancer Screening, Department of Health</p> <p>Dr Joanna Jenkinson, Board Secretary, Neurosciences and Mental Health Board (NMHB), Medical Research Council</p>	<p>Professor Julietta Patnick, Director, NHS Cancer Screening Programme, Sheffield</p> <p>Dr Kay Pattison, Senior NIHR Programme Manager, Department of Health</p>	<p>Professor Tom Walley, CBE, Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool</p>	<p>Dr Ursula Wells, Principal Research Officer, Policy Research Programme, Department of Health</p>
--	--	---	---

Disease Prevention Panel

Members

<p>Chair, Professor Margaret Thorogood, Professor of Epidemiology, University of Warwick Medical School, Coventry</p> <p>Dr Robert Cook, Clinical Programmes Director, Bazian Ltd, London</p> <p>Dr Colin Greaves, Senior Research Fellow, Peninsula Medical School (Primary Care)</p> <p>Mr Michael Head, Public contributor</p>	<p>Professor Cathy Jackson, Professor of Primary Care Medicine, Bute Medical School, University of St Andrews</p> <p>Dr Russell Jago, Senior Lecturer in Exercise, Nutrition and Health, Centre for Sport, Exercise and Health, University of Bristol</p> <p>Dr Julie Mytton, Consultant in Child Public Health, NHS Bristol</p>	<p>Professor Irwin Nazareth, Professor of Primary Care and Director, Department of Primary Care and Population Sciences, University College London</p> <p>Dr Richard Richards, Assistant Director of Public Health, Derbyshire County Primary Care Trust</p> <p>Professor Ian Roberts, Professor of Epidemiology and Public Health, London School of Hygiene & Tropical Medicine</p>	<p>Dr Kenneth Robertson, Consultant Paediatrician, Royal Hospital for Sick Children, Glasgow</p> <p>Dr Catherine Swann, Associate Director, Centre for Public Health Excellence, NICE</p> <p>Mrs Jean Thurston, Public contributor</p> <p>Professor David Weller, Head, School of Clinical Science and Community Health, University of Edinburgh</p>
---	--	--	--

Observers

<p>Ms Christine McGuire, Research & Development, Department of Health</p>	<p>Dr Kay Pattison, Senior NIHR Programme Manager, Department of Health</p>	<p>Professor Tom Walley, CBE, Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool</p>
---	---	---

External Devices and Physical Therapies Panel

Members

Chair, Dr John Pounsford, Consultant Physician North Bristol NHS Trust	Dr Dawn Carnes, Senior Research Fellow, Barts and the London School of Medicine and Dentistry	Dr Shaheen Hamdy, Clinical Senior Lecturer and Consultant Physician, University of Manchester	Mr Jim Reece, Public contributor
Deputy Chair, Professor E Andrea Nelson, Reader in Wound Healing and Director of Research, University of Leeds	Dr Emma Clark, Clinician Scientist Fellow & Cons. Rheumatologist, University of Bristol	Professor Christine Norton, Professor of Clinical Nursing Innovation, Bucks New University and Imperial College Healthcare NHS Trust	Professor Maria Stokes, Professor of Neuromusculoskeletal Rehabilitation, University of Southampton
Professor Bipin Bhakta, Charterhouse Professor in Rehabilitation Medicine, University of Leeds	Mrs Anthea De Barton-Watson, Public contributor	Dr Lorraine Pinnigton, Associate Professor in Rehabilitation, University of Nottingham	Dr Pippa Tyrrell, Senior Lecturer/Consultant, Salford Royal Foundation Hospitals' Trust and University of Manchester
Mrs Penny Calder, Public contributor	Professor Nadine Foster, Professor of Musculoskeletal Health in Primary Care Arthritis Research, Keele University	Dr Kate Radford, Senior Lecturer (Research), University of Central Lancashire	Dr Nefyn Williams, Clinical Senior Lecturer, Cardiff University

Observers

Dr Kay Pattison, Senior NIHR Programme Manager, Department of Health	Dr Morven Roberts, Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council	Professor Tom Walley, CBE, Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool	Dr Ursula Wells, Principal Research Officer, Policy Research Programme, Department of Health
--	--	--	---

Interventional Procedures Panel

Members

Chair, Professor Jonathan Michaels, Professor of Vascular Surgery, University of Sheffield	Mr Seumas Eckford, Consultant in Obstetrics & Gynaecology, North Devon District Hospital	Dr Fiona Lecky, Senior Lecturer/Honorary Consultant in Emergency Medicine, University of Manchester/Salford Royal Hospitals NHS Foundation Trust	Professor Jon Moss, Consultant Interventional Radiologist, North Glasgow Hospitals University NHS Trust
Deputy Chair, Mr Michael Thomas, Consultant Colorectal Surgeon, Bristol Royal Infirmary	Professor Sam Eljamel, Consultant Neurosurgeon, Ninewells Hospital and Medical School, Dundee	Dr Nadim Malik, Consultant Cardiologist/Honorary Lecturer, University of Manchester	Dr Simon Padley, Consultant Radiologist, Chelsea & Westminster Hospital
Mrs Isabel Boyer, Public contributor	Dr Adele Fielding, Senior Lecturer and Honorary Consultant in Haematology, University College London Medical School	Mr Hisham Mehanna, Consultant & Honorary Associate Professor, University Hospitals Coventry & Warwickshire NHS Trust	Dr Ashish Paul, Medical Director, Bedfordshire PCT
Mr Sankaran Chandra Sekharan, Consultant Surgeon, Breast Surgery, Colchester Hospital University NHS Foundation Trust	Dr Matthew Hatton, Consultant in Clinical Oncology, Sheffield Teaching Hospital Foundation Trust	Dr Jane Montgomery, Consultant in Anaesthetics and Critical Care, South Devon Healthcare NHS Foundation Trust	Dr Sarah Purdy, Consultant Senior Lecturer, University of Bristol
Professor Nicholas Clarke, Consultant Orthopaedic Surgeon, Southampton University Hospitals NHS Trust	Dr John Holden, General Practitioner, Garswood Surgery, Wigan		Dr Matthew Wilson, Consultant Anaesthetist, Sheffield Teaching Hospitals NHS Foundation Trust
Ms Leonie Cooke, Public contributor			Professor Yit Chiun Yang, Consultant Ophthalmologist, Royal Wolverhampton Hospitals NHS Trust

Observers

Dr Kay Pattison, Senior NIHR Programme Manager, Department of Health	Dr Morven Roberts, Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council	Professor Tom Walley, CBE, Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool	Dr Ursula Wells, Principal Research Officer, Policy Research Programme, Department of Health
--	--	--	---

Pharmaceuticals Panel

Members

<p>Chair, Professor Imti Choonara, Professor in Child Health, University of Nottingham</p> <p>Deputy Chair, Dr Yoon K Loke, Senior Lecturer in Clinical Pharmacology, University of East Anglia</p> <p>Dr Martin Ashton-Key, Medical Advisor, National Commissioning Group, NHS London</p> <p>Dr Peter Elton, Director of Public Health, Bury Primary Care Trust</p> <p>Dr Ben Goldacre, Research Fellow, Epidemiology London School of Hygiene and Tropical Medicine</p>	<p>Dr James Gray, Consultant Microbiologist, Department of Microbiology, Birmingham Children's Hospital NHS Foundation Trust</p> <p>Dr Jurjees Hasan, Consultant in Medical Oncology, The Christie, Manchester</p> <p>Dr Carl Heneghan, Deputy Director Centre for Evidence-Based Medicine and Clinical Lecturer, Department of Primary Health Care, University of Oxford</p> <p>Dr Dyfrig Hughes, Reader in Pharmacoeconomics and Deputy Director, Centre for Economics and Policy in Health, IMSCaR, Bangor University</p>	<p>Dr Maria Kouimtzi, Pharmacy and Informatics Director, Global Clinical Solutions, Wiley-Blackwell</p> <p>Professor Femi Oyeboode, Consultant Psychiatrist and Head of Department, University of Birmingham</p> <p>Dr Andrew Prentice, Senior Lecturer and Consultant Obstetrician and Gynaecologist, The Rosie Hospital, University of Cambridge</p> <p>Ms Amanda Roberts, Public contributor</p> <p>Dr Gillian Shepherd, Director, Health and Clinical Excellence, Merck Serono Ltd</p>	<p>Mrs Katrina Simister, Assistant Director New Medicines, National Prescribing Centre, Liverpool</p> <p>Professor Donald Singer, Professor of Clinical Pharmacology and Therapeutics, Clinical Sciences Research Institute, CSB, University of Warwick Medical School</p> <p>Mr David Symes, Public contributor</p> <p>Dr Arnold Zermansky, General Practitioner, Senior Research Fellow, Pharmacy Practice and Medicines Management Group, Leeds University</p>
---	--	--	---

Observers

<p>Dr Kay Pattison, Senior NIHR Programme Manager, Department of Health</p> <p>Mr Simon Reeve, Head of Clinical and Cost- Effectiveness, Medicines, Pharmacy and Industry Group, Department of Health</p>	<p>Dr Heike Weber, Programme Manager, Medical Research Council</p> <p>Professor Tom Walley, CBE, Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool</p>	<p>Dr Ursula Wells, Principal Research Officer, Policy Research Programme, Department of Health</p>
---	--	---

Psychological and Community Therapies Panel

Members

<p>Chair, Professor Scott Weich, Professor of Psychiatry, University of Warwick, Coventry</p> <p>Deputy Chair, Dr Howard Ring, Consultant & University Lecturer in Psychiatry, University of Cambridge</p> <p>Professor Jane Barlow, Professor of Public Health in the Early Years, Health Sciences Research Institute, Warwick Medical School</p> <p>Dr Sabyasachi Bhaumik, Consultant Psychiatrist, Leicestershire Partnership NHS Trust</p>	<p>Mrs Val Carlill, Public contributor</p> <p>Dr Steve Cunningham, Consultant Respiratory Paediatrician, Lothian Health Board</p> <p>Dr Anne Hesketh, Senior Clinical Lecturer in Speech and Language Therapy, University of Manchester</p> <p>Dr Peter Langdon, Senior Clinical Lecturer, School of Medicine, Health Policy and Practice, University of East Anglia</p> <p>Dr Yann Lefeuvre, GP Partner, Burrage Road Surgery, London</p>	<p>Dr Jeremy J Murphy, Consultant Physician and Cardiologist, County Durham and Darlington Foundation Trust</p> <p>Dr Richard Neal, Clinical Senior Lecturer in General Practice, Cardiff University</p> <p>Mr John Needham, Public contributor</p> <p>Ms Mary Nettle, Mental Health User Consultant</p> <p>Professor John Potter, Professor of Ageing and Stroke Medicine, University of East Anglia</p> <p>Dr Greta Rait, Senior Clinical Lecturer and General Practitioner, University College London</p>	<p>Dr Paul Ramchandani, Senior Research Fellow/Cons. Child Psychiatrist, University of Oxford</p> <p>Dr Karen Roberts, Nurse/Consultant, Dunston Hill Hospital, Tyne and Wear</p> <p>Dr Karim Saad, Consultant in Old Age Psychiatry, Coventry and Warwickshire Partnership Trust</p> <p>Dr Lesley Stockton, Lecturer, School of Health Sciences, University of Liverpool</p> <p>Dr Simon Wright, GP Partner, Walkden Medical Centre, Manchester</p>
--	--	--	--

Observers

<p>Dr Kay Pattison, Senior NIHR Programme Manager, Department of Health</p>	<p>Dr Morven Roberts, Clinical Trials Manager, Health Services and Public Health Services Board, Medical Research Council</p>	<p>Professor Tom Walley, CBE, Director, NIHR HTA programme, Professor of Clinical Pharmacology, University of Liverpool</p>	<p>Dr Ursula Wells, Principal Research Officer, Policy Research Programme, Department of Health</p>
---	---	---	---

Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (www.hta.ac.uk) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.