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 Oostenbrink[4] and D Mant[1]

[1]Department of Primary Care Health Sciences, Oxford University, Oxford, UK
[2]Department of General Practice, Katholieke Universiteit Leuven, Leuven, Belgium
[3]Department of Paediatrics, Division of Medical Education and Social Care, University of Leicester, Leicester, UK
[4]Department of Pediatrics, Erasmus MC-Sophia, Rotterdam, the Netherlands
[5]Department of General Practice, Groningen, the Netherlands

*Corresponding author

Executive summary

Health Technology Assessment 2012; Vol. 16: No. 15
DOI: 10.3310/hta16150

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

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 LRs+ ranging from 2.66 to 52.2, and poor peripheral circulation had LRs+ 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 ≥ 40 °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.

Publication

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).

NIHR Health Technology Assessment programme

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.

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

ISSN 2046-4932 (DVD)

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.