Royal College of Obstetricians and Gynaecologists

Bringing to life the best in women's health care

RCPCH Royal College of Paediatrics and Child Health

Leading the way in Children's Health

Feverish illness in children: assessment and initial management in children younger than 5 years

May 2013

NICE Clinical Guideline

National Collaborating Centre for Women's and Children's Health

NHS Evidence accredited provider

NHS Evidence - provided by NICE
www.evidence.nhs.uk

Feverish illness in children: assessment and initial management in children younger than 5 years

National Collaborating Centre for Women's and Children's Health

Commissioned by the National Institute for Health and Care Excellence

NICE's original guidance on Feverish illness in children was published in 2007. It was updated in 2013, 2017 and 2019.

See the NICE website for the guideline recommendations and evidence review for the 2019 update.

This document preserves evidence reviews and committee discussions for areas of the guideline that were not updated in 2019.

May 2013

Published by the Royal College of Obstetricians and Gynaecologists, 27 Sussex Place, Regent's Park, London NW1 4RG

www.rcog.org.uk

Registered charity no. 213280

First published May 2013

1st edition published in 2007

This guideline has been fully funded by NICE. Healthcare professionals are expected to take it fully into account when exercising their clinical judgement. However, the guidance does not override the individual responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient.

Implementation of this guidance is the responsibility of local commissioners and/or providers.

NCC-WCH Editor: Karen Packham

Contents

1 Guideline summary

This section was partially updated in 2013.

1.1 Guideline development group membership, NCC-WCH staff and acknowledgements

GDG members [2013]

Leah Bowen	Lay member
Richard Bowker	Consultant paediatrician
John Crimmins	General practitioner
Penny McDougall	Nurse
Edward Purssell	Lecturer in Children's Nursing
Debra Quantrill	Lay member
Martin Richardson	Consultant paediatrician, GDG Chair
Andrew Riordan	Consultant in Paediatric Infectious Diseases and Immunology
Damian Roland	NIHR Doctoral research fellow in Paediatric Emergency Medicine

National Collaborating Centre for Women's and Children's Health (NCC-WCH) [2013]

Zosia Beckles	Information scientist
Jiri Chard	Senior research fellow
Hannah-Rose Douglas	Associate director, Health Economics
Ella Fields	Research fellow
Zipporah Iheozor-Ejiofor	Research assistant
M Stephen Murphy	Co-director in Child Health
Nitara Prasannan	Research assistant
Cristina Visintin	Project manager

Acknowledgements

We would like to thank the following people for providing additional information:

Carlos Luaces Cubells, Monica Lakhanpaul, Cathy Pierce, Elizabeth Southey, Anne-Marie Stephani, Matthew Thompson and Tiffany Wong.

GDG members [2007]

Martin Richardson	Consultant paediatrician, GDG Chair
Richard Bowker	Paediatric specialist registrar
James Cave	General practitioner
Jean Challiner	Associate medical director – NHS Direct
Sharon Conroy	Paediatric clinical pharmacist
John Crimmins	General practitioner
Annette Dearnum	Children's nursing practitioner (deputy of Jane Houghton for 2 months)
Jennifer Elliott	Patient/carer representative
Jane Houghton	Nurse consultant in Paediatric Ambulatory Care
Edward Purssell	Lecturer in Children's Nursing
Andrew Riordan	Consultant in Paediatric Infectious Diseases and Immunology
Peter Rudd	Consultant paediatrician
Ben Stanhope	Consultant in Paediatric Emergency Medicine
Bridie Taylor	Patient representative (attending meetings till February 2006)

National Collaborating Centre for Women's and Children's Health (NCC-WCH) [2007]

Adebayo Akande	Research fellow
Monica Lakhanpaul	Co-director in Child Health
Chia-Wen Lee	Research fellow
Michael Corkett	Information specialist
Rosie Crossley	Work-Programme coordinator
Hannah-Rose Douglas	Health economist
Peny Retsa	Health economist

Acknowledgements

We would like to thank the Patient and Public Involvement Programme (PPIP) of the National Institute for Health and Clinical Excellence (NICE) whose glossary was adapted for use in this guideline. Francoise Cluzeau and Bobbie Lloyd also gave us support in conducting the Delphi consensus technique. We are grateful to all the healthcare professionals and parents and carers that took part in the consensus exercise. Diane Crawford gave us invaluable information about thermometers. We obtained information about the burden of infectious diseases from Roderick MacFaul, and Matthew Thompson kindly visited us to talk about his research on feverish illnesses in children presenting to primary care.

1.2 Definitions used in the guideline

Definitions used in the guideline

At the first stage of the guideline development process, the GDG recognised that it was necessary to have a definition of fever and also to decide what outcomes they would look for in terms of serious illness.

It was necessary for the GDG to define certain terms that could be used as inclusion or exclusion criteria for the guideline and literature searches.

Definition of fever

The GDG considered several definitions of fever that have been used in the scientific literature. The GDG was aware that normal body temperature varies within and between individuals. It was also recognised that the measurement of body temperature can vary with the site of measurement and type of thermometer used. Accordingly, it was acknowledged that any definition of fever based on a fixed body temperature would be arbitrary. It was therefore decided to use a well-recognised physiological definition.[16] For the purposes of this guideline, fever was thus defined as 'an elevation of body temperature above the normal daily variation'.

It was also decided that the entry point into the guideline would be a child presenting to health services with a measured or perceived fever. It was recognised that not all parents and carers have access to thermometers and it was considered appropriate that the definition and entry point allow the inclusion of children who are deemed to have a fever, with or without the use of a thermometer.

Despite agreeing on the above definition, the GDG recognised that other definitions of fever are used in most of the scientific studies that appear in the literature searches and evidence tables. For these studies, the inclusion criteria typically defined a fixed body temperature such as = 38°C or higher.

Definition of serious illness

Much of this guideline is devoted to identifying children with serious illnesses from among the many who present to healthcare professionals with a fever. The GDG recognised that it would be necessary to have a definition of serious illness to be used as an outcome measure in literature searches, etc. In addition to mortality and morbidity, it was agreed that a list of diagnoses that represented serious illnesses was needed. For the purposes of this guideline, serious illness with fever is defined as 'an illness with fever that could cause death or disability if there were a delay in diagnosis and treatment'.

The GDG also considered which diagnoses would fulfil this definition and, after consulting the literature, the following list of terms and diagnoses was included in literature searches:

- bacterial infection
- serious bacterial infection
- meningitis
- septicaemia
- bacteraemia
- pneumonia
- urinary tract infection
- septic arthritis
- osteomyelitis
- Kawasaki disease
- encephalitis (herpes simplex).

1.3 Who is covered by this guideline

This section was partially updated in 2013.

The scope of the guideline outlines who is and who is not covered by this guideline.

Groups that will be covered by this guideline are:

- Children from birth up to their 5th birthday presenting with a fever that has not been previously diagnosed.

No patient subgroups have been identified as needing specific consideration.

Groups that will not be covered by this guideline are:

- Children already admitted to hospital.
- Children with a pre-existing comorbidity for which fever is already covered by an established management plan by their specialist team; for example cystic fibrosis, immunosuppression, sickle cell disease and cerebral shunts.
- Children with recurring fever.
- Children diagnosed with tropical diseases.

1.4 Care pathway

A care pathway was used to identify patient flows and key decision points which informed the development of clinical questions.

The GDG designed an outline care pathway early in the development process to explore how a child with feverish illness might access, and be dealt with by, the health services. The resulting pathway is shown in Figure 1.3. The pathway starts with a child at home with fever, and the pathway and guideline come into effect when parents or carers decide to access the health services. The figure also illustrates a number of other concepts that were crucial to the guideline development process. More detailed clinical questions evolved from the pathway and the pathway was modified at the end of the development process to incorporate the recommendations derived from the updated clinical questions.

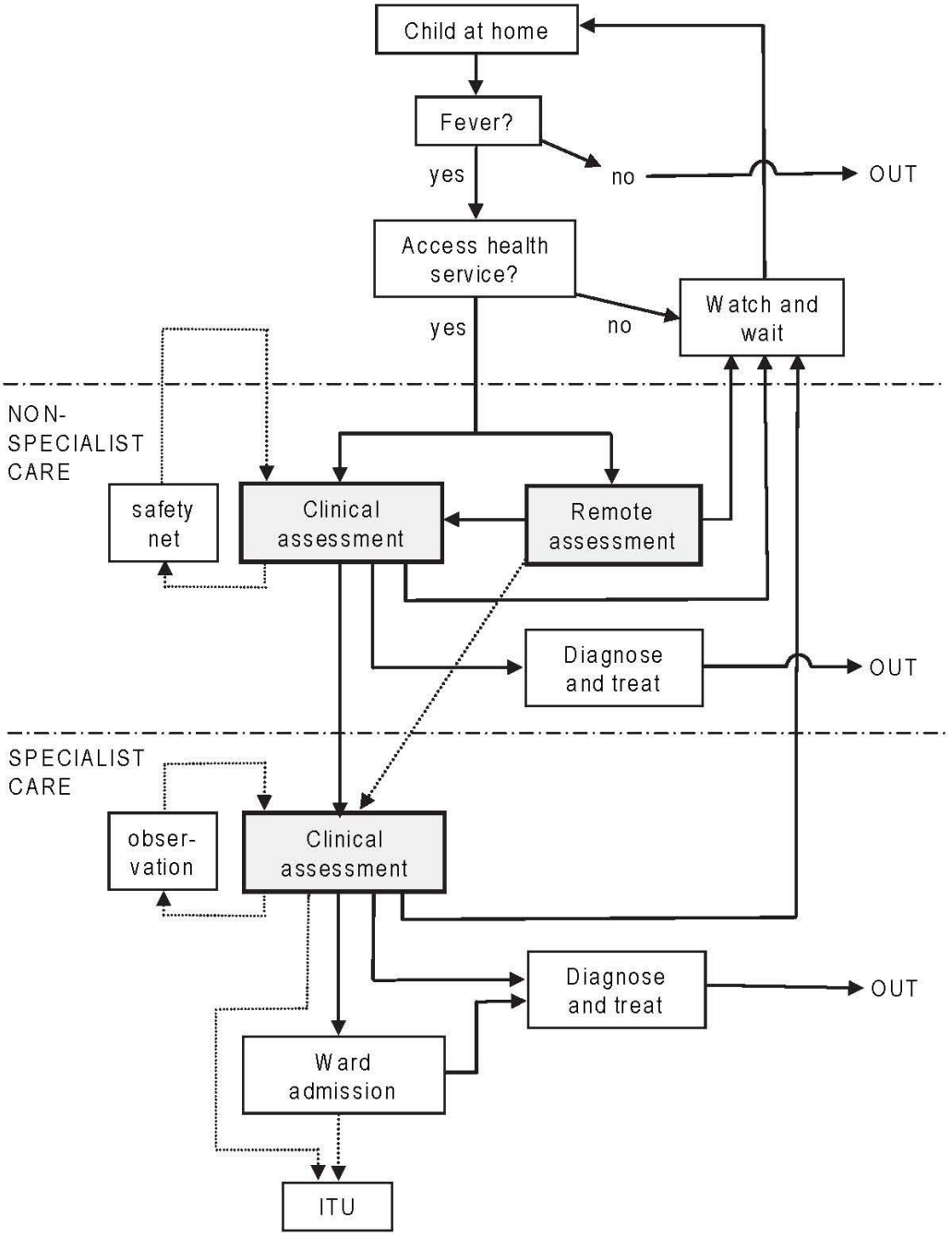
It was recognised that children with fever may currently be assessed by healthcare professionals who either have or do not have recognised training and/or expertise in the management of children and childhood diseases. In this guideline, professionals with specific training and/or expertise are described as paediatric specialists and they are said to be working in specialist care. Those without specific training and/or expertise are described as non-paediatric practitioners although it is acknowledged that such practitioners may be managing children and their illnesses on a regular basis. Non-paediatric practitioners are said to be working in non-specialist care.

For most children with feverish illness, the initial contact will be in non-specialist care. These contacts will mostly be in primary care but some non-specialist contacts may also be made in secondary care, for example in a general emergency department. A minority of these patients will then be referred on to specialist care, for example in a paediatric assessment unit.

The GDG recognised that assessments of children with feverish illness can take place in three main situations. These are represented by the shaded boxes on the care pathway in Figure 1.3. Broadly, assessments can take place in two ways in non-specialist care. The first is a traditional face-to-face encounter where the child undergoes a full clinical assessment, including history and physical examination. This usually occurs in general practice but it could equally occur in a walk-in centre or a hospital emergency department. Alternatively, the first point of contact could be with what has been described as a remote assessment. This is where the child is assessed by a healthcare professional who is unable to examine the child because the child is geographically remote from the assessor. Remote assessments are becoming increasingly important in the health service and they are used both in and out of normal working hours. Examples include NHS Direct and other telephone advice services. In some circumstances, although the child is not geographically remote from the assessor, it may not fall within the scope of practice for a particular healthcare professional to carry out a physical examination of the child, for example a pharmacist. In these circumstances, the healthcare professional may choose to follow the remote assessment guidance rather than the face-to-face guidance that takes into account signs found on physical examination. In specialist care, the clinical assessment will be undertaken by individuals trained in the care of sick children and the assessment may take place in a paediatric assessment unit, on a children's ward or in a dedicated paediatric emergency department.

The care pathway demonstrates a number of possible outcomes from each type of encounter with the health services. From a remote assessment, parents and carers will either be advised how to care for their child at home with appropriate advice as to when to seek further attention, or they will be advised to bring the child in for a formal clinical assessment. For the small number of children who have symptoms suggestive of an immediately life-threatening illness, the parents or carers will be advised to take the child for an immediate specialist assessment, for example by calling an ambulance. From a clinical assessment in non-specialist care, a child may again be returned home with appropriate advice. Alternatively, the child may be discharged with a 'safety net' that ensures that the child has some kind of clinical review or planned further contact with the health services (see Chapter 7). If the child is considered to be sick or potentially at risk of serious illness, the child will be referred to specialist care. In many cases, a firm diagnosis will be made by the non-paediatric practitioner and the child will be managed and treated accordingly. In these circumstances, the child progresses beyond the scope of this guidance and it is expected that the child would be treated according to relevant national or local guidelines.

Figure 1.1 Care pathway for feverish illness in children



The guideline will assume that prescribers will use a drug's summary of product characteristics to inform decisions made with individual patients.

1.5 Foreword

This section was partially updated in 2013.

This guidance is a partial update of NICE clinical guideline 47 (published in 2007) and will replace it.

New recommendations have been added on the assessment and initial management in children younger than 5 years with no obvious cause of feverish illness.

Recommendations are marked to indicate the year of the last evidence review:

- [2007] if the evidence has not been reviewed since the original guideline
- [2007, amended 2013] if the evidence has not been reviewed, but an essential change has been made that affects the meaning of the recommendation
- [2013] if the evidence has been reviewed but no change has been made to the recommendation
- [new 2013] if the evidence has been reviewed and the recommendation has been updated or added.

The original NICE guideline and supporting documents are available from www.nice.org.uk/CG47.

In the 2013 guideline the term meningitis has been replaced with bacterial meningitis, where appropriate.

Appendix K contains recommendations from the [2007] guideline that NICE deleted in the [2013] update. This is because the evidence has been reviewed and the recommendation has been updated, or because NICE has updated other relevant guidance and has replaced the original recommendations. Where there are replacement recommendations, details are provided. Where there is no replacement recommendation, an explanation for the proposed deletion is given.

A grey bar down the side of the page indicates those sections of the guideline which are new or have been updated. Material from the original guideline which has been deleted can be found in Appendix J.

1.6 Key priorities for implementation

The current recommendations can be found at www.nice.org.uk/guidance/ng143

1.7 Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

1.8 Key research recommendations

This section was partially updated in 2019. See the current recommendations at www.nice.org.uk/guidance/ng143

Number	Research recommendation	See section
	Symptoms and signs of serious illness	
RR 2	The GDG recommends a UK-based epidemiological study on the symptoms and signs of serious illness. [new 2013] Why this is important The current recommendations on symptoms and signs in the NICE guideline are based on a series of heterogeneous studies (using different methods, populations, outcomes and of varying quality) and a degree of subjectivity was needed to bring these together in the guideline. Therefore, the GDG recommends that a large prospective UK-wide study (n = 20,000 plus) should be undertaken comparing all of these symptoms and signs covered in the guideline. This would allow for a standardised comparison of each symptom and sign, and for validation of the existing 'traffic light' table. The study should use a standardised data collection protocol. Where possible the study should link with routinely collected data sets, such as Hospital Episode Statistics. The study should include a variety of settings and locations – that is, wherever children present, including primary care. The primary outcome of the study should be the final diagnosis and results of treatment.	5.5
RR3	**Management by remote assessment** The GDG recommends that a UK study is undertaken to determine the validity of symptoms reported on remote assessment for children with fever. [2007] Why this is important Traditionally, symptomatic patients have been assessed in a face-to-face setting but increasingly, remote assessment (for example, assessment over the telephone) determines the urgency of the patient's need, the level of care required and from that the most appropriate next step for the patient. This might include referral to emergency services, referral to acute or non-acute services or closing the call with self-care advice/support. Clinical and cost effectiveness will only be achieved through remote assessment if perceived need equates to actual need. There is currently a lack of data available that demonstrate the validity of remote assessment	6.2

RR5

Management by the paediatric specialist

8.3

Diagnosis

The GDG recommends that a UK study of the performance characteristics and cost-effectiveness of procalcitonin versus C-reactive protein in identifying serious bacterial infection in children with fever without apparent source be carried out. [2007].

Why this is important

Many young children with fever appear well with no symptoms or signs of serious illness. The vast majority of these children will have self-limiting illnesses. However, a few will have serious bacterial infections which may not be identifiable by clinical assessment alone. Investigations that help to identify these children with serious bacterial infections could lead to prompt antibiotic treatment, which may improve their outcome. These investigations need to be both sensitive and specific so that most serious bacterial infections are identified and so that antibiotics are not given to children who don't need them. The inflammatory markers C-reactive protein and procalcitonin have shown varying performance characteristics for identifying bacterial infection in a variety of populations. If either or both were found to be sensitive and specific for identifying serious bacterial infection in children with fever without apparent source, there would be evidence for their more widespread use. The cost effectiveness of this approach would need to be calculated

RR6

8.3

Antipyretics

The GDG recommends that studies are conducted in primary care and secondary care to determine whether examination or re-examination after a dose of antipyretic medication is of benefit in differentiating children with serious illness from those with other conditions. [2007]

Why this is important

Antipyretic medications are widely used in primary and secondary settings by parents and healthcare professionals. Children may therefore present to healthcare facilities having had a dose of antipyretics. Furthermore, the child's response to antipyretic drugs may be used as an indication of severity of illness, the rationale being that those with milder illness will either show greater improvement in condition or a greater reduction in their fever than children with more serious illnesses. However, it is not clear if such changes in condition are a valid and reliable method of differentiating children with serious illness from those with less serious conditions.

RR7

Advice for home care

10.1

Home-based antipyretic use

The GDG recommends studies on home-based antipyretic use and parental perception of distress caused by fever. [new 2013].

Why this is important

relieve distress in children. However, the concept of ‘distress’ and how parents act on it is little understood. Therefore, the GDG recommends that a study is undertaken to investigate ‘distress’ in children with feverish illness. The study should include parents’ and carers’ interpretation of this, including: help-seeking behaviour, what triggers presentation to a healthcare professional, what triggers the decision to give a dose of antipyretic, and what triggers the decision to change from one antipyretic to another.

1.9 Research recommendations

Number	Research recommendation	See section
	Thermometers and the detection of fever	
RR1	Measuring temperature in young babies: tympanic versus axilla electronic versus axilla chemical dot versus temporal artery. [2007]	4.2
	Clinical assessment of the child with fever	
	Symptoms and signs of serious illness	
RR 2	The GDG recommends a UK-based epidemiological study on the symptoms and signs of serious illness. [new 2013].	5.5
RR3	**Management by remote assessment** The GDG recommends that a UK study is undertaken to determine the validity of symptoms reported on remote assessment for children with fever. [2007]	6.2
	Management by the non-paediatric practitioner	
	Management according to risk of serious illness	7.2
RR4	The GDG recommends that research is carried out on referral patterns between primary and secondary care for children with fever, so the health economic impact of this and future guidelines can be estimated	
	Management by the paediatric specialist	8.3
RR5	Diagnosis The GDG recommends that a UK study of the performance characteristics and cost-effectiveness of procalcitonin versus C-reactive protein in identifying serious bacterial infection in children with fever without apparent source be carried out. [2007].	
RR6	Antipyretics The GDG recommends that studies are conducted in primary care and secondary care to determine whether examination or re-examination after a dose of antipyretic medication is of benefit in differentiating children with serious illness from those with other conditions. [2007]	8.3
	Advice for home care	
RR7	Home-based antipyretic use The GDG recommends studies on home-based antipyretic use and parental perception of distress caused by fever. [new 2013].	10.1

1.10 Other versions of the guideline

This section will be completed following the stakeholder consultation.

1.11 Schedule for updating the guideline

NICE clinical guidelines are updated so that recommendations take into account important new information. New evidence is checked after publication, and healthcare professionals and patients are

asked for their views; we use this information to decide whether all or part of a guideline needs updating. If important new evidence is published at other times, we may decide to do a more rapid update of some recommendations. Please see NICE website for information about updating the guideline.

2 Introduction

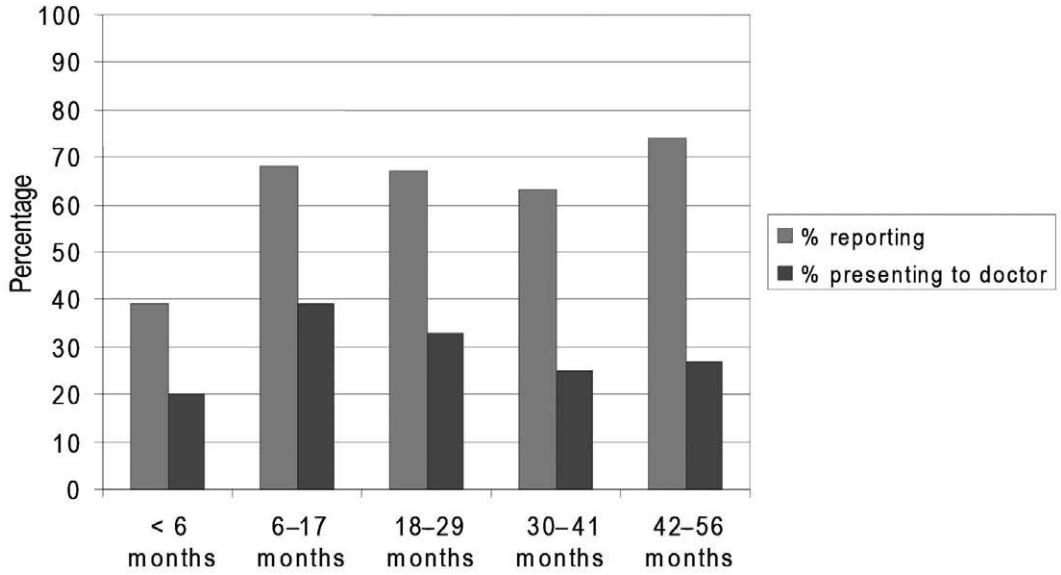
2.1 Feverish illness in children

Feverish illness in young children usually indicates an underlying infection of some kind and, as such, the condition is a cause of concern for parents and carers. The condition can be diagnostic challenge for healthcare professionals, and infectious diseases remain a major cause of child-hood mortality and morbidity in the UK. As a result, there is a perceived need to improve the recognition, evaluation and immediate treatment of feverish illnesses in children.

Incidence and prevalence

Feverish illness is very common in young children. Figure 2.1 shows the proportions of children from a birth cohort of all infants born in one English county (Avon) whose parents either reported a high temperature or presented to a doctor for this reason.[1] It can be seen that a high temperature is reported by nearly 40% of parents of children aged under 6 months, and in over 60% of children in the other age ranges between 6 months and 5 years. Between 20% and 40% of children in the various age ranges are taken to a doctor because of fever, with the highest proportions presenting between the ages of 6 and 18 months. It has been estimated that an average of eight infective episodes occur in otherwise healthy children during the first 18 months of life.[2]

Figure 2.1 Proportions of children reporting and presenting to doctors with high temperature by age range; data from Hay[1]



The prevalence of feverish illness in children is reflected by statistics from primary care. Fever is probably the most common reason for a child to be taken to the doctor. In a study of 1% of the national child population, the mean general practice (GP) consultation rate was 3.7 per child per year and almost double that rate for children aged under 4 years. Infections and respiratory disorders made up over 40% of the consultations.[3] In the fourth national study of morbidity in general practice, which included nearly 10 000 children, the annual consultation rates for infections were 60% of the population aged less than 12 months, 36% aged 1–4 years and 20% aged 5–15 years.[4] Not surprisingly, fever in children is also a common reason for seeking health advice out of hours. In one service, 34% of calls concerned children under 5 years of age.[5] Fever was a concern in 52% of calls about children aged under 12 months and in 64% of calls about children aged 1–5 years.

Feverish illness is also one of the most common reasons for children to be seen in hospital emergency departments and it is a leading cause of admission to children's wards. In a study from an emergency department in Nottingham, 32% of the 120 000 annual total attendances were for children.[6] Febrile illness was the second most common medical reason for attendance, accounting for 20% of such cases. On children's wards, at least 48% of admissions are associated with infection. Most of these infections

present with a feverish illness with or without other symptoms such as breathing difficulty, fit, rash or cough. Feverish illness is second only to breathing difficulty as the most common presenting problem leading to acute hospital admission in childhood.[7]

Issues for healthcare professionals

Feverish illness in young children can be a diagnostic challenge for healthcare professionals because it is often difficult to identify the cause. In most cases, the illness is due to a self-limiting virus infection and the child will recover quickly without intervention. However, fever may also be the presenting feature of serious bacterial illnesses such as meningitis, septicaemia, urinary tract infections and pneumonia. Estimates of the incidence of these and other serious infections are given in Table 2.1. Although there is quite a large variation in the estimated incidences according to the source of data, it appears that up to 1% of children aged 0–5 years may have one of these infections each year.

In some children with fever there will be symptoms and signs that suggest a particular infection, such as an inflamed eardrum in a child with otitis media or a non-blanching rash in a child with meningococcal septicaemia. When these features are identified, the diagnosis can be established relatively easily and the child can be treated appropriately. There will remain a significant number of children, however, who have no obvious cause of fever despite careful assessment and investigation. These children with fever without apparent source (FWS), are a particular concern to healthcare professionals because it is especially difficult to distinguish between simple viral illnesses and life- threatening bacterial infections in this group.[8] In general, FWS tends to be a problem in young children, and the younger the child the more difficult it is to establish a diagnosis and assess the severity of illness. Because of these problems, a number of diagnostic and management strategies have been developed for feverish illness without obvious source in young children.[9]

Table 2.1 Estimated incidence of serious infections in children aged 0–5 years in the UK; data from Hospital Episode Statistics (HES)

	Incidence (per 100 000)	
Diagnosis group	**HES data**	**Published data**
Pneumonia	664	92[a]
Septicaemia	388	20–50[b]
Urinary tract infection	333	
Meningitis	30.2	
Septic arthritis	9.25	3.75–5.0
Osteomyelitis	6.17	2.9
Other bacterial infection	0.66	
Encephalitis	3.65	0.8[c]
Kawasaki disease[d]	10.2	8.1
Total	**1445**	

[a] Pneumococcal pneumonia.

[b] Meningococcal septicaemia.

[c] Herpes simplex encephalitis.

[d] Kawasaki disease is not a confirmed infectious disease but it is believed to be caused by a microbiological toxin.

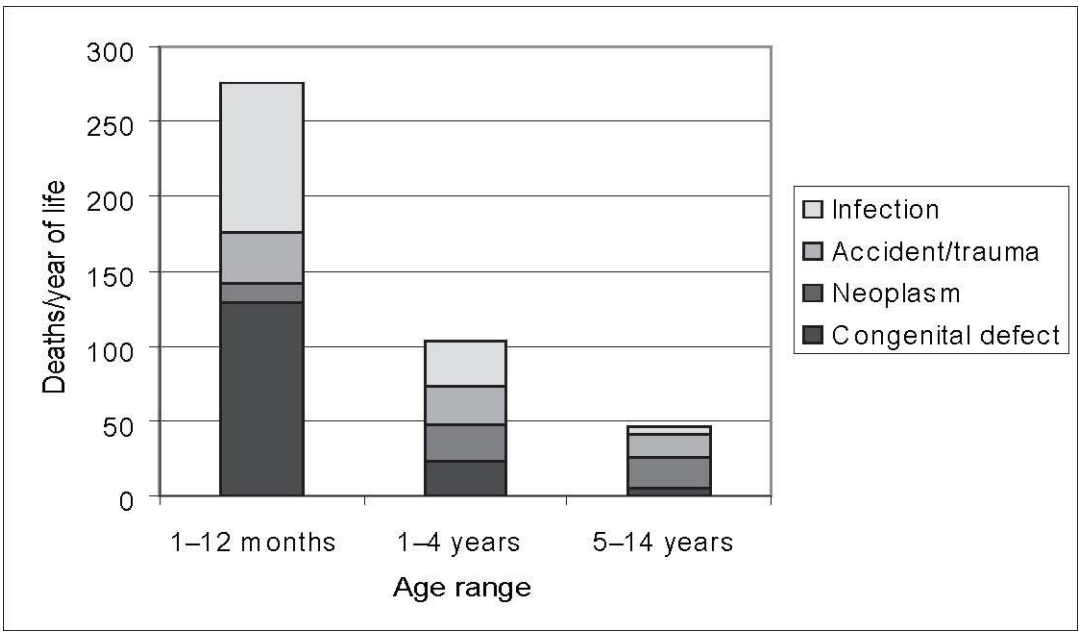
To further complicate the problem of assessment and diagnosis, the clinical picture often changes rapidly in young children. The condition of young children with serious illness may deteriorate within hours of onset but, on the other hand, an ill-appearing child with a viral illness may make a rapid recovery. Thus, another challenge for healthcare professionals is to determine when to observe the child for a period of time, and when to investigate and begin treatment.

Most healthcare professionals are aware that infectious diseases were, and remain, an important cause of mortality and morbidity in childhood. In the past hundred years there have been impressive reductions in childhood mortality. The infant mortality rate in the UK, for example, has fallen from 20% to 0.5% since 1890. Much of this improvement has been due to public health measures, and immunisation against infectious disease has increasingly been an important factor. In recent years, the reduction in childhood mortality has changed only a little. In other countries, mortality rates have continued to fall and some European countries now have childhood mortality rates that are 30–40% lower than that of the UK. These figures suggest that more can be done to reduce childhood mortality in this country.

Figure 2.2 shows that infection is a major cause of mortality in children aged 0–5 years. There are over 100 deaths from infection in children aged 1–12 months each year in England and Wales. In the first year of life, infection is second only to congenital defects as a cause of death. In children aged 1– 4 years there are around 30 deaths from infection per year of life, and infection is the most common cause of death in this age group.

It is possible that the childhood mortality rate in the UK could be reduced to a figure in line with other European countries if the proportion due to infections could be reduced. Immunisation will probably play an important part in this process. For example, the new pneumococcal conjugate vaccine, which was introduced into the UK schedule in 2006, has led to a dramatic reduction in invasive disease due to *Streptococcus pneumoniae* in other countries.[10] However, it is likely that improved recognition, evaluation and treatment of febrile illnesses in children could also lead to a reduction in mortality from infectious disease. For instance, a recent national study investigated deaths from meningococcal disease, which is the leading cause of mortality from infectious diseases in children.[11]The researchers found that mortality from meningococcal disease is often associated with late identification, sub- optimal treatment and other deficiencies in health care.

Figure 2.2 Contributions of the four major causative categories to childhood mortality, England and Wales, 2004; neonatal deaths and deaths due to perinatal events have been excluded; data from the Department of Health, courtesy of R MacFaul



Parental concern

This section was updated in 2013.

It is clear that febrile illnesses continue to have a considerable impact on childhood mortality and morbidity. This impact is reflected in the concerns of parents and carers. Several authors have conducted surveys of parents' responses to acute illness in their children and found that fever, cough and the possibility of meningitis were parents' primary concerns when their children became acutely ill (Kai, 1996; Hugenholtz et al, 2009). Parents reported that they experienced high levels of worry when their children were ill, and had particular concerns that the presence of fever might herald potential harm. Parents also had anxieties relating to the outcome of fever, believing that it could indicate serious illness such as meningitis. They were fearful that fever itself could damage their children and could also result in fits which they believed would result in permanent brain damage and even death.

This concern, which can lead to what has been described as fever phobia (Karwowska et al, 2002), is quite widespread and tends to increase with the height of temperature, rapid onset and duration of fever (Enarson et al, 2012). Additionally, it is important to recognise that fear of fever among parents can be influenced by ethnicity and cultural beliefs. These, compounded with concerns parents have for their children's well-being and the need for reassurance, often prompt parents to request care from both primary and secondary healthcare services (Hugenholtz et al, 2009; Taveras et al, 2004; Sands et al, 2011).

In scientific terms, fever is a natural response to infection and is not harmful in itself. Instead, it is the underlying infection that has the potential to cause harm. Indeed, there are some theoretical grounds to suggest that fever is beneficial in the body's response to infection. In any event, it is clear that parents and carers could receive more useful advice about feverish illness from healthcare professionals as well as recognition that their concerns are valid. This could include information about detecting potential serious infections, how to manage fever appropriately at home and when to seek further advice (Taveras et al, 2004; Kai, 1996).

Need for guidance

This section was updated in 2013.

It is a requirement of the Children's National Service Framework that all ill children should have access to high-quality, cost-effective, evidence-based care.[15] Because it is difficult to evaluate the severity of the illness, there is a need for evidence-based guidance to inform healthcare professionals about how to judge whether a child who presents with a fever is likely to develop a serious illness. Healthcare professionals also need advice to support their decision on whether to observe the child, perform diagnostic tests, start treatment such as antibiotics or refer onwards for specialist care. The guidance should also include advice on the best ways to detect fever, the management of fever itself, and what to tell parents and carers who have made contact with healthcare services. The guidance should be applicable to primary and secondary care and should take account of the number of agencies that are involved in giving health care and giving advice to parents and carers. It is also important that parental preferences, as well as the child's best interests in terms of health outcomes, should be taken into account when considering the various options for investigation and treatment.

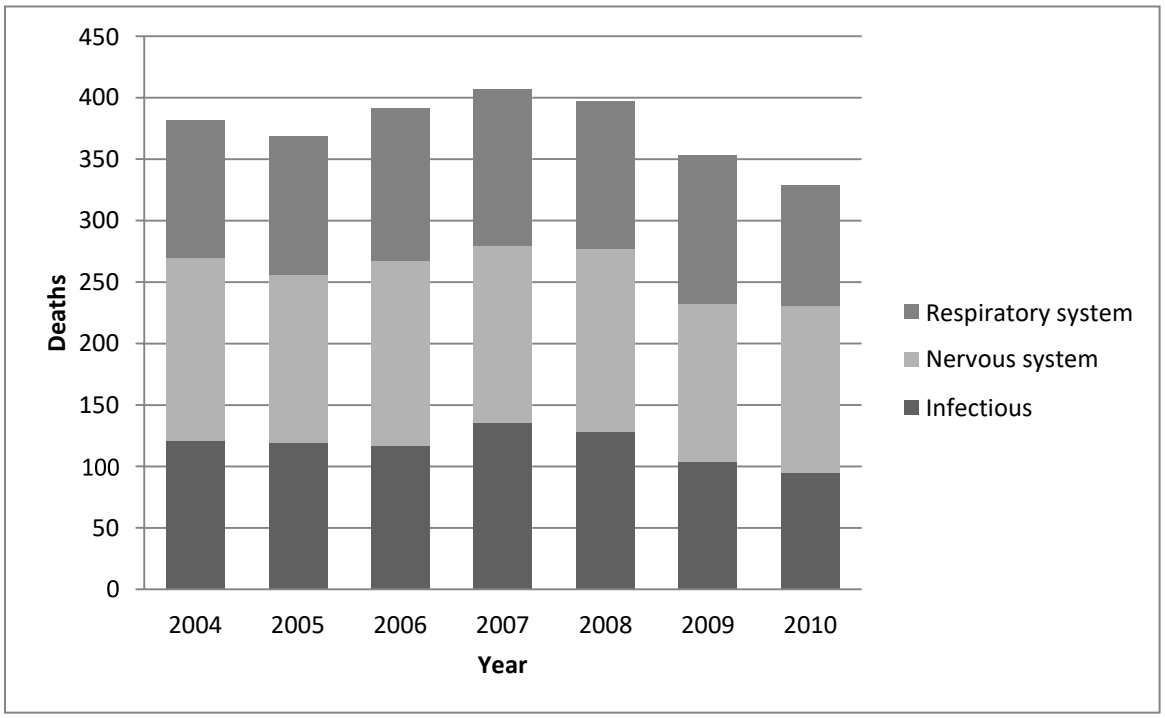
Need for 2013 update

This section was updated in 2013.

The decision to update the guideline was made based on deveopments in the NHS and new evidence becoming available that could affect existing recommendations.

The introduction of new vaccination programmes in the UK may have significantly reduced the level of admissions to hospital resulting from diseases covered by this guideline. For example, early analysis of the pneumococcal vaccination programme in England shows that the incidence of pneumococcal related disease has fallen 98% in children younger than 2 years since vaccination was introduced. However, evidence suggests a 68% increase in the prevalence of disease caused by sub-types of bacteria not covered by vaccination programmes. Also, potentially serious cases of feverish illness are likely to be rare, so it is important that information is in place to help healthcare professionals distinguish these from mild cases.

Figure 2.3 Mortality in children aged 1 month to 5 years in England and Wales caused by infection or diseases of the respiratory or nervous system, 2004 to 2010; data from Office of National Statistics.



In addition, new evidence is available on a number of the clinical questions covered by the guideline:

- the relationship of heart rate to fever in predicting the risk of serious illness in children
- clinical effectiveness of combination or alternating therapy with paracetamol and ibuprofen in the management of fever in children
- predictive value and accuracy of pro-calcitonin as a marker of serious bacterial illness in children with fever without apparent source.

2.2 For whom is this guideline intended

This clinical guideline is intended for use by all healthcare professionals who are involved in the care or management of young children with feverish illnesses. The guideline is intended for use in the full range of healthcare settings provided for children with acute illnesses, including both primary and secondary care. For the purposes of this guideline, primary care includes services such as NHS Direct, where the assessment of the child may not include a physical examination. The term specialist paediatric care has been used to define services where the child will be cared for and managed by trained paediatric staff. For the most part, the term refers to hospital paediatric departments and specialist children's emergency departments.

2.3 Related NICE guidance

- Urinary tract infection in children. NICE clinical guideline 54 (2007).
- Diarrhoea and vomiting in children under 5. NICE clinical guideline 84 (2009).
- Medicines adherence. NICE clinical guidance 76 (2011).
- Bacterial meningitis and meningococcal septicaemia, NICE clinical guideline 102(2010).
- Medicines adherence. NICE clinical guidance 76 (2011).

3 Guideline development methodology

This section was partially updated in 2013.

3.1 Methodology for the 2013 update

This partial update of guidance was commissioned by NICE and developed in accordance with the guideline development process outlined in the 2009 edition of The Guidelines Manual.

In accordance with NICE's Equality Scheme, ethnic and cultural considerations and factors relating to disabilities have been considered by the guideline development group (GDG) throughout the development process and specifically addressed in individual recommendations where relevant. Further information is available from: www.nice.org.uk/aboutnice/howwework/NICEEqualityScheme.jsp.

Developing review questions and protocols and identifying evidence

The scope for this update (see Appendix A) identified areas where substantial new evidence was available. The GDG formulated review questions based on the scope and prepared a protocol for each review question (see Appendix D). These formed the starting point for systematic reviews of relevant evidence. Published evidence was identified by applying systematic search strategies (see Appendix E) to the following databases: Medline (1948 onwards), Embase (1980 onwards), and four Cochrane databases (Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, the Database of Abstracts of Reviews of Effects and the Health Technology Assessment [HTA] database). Searches to identify economic studies were undertaken using the above databases and the NHS Economic Evaluation Database (NHS EED). Where appropriate, searches were limited by date to capture only studies published after the original guideline. Searches in Medline and Embase were limited to English language and studies in humans. None of the other searches were limited by language of publication (although publications in languages other than English were not reviewed). Search filters were used to identify particular study designs, such as randomised controlled trials (RCTs). There was no searching of grey literature, nor was hand searching of journals undertaken.

All the searches were updated and re-executed within 10 weeks of the start of the stakeholder consultation to ensure the reviews were up-to-date. This process was completed by 1 October 2012.

Reviewing and synthesising evidence

Evidence relating to clinical effectiveness was reviewed and synthesised according to the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach. In the GRADE approach, the quality of the evidence identified for each outcome listed in the review protocol is assessed according to the factors listed below, and an overall quality rating (high, moderate, low or very low) is assigned by combining the ratings for the individual factors.

- Study design (as an indicator of intrinsic bias; this determines the initial quality rating).
- Limitations in the design or execution of the study (including concealment of allocation, blinding, loss to follow up; these can reduce the quality rating).
- Inconsistency of effects across studies (this can reduce the quality rating).
- Indirectness (the extent to which the available evidence fails to address the specific review question; this can reduce the quality rating).

- Imprecision (reflects the confidence in the estimate of effect and this can reduce the quality rating). For continuous variables (such as change in temperature) the GDG was asked to predefine minimally important differences (the smallest difference between treatments that health professionals or patients think is clinically beneficial). However, the GDG was unable to agree these so imprecision was graded based on statistical differences.
- Other considerations (including large magnitude of effect, evidence of a dose–response relationship, or confounding variables likely to have reduced the magnitude of an effect; these can increase the quality rating in observational studies, provided no downgrading for other features has occurred).

The type of review question determines the highest level of evidence. For questions on therapy or treatment, the highest possible evidence level is a well-conducted systematic review or meta-analysis of RCTs, or an individual RCT. In the GRADE approach, a body of evidence based entirely on such studies has an initial quality rating of high, and this may be downgraded to moderate, low or very low if factors listed above are not addressed adequately. For questions on prognosis, the highest possible level of evidence is a controlled observational study (a cohort study or case–control study), and a body of evidence based on such studies would have an initial quality rating of high, which might be downgraded to moderate, low or very low, depending on the factors listed above. For diagnostic tests, studies examining the performance of the test were used if information on accuracy was required, but where an evaluation of the effectiveness of the test in the clinical management of the condition was required, evidence from RCTs or cohort studies was considered optimal. For studies evaluating the accuracy of a diagnostic test, summary statistics (sensitivity, specificity, positive predictive value [PPV], negative predictive value [NPV] and likelihood ratios for positive and negative test results [LR+ and LR–, respectively]) were calculated or quoted where possible (see Table 3.1). The following definitions were used when summarising the likelihood ratios for the GDG:

- Convincing: positive likelihood ratio (LR+) 10 or higher, negative likelihood ratio (LR-) 0.1 or lower
- Strong: LR+ 5 or higher (but less than 10), LR- 0.2 or lower (but higher than 0.1)
- Not strong: LR+ 4.9 or lower, LR- higher than 0.2

The following definitions were used when summarising the levels of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) for the GDG:

- High: 90% and above
- Moderate: 75% to 89%
- Low: 74% or below

All diagnostic outcomes (likelihood ratios, sensitivity, specificity and predictive values) were considered when discussing the evidence. However, particular emphasis was placed on the positive likelihood ratio, with a ratio of 5 or higher being considered a good indicator that a symptom or sign should be presented in the red column of the traffic light table.

For each review question the highest available level of evidence was sought. Where appropriate, for example if a systematic review, meta-analysis or RCT was identified to answer a question directly, studies of a weaker design were not considered. Where systematic reviews, meta-analyses and RCTs were not identified, other appropriate experimental or observational studies were sought.

The GRADE system described above covers studies of treatment effectiveness. However, it is less well established for studies reporting accuracy of diagnostic tests or prognostic factors. For such studies, NICE recommends using the Quality Assessment of Studies of Diagnostic Accuracy (QUADAS) methodology checklist or the NICE prognostic study checklist, respectively, to assess study quality (see the NICE guidelines manual). These were then mapped onto the GRADE system.

Some studies were excluded from the guideline reviews after obtaining copies of the publications because they did not meet inclusion criteria specified by the GDG (see Appendix G). The characteristics of each included study were summarised in evidence tables for each review question

(see Appendix H). Where possible, dichotomous outcomes were presented as relative risks (RRs) or odds ratios (ORs) with 95% confidence intervals (CIs), and continuous outcomes were presented as mean differences with 95% CIs or standard deviations (SDs).

The body of evidence identified for each therapy or treatment review question (or part of a review question) was presented in the form of a GRADE evidence profile summarising the quality of the evidence and the findings (pooled relative and absolute effect sizes and associated CIs).

Where appropriate, the body of evidence corresponding to each outcome specified in the review protocol was subjected to quantitative meta-analysis. In such cases, pooled effect sizes were presented as pooled risk ratios (RRs), pooled ORs or weighted mean differences. By default, meta- analyses were conducted by fitting fixed effects models, but where statistically significant heterogeneity was identified, random effects models were used to investigate the impact of the heterogeneity. As Review Manager does not support formal meta-analysis of diagnostic studies this was undertaken using the Stata® software package using the METANDI and MIDAS commands.

Where quantitative meta-analysis could not be undertaken (for example because of heterogeneity in the included studies) the range of effect sizes reported in the included studies was presented. The GRADE evidence profiles are not directly applicable to epidemiological studies or non-comparative cohort studies. Where these studies are presented, they are included in descriptive paragraphs and/or tables as appropriate.

Table 3.1 '2 x 2' table for calculation of diagnostic accuracy parameters

	Reference standard positive	**Reference standard negative**	**Total**
Index test result positive	a (true positive)	b (false positive)	a+b
Index test result negative	c (false negative)	d (true negative)	c+d
Total	a+c	b+d	a+b+c+d = N (total number of tests in study)

Identification of serious illness

The following serious illnesses were identified as being the main focus of the diagnostic reviews:

- bacterial meningitis
- meningococcal septicaemia
- bacteraemia
- pneumonia
- urinary tract infection
- encephalitis (herpes simplex)
- septic arthritis/osteomyelitis
- Kawasaki disease.

Outcome measures

For this guideline update, the review questions were judged on a number of outcomes. The justification for using these outcomes was based on their relevance to the groups covered by the guideline and consensus among members of the GDG. Outcomes include those that were felt to be desirable (for example early detection of serious illness) and unwanted effects of treatment that it would be important to reduce to a minimum. When assessing the accuracy of a test or the

effectiveness of a particular treatment, appropriate information about the effect on one or more primary outcomes was sought.

The primary outcomes considered in the guideline were:

- accuracy in identifying serious illness
- change in the child's 'distress'
- change in child's temperature
- adverse events.

The GDG stated that the overarching aim of the guideline was the early and accurate detection of serious illness in children with fever. This allows for suitable treatment to begin, which should then reduce morbidity and mortality.

Incorporating health economics

The aims of the health economic input to the guideline were to inform the GDG of new economic issues relating to fever in children, and to consider whether the recommendations continued to represent a cost-effective use of healthcare resources. Health economic evaluations aim to integrate data on benefits (ideally in terms of quality adjusted life years [QALYs]), harms and costs of different care options.

Systematic searches for published economic evidence were undertaken for all clinical questions in the guideline update. For economic evaluations, no standard system of grading the quality of evidence exists and included papers were assessed using a quality assessment checklist based on good practice in economic evaluation. Reviews of the relevant published health economic literature identified in the literature search are presented alongside the clinical effectiveness reviews.

The GDG prioritised a number of review questions where it was thought that economic considerations would be particularly important in formulating recommendations. The plan was to provide additional health economic analyses where data were available and health economic analysis was warranted as part of the development process. Cost effectiveness analysis can be useful where there are alternative clinical strategies, one or more of which is associated with potentially higher costs and evidence of improved effectiveness. For this guideline the areas prioritised for economic analysis were:

- the predictive value of pro-calcitonin and/or C reactive protein markers
- the efficacy of paracetamol and non-steriodal anti-inflammatory drugs (NSAIDs) alone and in combination in reducing fever
- whether reducing fever with paracetamol or NSAIDs affects the course of the illness.

Evidence to recommendations

Recommendations for clinical care were derived using, and linked explicitly to, the evidence that supported them. In the first instance, informal consensus methods were used by the GDG to agree short clinical and, where appropriate, cost effectiveness evidence statements which were presented alongside the evidence profiles. Statements summarising the GDG's interpretation of the evidence and any extrapolation from the evidence used when making recommendations were also written to ensure transparency in the decision-making process. The criteria used in moving from evidence to recommendations were:

- relative value placed on the outcomes considered
- consideration of clinical benefits and harms consideration of net health benefits and resource use
- quality of the evidence
- other considerations (including equalities issues).

The GDG also identified areas where evidence to answer its review questions was lacking and used this information to formulate recommendations for future research.

Towards the end of the guideline development process, formal consensus methods were used to consider all the clinical care recommendations and research recommendations that had been drafted. The GDG identified 10 'key priorities for implementation' (key recommendations) and five high-priority research recommendations. The key priorities for implementation were those recommendations thought likely to have the greatest impact on clinical care and outcomes in the NHS as a whole; they were selected using a variant of the nominal group technique (see the NICE guidelines manual). The priority research recommendations were selected in a similar way.

Stakeholder involvement

Registered stakeholder organisations were invited to comment on the draft scope and the draft guideline. The GDG carefully considered and responded to all comments received from stakeholder organisations. The comments and responses were reviewed by NICE in accordance with the NICE guideline development process.

3.2 Methodology for the 2007 guideline

This section outlines the methodology used in the development of the 2007 guideline and applies only to those parts of the guideline that were developed in 2007.

This guideline was commissioned by NICE and developed in accordance with the guideline development process outlined in the 2005 NICE *Guidelines Manual*.[17]

Literature search strategy

Initial scoping searches were carried out to identify relevant guidelines (local, national, international) produced by other development groups. The reference lists in these guidelines were checked against subsequent searches to identify missing evidence.

Systematic searches to answer the clinical questions formulated and agreed by the GDG were carried out using the following databases via the OVID platform: MEDLINE (1966 onwards), Embase (1980 onwards), Cumulative Index to Nursing and Allied Health Literature (1982 onwards) and PsycINFO (1967 onwards). The most recent search conducted for the three Cochrane databases (Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and the Database of Abstracts of Reviews of Effects) was Quarter 3, 2006. Searches to identify economic studies were undertaken using the above databases and the NHS Economic Evaluations Database (NHS-EED).

Relevant published evidence to inform the guideline development process and answer the clinical questions was identified by systematic search strategies. The clinical questions are shown in the relevant sections. Additionally, stakeholder organisations were invited to submit evidence for consideration by the GDG, provided it was relevant to the clinical questions and of equivalent or better quality than evidence identified by the search strategies. GDG members also contributed evidence under the same conditions.

Search strategies combined relevant controlled vocabulary and natural language in an effort to balance sensitivity and specificity. Both generic and specially developed methodological search filters were used appropriately. Unless advised by the GDG, searches were not date specific.

There was no systematic attempt to search grey literature (conferences, abstracts, theses and unpublished trials). Hand searching of journals not indexed on the databases was not undertaken. Ongoing trials were identified and the principal investigators asked to share their research proposals and outcomes, if available.

Although search strategies were devised for children under the age of 5 years, evidence beyond this age group was considered when no other evidence was available for children under 5 years. Refer to the evidence tables outlining these studies on the accompanying CD-ROM.

Searches were updated and re-run 6–8 weeks before the stakeholder consultation, thereby ensuring that the latest relevant published evidence was included in the database. Any evidence published

after this date was not included. For the purposes of updating this guideline, 1 September 2006 should be considered the starting point for searching for new evidence.

Further details of the search strategies, including the methodological filters used, are provided on the accompanying CD-ROM.

Synthesis of clinical effectiveness evidence

The NICE *Guidelines Manual* was largely abided by. However, because this is a symptom-based guideline with un-established methodology, the methodology used is stated where it was not covered in the NICE *Guidelines Manual*. Evidence relating to clinical effectiveness was reviewed using established guides[17–24] and classified using the established hierarchical system shown in Table 3.2.[24] This system reflects the susceptibility to bias that is inherent in particular study designs.

The type of clinical question determines the highest level of evidence that may be sought. In assessing the quality of the evidence, each study receives a quality rating coded as '++', '+' or '-'. For issues of therapy or treatment, the highest possible evidence level (EL) is a well-conducted systematic review or meta-analysis of randomised controlled trials (RCTs; EL = 1++) or an individual RCT (EL = 1+). Studies of poor quality are rated as '-'. Usually, studies rated as '-' should not be used as a basis for making a recommendation, but they can be used to inform recommendations. For issues of prognosis, the highest possible level of evidence is a cohort study (EL = 2) since this is the most appropriate methodology to address prognosis. There are no specific ELs for prognosis and therefore all the prognostic studies were rated according to Table 3.2.

Table 3.2 Levels of evidence for intervention studies[17]

Level	Source of evidence
1++	High-quality meta-analyses, systematic reviews of randomised controlled trials (RCTs) or RCTs with a very low risk of bias
1+	Well-conducted meta-analyses, systematic reviews of RCTs or RCTs with a low risk of bias
1−	Meta-analyses, systematic reviews of RCTs or RCTs with a high risk of bias
2++	High-quality systematic reviews of case–control or cohort studies; high-quality case–control or cohort studies with a very low risk of confounding, bias or chance and a high probability that the relationship is causal
2+	Well-conducted case–control or cohort studies with a low risk of confounding, bias or chance and a moderate probability that the relationship is causal
2−	Case–control or cohort studies with a high risk of confounding, bias or chance and a significant risk that the relationship is not causal
3	Non-analytical studies (for example case reports, case series)
4	Expert opinion, formal consensus

For each clinical question, the highest available level of evidence was selected. Where appropriate, for example if a systematic review, meta-analysis or RCT existed in relation to a question, studies of a weaker design were not included. Where systematic reviews, meta-analyses and RCTs did not exist, other appropriate experimental or observational studies were sought, such as diagnostic studies, which examined the performance of the clinical test if the efficacy of the test was required (see Table 3.3). Where an evaluation of the effectiveness of the test in the clinical management of patients and the outcome of disease was required, evidence from RCTs or cohort studies was used.

The system in Table 3.2 covers studies of treatment effectiveness. However, it is less appropriate for studies reporting diagnostic tests of accuracy. In the absence of a validated hierarchy for this type of test, NICE suggests levels of evidence that take into account the factors likely to affect the validity of these studies (see Table 3.3).

Table 3.3 Levels of evidence for studies of the accuracy of diagnostics tests[17]

Level	Type of evidence
Ia	Systematic reviews (with homogeneity)[a] of level-1 studies[b]
Ib	Level-1 studies[b]
II	Level-2 studies;[c] systematic reviews of level-2 studies
III	Level-3 studies;[d] systematic reviews of level-3 studies
IV	Consensus, expert committee reports or opinions and/or clinical experience without explicit critical appraisal; or based on physiology, bench research or 'first principles'

[a] Homogeneity means there are no or minor variations in the directions and degrees of results between individual studies that are included in the systematic review.

[b] Level-1 studies are studies that use a blind comparison of the test with a validated reference standard (gold standard) in a sample of patients that reflects the population to whom the test would apply.

[c] Level-2 studies are studies that have only one of the following:

- narrow population (the sample does not reflect the population to whom the test would apply)
- use a poor reference standard (defined as that where the 'test' is included in the 'reference', or where the 'testing' affects the 'reference')
- the comparison between the test and reference standard is not blind
- case–control studies.

[d] Level-3 studies are studies that have at least two or three of the features listed above.

Prognostic studies

A substantial part of the evidence for this guideline was derived from prognostic studies. It is worth noting that there is very limited research on prognostic studies and on methods for assessing their quality. The 2005 version of the NICE *Guidelines Manual* contains virtually no advice on how to assess such studies. These limitations were recognised from the outset and the NICE methodology was adapted to account for these deficiencies, as outlined in Table 3.3.

For searching, a highly sensitive evidence-based prognostic study search strategy developed by McMaster University was adopted. Searches for this evidence utilised a prognostic search filter by Wilczynski*et al*.[25] full details of the search strategy are provided on the accompanying CD-ROM.

The search identified 3151 prognostic studies. After filtering double references, 300 different abstracts were screened for inclusion.

Studies were appraised using the checklist for cohort studies in Appendix D of the 2005 version of the NICE *Guidelines Manual*, and the evidence level was allocated using the hierarchy described in Table 3.2. According to this system, the best quality evidence would usually be of evidence level 2 because RCTs are not usually used to address questions of prognosis. Prospective cohort studies are generally the preferred type of study. Lower evidence level studies were included on an individual basis if they contributed information that was not available in the higher evidence level studies but yielded important information to inform the GDG discussions for formulating recommendations.

Delphi consensus

In areas where important clinical questions were identified but no substantial evidence existed, a two-round Delphi consensus method was used to derive recommendations that involved the participation of over 50 clinicians, parents and carers from appropriate stakeholder organisations. The participants rated a series of statements developed by the GDG using a scale of 1–9 (1 being strongly disagree, 9 being strongly agree). Consensus was defined as 75% of ratings falling in the 1–3 or 7–9 categories. Results and comments from each round were discussed by the GDG and final recommendations were made according to predetermined criteria. Full details of the consensus process are presented in Appendix A.

For economic evaluations, no standard system of grading the quality of evidence exists. Economic evaluations that are included in the review have been assessed using a quality assessment checklist based on good practice in decision-analytic modelling.[26] Evidence was synthesised qualitatively by

summarising the content of identified papers in evidence tables and agreeing brief statements that accurately reflected the evidence. Quantitative synthesis (meta-analysis) was not performed in this guideline due to methodological and statistical heterogeneity of the studies identified.

Summary results and data are presented in the guideline text. More detailed results and data are presented in the accompanying evidence tables. Where possible, dichotomous outcomes are presented as relative risks (RRs) with 95% confidence intervals (CIs), and continuous outcomes are presented as mean differences with 95% CIs or standard deviations (SDs). Moreover, RRs were also calculated as positive predictive values (PPV)/(1 - negative predictive value [NPV]) in diagnoses and prognoses when appropriate.

The quality of cohort studies was appraised based on Appendix B in the 2005 NICE *Guideline Manual*, and Appendix F for diagnostic studies.

Health economics

The aim of the economic input into the guideline was to inform the GDG of potential economic issues relating to fever in children. The health economist helped the GDG by identifying topics within the guideline that might benefit from economic analysis, reviewing the available economic evidence and, where necessary, conducting economic analysis. Where published economic evaluation studies were identified that addressed the economic issues for a clinical question, these are presented alongside the clinical evidence. However, this guideline addressed only assessment and initial management of fever in children. Economic evaluation requires assessment of healthcare resources (costs) alongside health outcomes, preferably quality-adjusted life years (QALYs). Since clinical outcomes of treatment were outside the scope of the guideline, it was anticipated that the economic literature that addressed the guideline questions would be very limited.

Apart from the review of the literature, additional health economic analysis was undertaken for specific questions in the guideline which the GDG identified as requiring economic evaluation. Specifically, health economic analysis was undertaken on the cost of thermometers, and the cost- effectiveness of specific investigations in specialist care (C-reactive protein versus procalcitonin). Additional economic models were developed to assess the impact of changing the pattern of referrals to secondary care but the lack of data prevented any meaningful analysis and conclusions to be drawn from this.

For the analysis that was undertaken, clinical data reported in the guideline were used, and UK cost data were collected. The perspective adopted is the NHS and cost data are reported for 2005/06.

Health economic analysis carried out as part of the guideline development is presented within the relevant clinical chapter, with readers being referred forward to appendices which provide more detailed explanation of methods and results.

Health economic statements are made in the guideline in sections where the use of NHS resources is considered.

Forming recommendations

For each clinical question, the recommendations were derived from the evidence statements presented to the GDG as summaries from the studies reviewed. The link between the evidence statements and recommendation were made explicit in the translation of the evidence statement. The GDG agreed the final recommendation through informal consensus. In the first instance, informal consensus methods were used by the GDG to agree evidence statements and recommendations. Additionally, in areas where important clinical questions were identified but no substantial evidence existed, formal consensus methods were used to identify current best practice (see the section above). Shortly before the consultation period, five to ten key priorities were selected using a nominal group technique for implementation (details available at the NCC-WCH). To avoid giving the impression that higher grade recommendations are of higher priority for implementation, NICE no longer assigns grades to recommendations.

External review

This guideline has been developed in accordance with the NICE guideline development process. This has included giving registered stakeholder organisations the opportunity to comment on the scope of the guideline at the initial stage of development and on the evidence and recommendations at the concluding stage. This involved reviewing by two independent reviewers as part of NICE's external expert review process for its guidelines. The developers have carefully considered all of the comments during the stage of the consultation by registered stakeholders and expert external reviewers and validation by NICE.

4 Thermometers and the detection of fever

Introduction

Body temperature in children can be measured at a number of anatomical sites using a range of different types of thermometers. Sites used to measure temperature include the mouth, rectum and axilla. The types of thermometers available include mercury-in-glass, electronic, chemical and infrared. Mercury-in-glass thermometers were the traditional type of thermometer used to measure body temperature but they are no longer recommended for use in infants and young children because of the risks of breakage and mercury spillage.[27] Furthermore, UK health and safety regulations require that mercury-containing medical devices should not be used whenever a suitable alternative exists.[28] Mercury-in-glass thermometers will not be considered further in this guideline except as a comparator in diagnostic studies.

Electronic thermometers are widely used by healthcare professionals as an alternative to mercury-in-glass thermometers. Electronic thermometers have the advantages of being accurate and very quick to use but they are often complex and quite expensive pieces of medical equipment. Recently, cheaper compact electronic thermometers have been produced and these are available for use by the public as well as healthcare professionals. Chemical phase-change thermometers measure body temperature by using a combination of chemicals that change colour in response to variations in temperature. These can either be chemical dot thermometers where the chemicals are contained in cells on a plastic stick, or chemical forehead thermometers which consist of a patch of chemicals in a plastic pouch that is placed on the forehead. Chemical dot thermometers are usually designed for single use but reusable types are available. All types of chemical thermometers can be used by the public. In recent years, infrared thermometers have been used more and more frequently. This type of thermometer detects infrared radiation from blood vessels and this is then used to estimate central body temperature. Most thermometers of this type measure temperature at the eardrum (infrared tympanic thermometers) but temporal artery thermometers are now available where temperature is measured on the scalp. Infrared thermometers are quick, non-invasive and simple to use. They are relatively expensive, however.

In this chapter, the different sites and thermometers are compared with regard to their accuracy in measuring true body temperature and their ability to detect fever. In general, the various sites and different types of thermometers are compared in their diagnostic ability against a traditional gold standard. The gold standard is usually a measurement with a mercury-in-glass or electronic thermometer using the mouth in older children and the rectum in young children and infants. This chapter also looks at the ability of parents and carers to detect fever in young children using subjective means such as palpation of the child's brow.

4.1 Thermometers and the site of measurement

Review questions

How accurate are the different types of thermometer in the measurement of body temperature in young children, and how do they compare in their ability to detect fever?

How accurate are the readings of temperature from different sites of the body in young children, and how do these sites compare in the ability to detect fever?

Body temperature can be recorded from a number of sites in the body in babies and young children. Traditionally, temperature was taken by the oral route in older children and adults, while the rectal

route was used in infants and young children. Alternatives methods include using the axilla or using a tympanic thermometer. These methods are generally considered to not be as accurate as traditional measurement[29,30] but they are often quicker and easier to use in young children.[31] Axillary and tympanic measurements may also be better accepted by children and their carers.[31,32]

Oral and rectal temperature measurements

Review question

How accurate are the different types of thermometer in the measurement of body temperature in young children, and how do they compare in their ability to detect fever?

How accurate are the readings of temperature from different sites of the body in young children, and how do these sites compare in the ability to detect fever?

Narrative evidence

An attempt was made to find evidence of the comparative accuracy of oral and rectal temperature measurements using mercury-in-glass or electronic thermometers. Two EL II studies were found that looked at the diagnostic accuracy of an electronic thermometer embedded in an infant pacifier.[33,34] The studies recruited children of different ages (e.g. 10 days to 24 months[33] to < 2 years[34]). The reported sensitivity was 10% and 63.3%, respectively.

The GDG did not consider these studies to be applicable to UK practice because these thermometers are not available and the evidence for their usefulness is weak.

Evidence summary

The GDG was aware that temperature measurements by the oral and rectal routes were rarely used in young children by healthcare professionals in the UK. These sites are probably the most accurate for temperature measurement but there are concerns about their safety and acceptability. The GDG could not reach a consensus among themselves as to whether these routes should be used and it was therefore decided to use the Delphi technique in an attempt to achieve formal consensus.

Regarding oral thermometers, the following background information and statement was put to the Delphi panel.

Background

In older children and adults, the inside of the mouth is considered to be one of the most accurate sites for the measurement of body temperature. When temperature is measured via the mouth, it is necessary for the thermometer to be held in place under the tongue while the measurement is taken. Most children's nurses are taught that children under the age of 5 years cannot cooperate with this procedure and that inaccurate measurements will be obtained. There are also concerns that some young children will bite the thermometer, and others find the technique uncomfortable or even painful.

Delphi statement 7.2

Healthcare professionals should not routinely use the oral route (mouth) to measure body tem- perature in children under the age of 5 years. The following responses were obtained from the first round of the Delphi process (see section 3.2):

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
2 (4%)	4 (8%)	44 (85%)	2 (4%)	1	52	9

The statement therefore achieved consensus at the first round of the Delphi technique.

Regarding rectal thermometers, the following background information and statements were put to the Delphi panel. The results from the first round of the Delphi process are also shown.

Background

In this technique, the probe of an electronic thermometer is placed in the rectum (back passage). The rectum is often considered the most accurate site of measurement of body temperature; the rectal route is therefore a reliable way of detecting fever in babies and young children.

Some people find rectal thermometers unacceptable for routine use. In newborn babies there have been reports of injuries including perforation of the bowel after the use of rectal mercury thermometers. Some people are concerned that electronic thermometers could have the same effect. In newborn babies taking the temperature in the axilla (armpit) is almost as accurate as using the rectal route (back passage).

Delphi statement 7.3

Healthcare professionals should routinely use electronic thermometers by the rectal route (back passage) to measure body temperature in children aged: 0–3 months.

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
45 (87%)	3 (6%)	3 (6%)	1 (2%)	1	52	1

The statement therefore achieved consensus at the first round of the Delphi technique.

Delphi statement 7.4

Healthcare professionals should not routinely use electronic thermometers by the rectal route (back passage) to measure body temperature in children aged 3 months to 2 years.

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
46 (88%)	4 (8%)	1 (2%)	1 (2%)	1	52	1

The statement therefore achieved consensus at the first round of the Delphi technique.

Delphi statement 7.5

Healthcare professionals should routinely use electronic thermometers by the rectal route (back passage) to measure body temperature in children aged 2–5 years.

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
47 (92%)	3 (6%)	0	1 (2%)	1	52	1

The statement therefore achieved consensus at the first round of the Delphi technique.

Delphi evidence summary

There was a lack of evidence on the relative accuracy or ability to detect fever using the oral and rectal routes of temperature measurement. The Delphi panel achieved consensus at the first round on all statements relating to oral and rectal temperature measurements. Eight-five percent of the panel agreed with the statement that the oral route should not be used routinely in young children. On the three statements regarding the rectal route, between 87% and 92% of the panel disagreed with the recommendation that this route should be used routinely. (EL IV)

GDG translation

The GDG considered that the results of the Delphi process indicated strongly that the oral and rectal routes should not be used for routine temperature measurements in infants and young children.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

4.2 Measurement of body temperature at other sites

In the event of not recommending temperature measurements by the oral or rectal route, it was necessary for the GDG to recommend an alternative method of measurement. The GDG collected data on axillary measurements using electronic and chemical thermometers, infrared measurements at the tympanic and temporal artery sites, and on forehead crystal thermometers. The GDG looked at evidence on the accuracy and ability to detect fever of these sites and thermometers.

Narrative evidence

Axillary temperature measurement

One EL 2+ SR[29] and 20 prospective studies (two EL Ib,[35,36] ten EL II[37–46] and eight EL III[47–54]) were found. The EL reflects the quality of report but may not necessarily reflect the quality of the studies themselves. Therefore, all the EL III studies were judged to be adequate for inclusion to inform recommendation. There is tremendous methodological heterogeneity among the included studies. For instance, the age of included children varied from 12–48 hours after birth[36] to 6–14 years,[48] the setting also varied from birth registry,[55] paediatric ward,[44] and emergency department[56] to nursery.[43] There is also variation of the device (e.g. mercury[43] or digital[44] thermometry). Owing to the clinical and statistical heterogeneity, it was inappropriate to perform meta-analysis. The findings suggest that, on average, axillary temperature underestimates body temperature by at least 0.5°C (although the difference between the body temperature may be smaller when a mercury thermometer rather than an electronic one is used). There is also a wide range of variation between individuals. The mean difference between axillary temperature and body temperature varied between 0.09°C[57] and 1.52 °C,[40] and the SR29 showed that the upper limit of mean difference was 2°C if axillary temperature was taken by digital thermometers. Furthermore, the sensitivities for detecting fever ranged from 25%[35] to 98%.[39]

For studies with data specifically looking at neonates, the reported mean differences between rectal and axillary temperature were 0.09°C (95% CI 0.06 to 0.12°C),[43] 0.3°C,[58] and 0.2 °F.[36] There appeared to be a significant correlation between the rectal and axillary temperatures;[46,49,36] no sensitivity and specificity were reported in this subgroup. Moreover, one EL II study[37] reported that in infants younger than 1 month, the difference between the axillary and rectal temperatures varied with age. Least squares linear regression analysis showed that the rectal temperature was equal to the axillary temperature plus 0.2°C for each week of age up to 5 weeks.

Chemical dot (phase-change) thermometers

Three EL II prospective cohort studies[45,59,60] investigating the diagnostic accuracy of chemical dot thermometers were found. Only the diagnostic accuracy of chemical dot thermometers used in the axilla was looked at. The age and setting of children included varied from 0–102 days in neonatal ICU[60] to 3–36 months admitting to hospitals.[45] The mean difference in axillary temperature between chemical dot and mercury thermometer measurement was 0.32°C[59] to 0.93°C.[60] Moreover, the sensitivity ranged between 68%[45] and 92%,[59] with RR of 17.2[59] to detect fever.

Forehead crystal thermometers

Two EL II prospective cohort studies[61,62] and two EL III studies[63,64] investigating the diagnostic accuracy of forehead measurement were found. These studies varied at baseline. For example, one[61] recruited patients aged 0–14 years, the other[62] had children aged 12 days to 17 years. The authors also used different references for comparisons. For example, one study[62] compared forehead temperature with either rectal temperature (< 4 years) or oral temperature (> 4 years) measured by mercury glass thermometer and another[64] oral temperature measured by digital thermometer. The limited data suggest that forehead measurement underestimated body temperature by 1.2 °C on average.

Infrared tympanic thermometers

Two EL II SRs[30,65] and 21 prospective cohort studies (two EL Ib,[66,67] eight EL II[38,40,43,68–72] and ten EL III studies[73–83]) investigating the diagnostic accuracy of tympanic temperature measurement were found. The SR[30] included 4441 children aged 0–16 years. Other prospective cohort studies[38,40,43,66–82] had very different baselines in terms of sampling frame, age, condition of children recruited and method of temperature measurement. For instance, one study[66] recruited children aged 0–18 years

from a paediatric clinic, another study[77] recruited injured children aged 1–14 years, and another recruited babies from a well-baby nursery.[69] Based on pooled analysis, tympanic measurement differs on average from body temperature by 0.29°C.[30] The difference between tympanic temperature and body temperature can be up to 0.74°C below to 1.34°C[30] above and this varies with age, mode, environment temperature and devices. Moreover, the pooled estimates of sensitivity and specificity from random effect models were 63.7% (95% CI 55.6% to 71.8%) and 95.2% (93.5% to 96.9%).[30] Refer to the evidence tables on the accompanying CD-ROM for details.

Some studies[67,69] suggested that tympanic thermometers were unreliable in infants under 3 months because of difficulties in ensuring that the probe is correctly positioned in the ear canal. The GDG was unable to achieve consensus on the cut-off point of age using tympanic thermometers and thus this issue was put forward for Delphi consensus. The background information and statement below were put to the Delphi panel.

Background

These thermometers use a probe in the ear canal to measure the temperature of the eardrum. Infrared tympanic thermometers are licensed for use in people of all ages, including babies and young children. Some researchers and many users have suggested that tympanic thermometers may be inaccurate in babies under the age of 3 months because it is difficult to ensure that the probe is correctly positioned. Other researchers have found that tympanic thermometers can be used reliably in children of all ages as long as the user ensures that the ear canal is straight and the probe is pointing at the eardrum. In young babies this is achieved by tugging gently on the outer ear.

Delphi statement 7.1

Infrared tympanic thermometers can be used in babies under the age of 3 months as long as it is ensured that the probe is positioned correctly.

The following responses were obtained from two rounds of the Delphi process (see section 3.2).

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
11 (21%)	8 (15%)	28 (54%)	5 (10%)		52	7

There was no consensus for this statement.

Temporal artery thermometers

Only one EL III prospective cohort study[84] meeting the inclusion criteria investigating the accuracy of temporal artery thermometers was found. The researchers recruited 332 parents with children under 2 years and there were 327 sets of complete data. They found that the temporal artery thermometer detected 81% rectal temperature = 38.0°C, 88% (89/101) rectal temperature = 38.3°C.

Evidence summary

Axillary temperature

On average, axillary temperature measurement using an electronic thermometer underestimates body temperature by at least 0.5°C. There is also a wide range of variation in the difference between axillary and body temperature between individuals. The difference can be as much as 2°C in some children. In different EL Ib and EL II studies, the axillary route has variable sensitivities for detecting fever compared with the rectal or oral routes (25–89%). (EL II)

In neonates the axillary route appears to be more accurate, with a difference from rectal temperature of around 0.5°C. (EL II) In the one study to report the ability to detect fever in neonates, the axillary route was reported to have a sensitivity of 98%. (EL II)

Chemical dot thermometers (axillary route)

Three EL II studies that reported on the use of chemical dot thermometers in children were found. Axillary temperatures were measured in all three studies. The studies varied in terms of settings, the ages of children included and the methods of analysis. Only two of the studies assessed ability to detect fever. Given the above limitations, the accuracy of chemical dot thermometers is usually

reported to be comparable with other thermometers used in the axilla. In the one study to compare the ability to detect fever against rectal temperature, the sensitivity was 68%. (EL II)

Tympanic temperature (by infrared thermometer)

Tympanic measurement differs on average from body temperature by 0.3°C. From EL Ib and EL II studies the difference between tympanic temperature and body temperature can be up to 0.74°C below to 1.34°C above and this varies with age, mode, environment temperature and device. The sensitivity to detect fever ranged from 51% to 97% in these studies.

Some studies reported that tympanic measurements are difficult or inaccurate in infants under the age of 3 months. Other studies reported that the technique could be used in infants of all ages, including neonates. A statement that tympanic measurements should not be used in infants under the age of 3 months was put to the Delphi panel. Consensus was not attained.

Forehead temperature (by chemical thermometer)

Data on the measurement of forehead temperature is sparse. The limited data suggests that forehead measurement appears to be inaccurate (underestimates body temperature by 1.2°C on average). (EL II) Forehead thermometers may be poor at detecting fever (sensitivity 27–88%). (EL II)

Temporal artery temperature (by infrared thermometer)

Measurement of temporal artery temperature has not been extensively studied. The available data suggest this technique has fair sensitivity (81%) to detect fever. (EL III)

Health economics profile

Cost analysis of thermometers was undertaken for this guideline (chapter 11). The analysis was based on the data from hospital setting as regards the annual number of measurements.[85]The results of the analysis are summarised in Table 5.3. The results are discounted to show the present value of costs which accrue in the future (up to 10 years). The analysis showed that the contact/electronic thermometers are the least costly option when staff costs are not included in the analysis. When the staff cost are included, the total cost of electronic/compact, contact/compact electronic and tympanic thermometers are comparable. Contact/electronic thermometers have a high purchase price but the fact that they can be used repeatedly means that they may be less costly per test than the chemical thermometers, which have a low purchase price but can be used only once (or can be reused only a limited number of times). Since the cost per test is dependent on the volume of tests undertaken, chemical thermometers may be a better use of resources than either electronic thermometer in very low volume settings, such as some primary care providers.

GDG translation

The GDG noted that the alternatives to oral and rectal thermometers can all give inaccurate readings and have variable sensitivity in detecting fever. Taking temperatures by the axillary route using an electronic or chemical dot thermometer underestimates body temperature by 0.5°C on average. Tympanic temperatures measured with an infrared thermometer differ from body temperature by 0.3°C on average. The GDG noted that these three types of measurements had not been compared with each other and therefore decided that they could not recommend one type over another. Data from neonates suggests that axillary measurements are more accurate in this age group and it was therefore decided to recommend this route at that age.

The GDG was aware that some authorities suggest that tympanic measurements are unreliable or impossible to perform in infants under the age of 3 months. The evidence was inconclusive on this issue and when the question was put to the Delphi panel there was no consensus. Accordingly, the GDG felt that they could not suggest age limits on the use of tympanic thermometers. The GDG considered that more research was needed in this area. Moreover, it would be helpful if direct comparisons were made between all of the different thermometers that were recommended for use in young children.

Table 4.1 Estimated 10 year expenditure on thermometers suitable for axillary and tympanic measurement in a large teaching hospital, discounted at 3.5% (see Appendix B for details)

	Chemical (single use)	Chemical (reusable)	Contact/electronic	Contact/compact electronic	Infrared sensing (tympanic)
Minimum priced model (with staff cost)		£12,260,326	£758,535	£4,137,153	£1,064,403
Maximum priced model (with staff costs)		£688,596	£941,610	£877,437	£732,427
Minimum priced model (without staff costs)	£769,177	£173,260	£834,153	£108,131	£930,102
Maximum priced model (without staff costs)	£2,637,178	£371,899	£673,009	£541,865	£598,126

From the health economics estimates, the GDG noted that there was considerable overlap in the estimated costs of most types of thermometers. When staff costs were not included, compact electronic thermometers appeared to be the most cost effective. The health economics analysis was based on the cost of thermometers in an acute care setting, and the best choice of thermometer may differ across different clinical settings, such as primary care or accident and emergency triage. In the acute care setting analysis, when estimated staff costs were included, the costs of electronic, compact electronic and tympanic thermometers were comparable. Single-use chemical thermometers appeared expensive. This is partly because a new thermometer is needed for each measurement and estimated staff costs are very high because they take longer to read than the other types of thermometers. The model assumes that healthcare professionals are not engaged in other activities while waiting to read the thermometer, which may not reflect actual practice and may therefore overestimate the cost. Furthermore, the GDG noted that the economic model uses an assumption of 18 recordings per admission. The GDG decided that single-use chemical thermometers may be a cost-effective choice in situations where repeated measurements are unlikely to be needed.

On the use of temporal artery thermometers, the GDG considered that there was insufficient evidence at present from which to make a recommendation. The GDG did not believe that forehead crystal thermometers were accurate enough to be recommended for use by healthcare professionals.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

4.3 Subjective detection of fever by parents and carers

Not all families own a thermometer and parents and carers often attempt to confirm that their child has a fever by subjective means. This is usually done by placing a hand over the child's forehead or other part of the body surface. Most guidelines and review articles do not refer to subjective methods of detecting fever. The GDG considered it important to determine whether subjective detection of fever is accurate and should be considered a valid entry point into this guideline.

Review question

How accurate is the subjective detection of fever by parents and carers compared with the detection of fever with a thermometer?

Narrative evidence

Five EL II studies,[86–90] one EL III prospective cohort study[91] and one EL III research letter[59] inves- tigating the diagnostic accuracy of subjective measurement to detect fever were found. Overall, most of the studies were conducted in resource-poor settings such as Malawi[88] or Zimbabwe,[59] the age of children included varied (e.g. 2 days to 48 months[87] to 1 month to 18 years[90]) and the authors used different reference standards (for instance, one compared perceived fever with oral temperature = 37.8°C or rectal temperature = 38.3°C measured by either mercury or digital thermometer[86]). The other prospective cohort study[87] used tympanic temperature measured by non-contact tympanic thermometer and rectal temperature by mercury thermometer as standard. The overall finding suggested that parental perceived fever had reasonable diagnostic accuracy with the sensitivity of detection of fever ranging from 74%[86] to 97%[88] and specificity ranging from 19%[88] to 86%[86] in EL II studies. Sensitivities and specificities as high as 94% and 90.6%, respectively, have been reported by EL II studies.[59,91]

Evidence summary

Subjective detection of fever by parents and carers has been relatively well studied but there are no UK studies. The sensitivity of palpation for the detection of fever ranged from 74% to 97%. (EL II). Five of the six studies that quoted specificity gave values between 67% and 91%; the other gave a value of 19%. (EL II)

GDG translation

The GDG noted that, although there had been no direct comparisons, the sensitivity and specificity of detecting fever by palpation were comparable with those reported for axillary and tympanic thermometers. The GDG therefore decided that detection of fever by palpation was probably as good as the other alternatives to oral and rectal temperature measurements. The GDG considered that it was important for these facts to be recognised by healthcare professionals.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Research recommendations

Number	Research recommendation
RR1	Measuring temperature in young babies: tympanic versus axilla electronic versus axilla chemical dot versus temporal artery. [2007]

5 Clinical assessment of children with fever

This section was partially updated in 2013.

Introduction

Concerned parents or carers of young children commonly seek access to healthcare services when their child has a fever.

The initial assessment of the feverish child is very important. The majority of children presenting with fever will have either a self-limiting viral condition or an obvious cause for their fever for which specific treatment can be given. A minority will present with fever with no obvious underlying cause, and a small number of these will have a serious illness.

Initial contact may be made remotely (e.g. by telephone) or the child may present directly to a facility where a face-to-face assessment can take place. Wherever the assessment is carried out, the assessor needs to understand the significance of certain symptoms and signs. A careful and thorough assessment should mean that in the majority of cases:

- the child with a potentially serious illness is recognised and managed appropriately
- the child with a minor self-limiting illness is not burdened with unnecessary medical intervention and the parents/carers are supported with appropriate self-care advice.

5.1 Priorities in the clinical assessment of feverish illness in children

Although most children with a fever will have a self-limiting illness, a minority will have a serious or even life-threatening illness. The over-riding priority for healthcare professionals should be to reduce the mortality of children with feverish illness in the UK. The priorities for healthcare professionals should be to:

- identify any immediately life-threatening features
- assess the child's likelihood of having a serious illness or self-limiting illness, without necessarily diagnosing any one particular condition
- determine a source of the illness to direct appropriate management decisions based upon the results of the assessment.

The clinical assessment is similar wherever it takes place and is described in detail in this chapter. Adaptations will need to be made to the assessment if the child cannot be physically examined or if the parents or caregivers of the child are not present, but the priorities and principles remain the same. Care also needs to be taken when assessing children with learning disabilities, and healthcare professionals should be aware that some features of the traffic light table might not apply to these children. The management of children after assessment, however, will be determined not only by the results of the assessment but also by the facilities available to the healthcare professional (for example a nurse consultant on the phone at NHS Direct, a GP in a surgery or a paediatrician in a hospital). Management is therefore dealt with separately in subsequent chapters.

5.2 Life-threatening features of illness in children

Evidence was sought for symptoms and signs associated with fever which would predict serious illness in young children.

Review question

In children with fever, what signs or combination of symptoms and signs are associated with serious illness or mortality?

Are there any scoring systems that use symptoms and signs in children with fever to predict the risk of serious illness? How accurate are they?

Evidence summary

Although evidence was found to determine risk factors for serious illness, none of the features in isolation or combination were strongly associated with death.

GDG translation

The guideline development group (GDG) felt that recommending a specific list of life-threatening signs could result in under-recognition of cases if such a list was used in isolation. Healthcare providers are trained to follow the principles of the Resuscitation Council (UK) guidelines for resuscitation: i.e. assessment of airway, breathing, circulation and neurological dysfunction.[92] Although the GDG could not find any prospective comparison of using these priorities with any other resuscitation strategy, they have been developed with widespread consultation and are seen as best practice by all those involved in the acute management of children. The GDG agreed with stakeholder input to reinforce the principles to determine life-threatening features. However, the GDG has not produced a specific list of signs as this could have the result of removing the clinical judgement required to assess whether a child has an immediate threat to life.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

5.3 Assessment of risk of serious illness

Introduction

After assessing the presence or absence of immediately life-threatening features in a child with a fever, the next priority for the healthcare professional should be to make a further risk assessment based on the presenting symptoms and signs. Some symptoms and signs lead towards a diagnosis of a specific illness or focus of infection. Other symptoms and signs are non-specific but may indicate the severity of illness. Healthcare professionals need to be able to detect those children with non- specific features of serious illness as well as be able to consider the working diagnosis for each case. Healthcare professionals also need to know when to be reassured that children have a self-limiting illness, with parents or carers needing advice and support rather than the child needing specific treatments or admission to hospital.

Traffic light system

Process used to develop the 2007 traffic light table

For the 2007 guideline the GDG developed an evidence based ‘traffic light’ system to highlight graphically both non-specific and specific symptoms and signs of serious illnesses. The purpose of the traffic light system is to aid clinicians in identifying children who may have a serious illness. The ‘red’ features are the most worrying, followed by the ‘amber’ features, whereas the ‘green’ features are the most reassuring. It is not meant to provide a clear diagnosis of specific serious illness, but to highlight which children need further investigation and monitoring.

Evidence based reviews were undertaken to identify the relationship of individual symptoms and signs and the likely presence of any serious illness. The list of symptoms and signs that were identified

included being drowsy, moderate/severe chest recession, a respiratory rate greater than 60 breaths per minute, capillary refill time, respiratory rate, height of fever, duration of fever and signs of dehydration. The GDG members used their clinical experience to assign these symptoms and signs to the green, amber or red column of the traffic light table.

Evidence based reviews were also undertaken to identify evidence on existing scoring systems which determine the likelihood that serious illness was present. These found two that looked at clinical symptoms and signs rather than laboratory values (the Yale Observation Scale [YOS] and the Young Infant Observation Scale [YIOS]). Although neither scale alone could reliably detect serious illness, the YOS did improve the detection of serious illness when combined with an examination and history taken by a physician. Although designed for use with children under 3 years, the GDG agreed it was reasonable to extrapolate symptoms and signs from the YOS to the table for children up to 5 years. The symptoms and signs from the YOS that were associated with being well were added to the green column of the traffic light table, and symptoms and signs that were correlated with serious illness were added to the red column of the traffic light table (see Table 5.1 below for features of the YOS).

Finally, evidence-based reviews were undertaken to identify symptoms and signs of specific serious illnesses, namely bacterial meningitis, septicaemia, bacteraemia, pneumonia, urinary tract infection, encephalitis (herpes simplex), septic arthritis/osteomyelitis and Kawasaki disease. The most predictive symptoms and signs of these specific serious illnesses were added to the traffic light table.

Table 5.1 The features of the Yale Observation Scale (YOS)

Observation item	Normal = 1	Moderate impairment = 3	Severe impairment = 5
Quality of cry	Strong or none	Whimper or sob	Weak or moaning, high-pitched, continuous cry or hardly responds
Reaction to parent stimulation	Cries briefly or no cry and content	Cries on and off	Persistent cry with little response
State variation	If awake, stays awake or if asleep, awakens quickly	Eyes close briefly when awake or awakens with prolonged stimulation	No arousal and falls asleep
Colour	Pink	Pale extremities or acrocyanosis	Pale or cyanotic or mottled or ashen
Hydration	Skin and eyes normal and moist mucous membranes	Skin and eyes normal and mouth slightly dry	Skin doughy or tented and dry mucous membranes and/or sunken eyes
Response to social overtures	Smiles or alerts (consistently)	Brief smile or alert	No smile, anxious, dull; no alerting to social overtures

Process used for the 2013 traffic light table

The guideline update aimed to reassess the symptoms and signs contained in the 2007 traffic light table to ensure that the evidence supporting their inclusion was up to date, and to explore whether there was new evidence to add any symptoms and signs that were not included in the 2007 traffic light table. The reviews focused on diagnostic usefulness of signs and symptoms. This differed from the 2007 approach that focused on correlations between symptoms and serious illness. Therefore, the updated reviews acted as validation of the original traffic light table.

For each symptom or sign, the data found in the 2013 review was considered along with the GDG's expert opinion regarding the use of a symptom or sign in current clinical practice. Based on both the diagnostic outcome measures (positive likelihood ratio, negative likelihood ratio, sensitivity, specificity, specificity, positive predictive value, and negative predictive value) and the GDG's views, a decision was made whether to: add a new symptom or sign to the traffic light table; move an existing symptom or sign to a different column (for example, from the amber column to the red column); or remove an existing symptom or sign from the traffic light table.

Combinations of symptoms and signs were not considered for the updated reviews as they could be misinterpreted if they were included, and they could not easily be incorporated into the existing traffic

light table. The 2007 review on symptoms and signs of specific serious illnesses was not updated for 2013, and the original section can be found at the end of this chapter (see Section 5.4).

The update review was organised under the headings used in the 2007 traffic light table:

- colour
- activity
- respiratory
- hydration
- other.

An updated review on the Yale Observation Scale was also undertaken to ensure the evidence for its use as the basis of the traffic light table is still valid.

To ensure the recommendations follow a logical sequence, the updated traffic light table is provided here before the evidence and translations. The 2013 updated review is presented in Section 5.4. The recommendations are provided towards the end of this chapter and the reader is advised to refer back to the table whenever it is mentioned.

In summary, the updated review resulted in the following changes to the traffic light table:

- 'A new lump > 2cm' was removed from the table
- 'Bile-stained vomiting' was removed from the table
- 'Age 3–6 months, temperature ≥ 39°C' was moved from the red column to the amber column
- 'Rigors' was added to the table, in the amber column
- 'Tachycardia' was added to the table, in the amber column.

Table 5.2 Traffic light system for identifying risk of serious illness.*

Children with fever and **any** of the symptoms or signs in the 'red' column should be recognised as being at high risk. Similarly, children with fever and any of the symptoms or signs in the 'amber' column and none in the 'red' column should be recognised as being at intermediate risk. Children with symptoms and signs in the 'green' column and none in the 'amber' or 'red' columns are at low risk. The management of children with fever should be directed by the level of risk.

	Green – low risk	Amber – intermediate risk	Red – high risk
Colour (of skin, lips or tongue)	• Normal colour	• Pallor reported by parent/carer	• Pale/mottled/ashen/blue
Activity	• Responds normally to social cues • Content/smiles • Stays awake or awakens quickly • Strong normal cry/not crying	• Not responding normally to social cues • No smile • Wakes only with prolonged stimulation • Decreased activity	• No response to social cues • Appears ill to a healthcare professional • Does not wake or if roused does not stay awake • Weak, high-pitched or continuous cry
Respiratory		• Nasal flaring • Tachypnoea: RR > 50 breaths/ minute, age 6–12 months RR > 40 breaths/ minute, age > 12 months • Oxygen saturation ≤ 95% in air • Crackles in the chest	• Grunting • Tachypnoea: RR > 60 breaths/minute • Moderate or severe chest indrawing

Circulation and hydration	• Normal skin and eyes • Moist mucous membranes	• Tachycardia: > 160 beats/minute, age < 1 year > 150 beats/minute, age 1–2 years > 140 beats/minute, age 2–5 years • CRT ≥ 3 seconds • Dry mucous membranes • Poor feeding in infants • Reduced urine output	• Reduced skin turgor
Other	• None of the amber or red symptoms or signs	• Age 3–6 months, temperature ≥ 39°C • Fever for ≥ 5 days • Rigors • Swelling of a limb or joint • Non-weight bearing limb/not using an extremity	• Age < 3 months, temperature ≥ 38°C • Non-blanching rash • Bulging fontanelle • Neck stiffness • Status epilepticus • Focal neurological signs • Focal seizures

CRT capillary refill time; RR respiratory rate

* This traffic light table should be used in conjunction with the recommendations in this guideline on investigations and initial management in children with fever.

The traffic light table is used throughout the rest of the guideline as a basis for making management decisions based on risk rather than diagnosis. Once a working diagnosis has been reached, the healthcare professionals treating the child should stop using this guideline and follow national/local guidance on the management of the specific condition that has been diagnosed.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

.

5.4 Non-specific symptoms and signs of serious illness

2013 review of symptoms and signs

Review question

What is the value (as shown by likelihood ratios, sensitivity, specificity, positive predictive value and negative predictive value) of the following symptoms and signs, alone or in combination, as initial indications of serious illness?

- abnormal skin or mucosal colour (for example pallor or cyanosis)
- appearing ill to a healthcare professional or parent/carer
- altered responsiveness or cry
- altered breathing (for example nasal flaring, grunting, chest indrawing)
- abnormal respiratory rate, pulmonary (lung) crackles and other sounds
- oxygen desaturation
- dehydration
- prolonged capillary refill time, cold hands and feet
- poor feeding
- persistent fever (5 days or more)
- height of fever
- limb or joint swelling

- unwillingness to bear weight or use a limb
- bulging fontanelle
- rash (blanching or non-blanching)
- focal neurological signs
- focal seizures
- new lumps
- neck stiffness
- vomiting
- status epilepticus (prolonged or continuous fits).

If evidence is found on additional signs and symptoms they will be added to the above list.

Overview of updated review

A literature search was undertaken with no restrictions on date. The bibliographies of existing systematic reviews, including a Health Technology Assessment (HTA) report, were searched for relevant studies (Thompson et al, 2012). A total of 7,977 records were identified. In addition, studies included in the 2007 guideline were reviewed for inclusion in the updated review.

Description of included studies

Fifty-nine studies were identified that were relevant to the 2013 review of symptoms and/or signs, of which 42 were prospective studies (Akpede et al., 1992; Andreola et al., 2007; Baker et al., 1989; Baskin et al., 1992; Berger et al., 1996; Bleeker et al., 2007; Brent et al., 2011; Craig et al., 2010;Crain et al., 1982; Crocker et al., 1985; Factor et al., 2001; Galetto-Lacour et al., 2003; Ghotbi et al., 2009; Haddon et al., 1999; Hewson et al., 2000; Hsiao et al., 2006; Lacour et al., 2001; Morris et al., 2007; McCarthy et al., 1985; Nademi et al., 2001; Newman et al., 2002; Nielsen et al., 2001; Nijman et al., 2001; Owusu-Ofori et al., 2004; Pantell et al., 2004; Pratt et al., 2007; Rabasa Al et al., 2009; Rudinsky et al., 2009; Salleeh et al., 2010; Shaw et al., 1998; Shettigar et al., 2011; Shin et al., 2009; Singhi et al., 1992; Tal et al., 1997; Taylor et al., 1995; Thompson et al., 2009; Trautner et al., 2006; Weber et al., 2003; Wells et al., 2001; Yeboah-Antwi et al., 2008; YICSSG, 2008; Zorc et al., 2005) 17 were retrospective (Alpert et al., 1990; Batra et al., 2011; Bleeker et al., 2001; Bonadio et al., 1994; Chen et al., 2009; Fouzas et al., 2010; Gomez et al., 2010; Gomez et al., 2012; Joffe et al., 1983; Nguyen et al., 2002; Offringa et al., 1992; Olaciregui et al., 2009; Schwartz et al., 2009; Stanley et al., 2005; Stathakis et al., 2007; Teach et al., 1997; Zarkesh et al., 2011) and two studies used data that was collected both prospectively and retrospectively (Mandl et al., 1997;Maniaci et al., 2008).

The smallest study included 92 children (Offringa et al., 2002) and the largest study included 12,807 children (Craig et al., 2010). Studies reported on children of a variety of age ranges, and some studies included children older than 5 years. The settings of the studies varied, including GP surgeries, emergency departments and paediatric wards of general hospitals, emergency departments of paediatric hospitals, tertiary care paediatric units and tertiary care medical centres. The definition of fever used for inclusion ranged from higher than 37.2°C to higher than 41.1°C.

Some of the studies looked at specific illnesses, including bacterial meningitis, bacteraemia, urinary tract infection, pneumonia, meningococcal disease and salmonella enteritis. Some studies looked at a group of diagnoses, for example ‘serious illness’ or ‘serious bacterial infection’.

Twenty-one of the studies were undertaken in the USA (Alpert et al., 1990; Baker et al., 1989; Baskin et al., 1992; Bonadio et al., 1994; Crain et al., 1982; Crocker et al., 1985; Hsiao et al., 2006; Joffe et al., 1983; Mandl et al., 1997; Maniaci et al., 2008; McCarthy et al., 1985; Newman et al., 2002; Nguyen et al., 1984; Pantell et al., 2004; Pratt et al., 2007; Rudinsky et al., 2009; Shaw et al., 1998; Stanley et al., 2005; Teach et al., 1997; Trautner et al., 2006; Zorc et al., 2005), five in Australia (Craig et al., 2010; Haddon et al., 1999; Hewson et al., 2000; Stathakis et al, 2007; Taylor et al., 1995), five in the Netherlands (Berger et al., 1996; Bleeker et al., 2001; Bleeker et al., 2007; Nijman et al., 2011; Offringa et al., 1992), four in the UK (Brent et al., 2011; Nademi et al., 2001; Thompson et al., 2009; Wells et al., 2001), three in India (Batra et al., 2011; Shettigar et al., 2011; Singhi et al., 1992), three in Spain (Gomez et al., 2010; Gomez et al., 2012; Olaciregui et al., 2009), two each in Switzerland (Galetto-Lacour et al., 2003; Lacour et al., 2001), Ghana (Owusu-Ofori et al., 2004; Yeboah-Antwi et al., 2008), Nigeria (Akpede et al., 1992; Rabasa Al et al., 2009), Israel (Schwartz et al., 2009; Tal et al., 1997), and Iran (Ghotbi et al., 2009; Zarkesh et al., 2011), and one each in Bangladesh (Factor et al., 2001), Canada (Salleeh et al., 2010), Denmark (Nielsen et al., 2001), Greece (Fouzas et al., 2010), Italy (Andreola et al., 2007), Papua New Guinea (Morris et al., 2007), South Korea (Shin et al., 2009) and Taiwan (Chen et al., 2009). One study was conducted in Bangladesh, Bolivia, Ghana, India, Pakistan and South Africa

(YICSSG, 2008) and another in Ethiopia, the Gambia, Papua New Guinea and the Philippines (Weber et al., 2003).

Twelve studies were found that reported evidence on the Yale Observational Scale (Andreola et al., 2007; Baker et al., 1990; Bang et al., 2009; Galetto-Lacour et al., 2003; Haddon et al., 1999; Hsiao et al., 2006; McCarthy et al., 1980; McCarthy et al., 1981; McCarthy et al., 1982; Teach et al., 1995; Thayyil et al., 2005; Zorc et al., 1995).

More details on each individual study can be found in the evidence tables.

The GDG is aware of an HTA relevant to this review (Thompson et al., 2012). However, the review was completed before the HTA was published. All relevant studies cited in the HTA were included in this review.

Evidence profiles

The GRADE profiles in the tables that follow show results of included studies for various aspects of the review question.

- Table 5.3 – evaluation of colour
- Table 5.4 – evaluation of social cues
- Table 5.5 – evaluation of 'appears ill to a healthcare professional or parent/carer'
- Table 5.6 – evaluation of awake
- Table 5.7 – evaluation of decreased activity
- Table 5.8 – evaluation of no smile and/or abnormal cry
- Table 5.9 – evaluation of irritability
- Table 5.10 – evaluation of decreased consciousness/coma
- Table 5.11 – evaluation of restlessness
- Table 5.12 – evaluation of tachypnoea
- Table 5.13 – evaluation of crackles
- Table 5.14 – evaluation of respiratory symptoms
- Table 5.15 – evaluation of nasal symptoms
- Table 5.16 – evaluation of wheeze
- Table 5.17 – evaluation of chest findings/abnormal chest sounds
- Table 5.18 – evaluation of cough
- Table 5.19 – evaluation of poor feeding
- Table 5.20 – evaluation of capillary refill time
- Table 5.21 – evaluation of reduced urine output
- Table 5.22 – evaluation of duration of fever
- Table 5.23 – comparison of duration of fever
- Table 5.24 – evaluation of height of fever in children younger than 3 months
- Table 5.25 – evaluation of height of fever in all ages up to 5 years, including those less than 3 months
- Table 5.26 – comparison of height of fever in children with and without serious illness – all ages up to 5 years
- Table 5.27 – evaluation of bulging fontanelle
- Table 5.28 – evaluation of neck stiffness
- Table 5.29 – evaluation of focal seizures
- Table 5.30 – evaluation of non-blanching rash
- Table 5.31 – evaluation of diarrhoea
- Table 5.32 – evaluation of vomiting
- Table 5.33 – evaluation of abdominal pain

- Table 5.34 – evaluation of crying on micturition/dysuria
- Table 5.35 – evaluation of headache
- Table 5.36 – evaluation of conjunctivitis
- Table 5.37 – evaluation of poor peripheral circulation
- Table 5.38 – evaluation of bulging abdomen
- Table 5.39 – evaluation of paresis or paralysis
- Table 5.40 – evaluation of abnormal neurological findings
- Table 5.41 – evaluation of impression of tone
- Table 5.42 – evaluation of tenderness on examination
- Table 5.43 – evaluation of urinary symptoms
- Table 5.44 – evaluation of abnormal ear, nose and throat signs
- Table 5.45 – evaluation of rigors and/or chills
- Table 5.46 – evaluation of Yale Observation Scale
- Table 5.47 – comparison of Yale Observation Scores

Table 5.3 GRADE profile for evaluation of colour

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Colour (cyanotic or pale or flushed/mottled)								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	36 (20 to 53)[a]	40 (31 to 49)[a]	16 (8 to 24)[a]	67 (55 to 78)[a]	0.6 (0.4 to 1.0)[a]	1.6 (1.1 to 2.3)[a]	Low
For detecting urinary tract infection								
1 (Newman, 2002)	1666	9 (5 to 14)[a]	92 (90 to 93)[a]	11 (6 to 16)[a]	90 (89 to 92)[a]	1.1 (0.7 to 1.8)[a]	1.0 (0.9 to 1.0)[a]	Low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.4 GRADE profile for evaluation of social cues

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Decreased social interaction								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	24 (17 to 30)[a]	74 (71 to 76)[a]	9 (6 to 11)[a]	90 (88 to 92)[a]	0.9 (0.7 to 1.2)[a]	1.0 (0.9 to 1.1)[a]	Low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.5 GRADE profile for evaluation of 'appears ill to a healthcare professional or parent/carer'

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
At least mildly unwell (includes mildly unwell, moderately unwell and very unwell)								
For detecting urinary tract infection, pneumonia or bacteraemia								
1 (Craig, 2010)	12807	74 (72 to 77) [a]	42 (41 to 43) [a]	9 (9 to 10) [a]	95 (95 to 96) [a]	1.3 (1.2 to 1.3) [a]	0.6 (0.5 to 0.7) [a]	Low
At least moderately ill or moderately unwell (includes moderately ill/unwell and very ill/unwell)								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	58 (41 to 74) [a]	70 (61 to 78) [a]	37 (24 to 51) [a]	84 (76 to 92) [a]	1.9 (1.3 to 2.9) [a]	0.6 (0.4 to 0.9) [a]	Low
For detecting urinary tract infection, pneumonia or bacteraemia								
1 (Craig, 2010)	12807	22 (20 to 25) [a]	92 (91 to 92) [a]	17 (15 to 19) [a]	94 (93 to 94) [a]	2.7 (2.4 to 3.0) [a]	0.8 (0.8 to 0.9) [a]	Low
For detecting urinary tract infection								
1 (Newman, 2002)	1666	38 (30 to 45) [a]	65 (62 to 67) [a]	10 (8 to 13) [a]	91 (89 to 92) [a]	1.1 (0.9 to 1.3) [a]	1.0 (0.8 to 1.1) [a]	Low
For detecting occult infections								
1 (Pantell, 2004)	3066	NC	NC	NC	NC	NC	NC	Low
Not well-appearing								
For detecting serious bacterial infection								
1 (Gomez, 2010)	1018	26 (8 to 44) [a]	96 (95 to 97) [a]	13 (3 to 22) [a]	98 (97 to 99) [a]	6.2 (2.9 to 13.1) [a]	0.8 (0.6 to 1.0) [a]	Very low
Appears unwell								
For detecting serious bacterial infection								
1 (Nijuman, 2012)	1255	1 (0 to 4) [a]	97 (96 to 98) [a]	6 (0 to 15) [a]	89 (88 to 91) [a]	0.6 (0.1 to 2.5) [a]	1.0 (1.0 to 1.0) [a]	Very low
Poor appearance								
For detecting serious bacterial infection								
1 (Chen, 2009)	135	35 (19 to 51) [a]	82 (75 to 90) [a]	40 (22 to 58) [a]	79 (71 to 87) [a]	2.0 (1.1 to 3.6) [a]	0.8 (0.6 to 1.0) [a]	Very low
Ill appearance								
For detecting serious illness								
1 (McCarthy, 1985)	103	54 (35 to 73) [a]	90 (83 to 96) [a]	64 (44 to 84) [a]	85 (77 to 93) [a]	5.2 (2.5 to 10.9) [a]	0.5 (0.3 to 0.8) [a]	Very low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
1 (Baker, 1989)	190	47 (21 to 72) [a]	90 (80 to 99) [a]	64 (35 to 92) [a]	81 (70 to 93) [a]	4.6 (1.6 to 13.3) [a]	0.6 (0.4 to 1.0) [a]	Very low
For detecting serious invasive bacteraemia								
1 (Mandl, 1997)	411	100 (60 to 100)	88 (86 to 91)	11 (1 to 23)	100 (97 to 100)	8.6 (6.6 to 11.2) [a]	NC	Very low
For detecting serious bacterial infection								
1 (Schwartz, 2009)	449	21 (12 to 29) [a]	90 (87 to 93) [a]	33 (20 to 45) [a]	82 (79 to 86) [a]	2.0 (1.2 to 3.4) [a]	0.9 (0.8 to 1.0) [a]	Very low
1 (Shin, 2009)	221	37 (22 to 51) [a]	69 (62 to 76) [a]	22 (12 to 32) [a]	82 (76 to 88) [a]	1.2 (0.7 to 1.9) [a]	0.9 (0.7 to 1.2) [a]	Low
For detecting urinary tract infection								
1 (Shaw, 1998)	2411	49 (38 to 60) [a]	72 (71 to 74) [a]	6 (4 to 7) [a]	98 (97 to 98) [a]	1.8 (1.4 to 2.2) [a]	0.7 (0.6 to 0.9) [a]	Low
Very ill or very unwell appearance								
For detecting urinary tract infection, pneumonia or bacteraemia								
1 (Craig, 2010)	12807	3 (2 to 3) [a]	100 (100 to 100) [a]	45 (33 to 58) [a]	93 (93 to 93) [a]	10.6 (6.5 to 17.3) [a]	1.0 (1.0 to 1.0) [a]	Low
For detecting urinary tract infection								
1 (Newman, 2002)	1666	4 (1 to 7) [a]	97 (97 to 98) [a]	14 (4 to 24) [a]	90 (89 to 92) [a]	1.5 (0.6 to 3.4) [a]	13.6 (3.5 to 23.8) [a]	Low
For detecting occult infections								
1 (Pantell (2004)	3066	NC	NC	NC	NC	NC	NC	Low
Severely ill								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	33 (17 to 49) [a]	90 (85 to 96) [a]	52 (31 to 73) [a]	81 (74 to 88) [a]	3.5 (1.6 to 7.5) [a]	0.7 (0.6 to 0.9) [a]	Low
Toxicity								
For detecting bacterial meningitis								
1 (Ghotbi, 2009)	254	33 (7 to 60) [a]	97 (94 to 99) [a]	33 (7 to 60) [a]	97 (94 to 99) [a]	10.1 (3.5 to 28.8) [a]	0.7 (0.5 to 1.0) [a]	Very low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Suspicious physical findings								
For detecting bacterial meningitis								
1 (Joffe, 1983)	241	23 (16 to 30) [a]	97 (94 to 100) [a]	91 (81 to 100) [a]	52 (45 to 59) [a]	8.5 (2.7 to 27.2) [a]	0.8 (0.7 to 0.9) [a]	Very low

NC Not calculable

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.6 GRADE profile for evaluation of awake

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Drowsy on history or examination								
For detecting serious illness								
1 (Hewson, 2000)	313	51 (40 to 61)[a]	84 (79 to 89)[a]	55 (44 to 66)[a]	82 (77 to 87)[a]	3.2 (2.2 to 4.6)[b]	0.6 (0.5 to 0.7)[b]	Low
Increased sleepiness								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	34 (26 to 41) [b]	74 (71 to 76) [b]	12 (9 to 15) [b]	91 (90 to 93) [b]	1.3 (1.0 to 1.6) [b]	0.9 (0.8 to 1.0) [b]	Low
Drowsiness								
For detecting bacterial meningitis								
1 (Ghotbi, 2009)	254	25 (1 to 50) [b]	100 (100 to 100) [b]	100 (100 to 100) [b]	96 (94 to 99) [b]	NC [b]	0.8 (0.5 to 1.0) [b]	Very low
1 (Offringa, 1992)	92	25 (1 to 50) [b]	74 (64 to 84) [b]	14 (0 to 29) [b]	85 (76 to 94) [b]	1.0 (0.3 to 2.8) [b]	1.0 (0.7 to 1.4) [b]	Very low
Drowsiness at home								
For detecting bacterial meningitis								
1 (Offringa, 1992)	92	30 (12 to 49) [b]	94 (89 to100) [b]	64 (35 to 92) [b]	80 (72 to 89) [b]	5.3 (1.7 to 16.3) [b]	0.7 (0.6 to 1.0) [b]	Very low
Postictal drowsiness								
For detecting bacterial meningitis								
1 (Batra, 2011)	199	60 (17 to 100) [b]	96 (93 to 99) [b]	27 (1 to 54) [b]	99 (97 to 100) [b]	14.6 (5.4 to 39.0) [b]	0.4 (0.1 to 1.2) [b]	Very low

NC Not calculable

[a] Confidence intervals calculated by a member of the technical team at the NCC-WCH based on results reported in the study

[b] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.7 GRADE profile for evaluation of decreased activity

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Decreased activity								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	17 (12 to 23) [a]	82 (80 to 84) [a]	9 (6 to 12) [a]	90 (89 to 92) [a]	0.9 (0.7 to 1.3) [a]	1.0 (0.9 to 1.1) [a]	Low
Decreased activity level during examination								
For detecting bacteraemia								
1 (Crain, 1982)	175	NC	NC	NC	NC	NC	NC	Moderate
Looking around the room (moderately impaired)								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	21 (7 to 35) [a]	69 (60 to 78) [a]	18 (6 to 30) [a]	73 (64 to 82) [a]	0.7 (0.3 to 1.4) [a]	1.1 (0.9 to 1.4) [a]	Low
Looking around the room (severely impaired)								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	30 (15 to 46) [a]	92 (87 to 97) [a]	56 (33 to 79) [a]	81 (73 to 88) [a]	3.9 (1.7 to 9.0) [a]	0.8 (0.6 to 1.0) [a]	Low
Moving arms and legs spontaneously (moderately impaired)								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	27 (12 to 42) [a]	78 (70 to 86) [a]	28 (13 to 44) [a]	77 (69 to 85) [a]	1.2 (0.6 to 2.4) [a]	0.9 (0.7 to 1.2) [a]	Low
Moving arms and legs spontaneously (severely impaired)								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	24 (10 to 39) [a]	96 (93 to 100) [a]	67 (40 to 93) [a]	80 (73 to 87) [a]	6.4 (2.0 to 19.8) [a]	0.8 (0.6 to 1.0) [a]	Low
Reaching for objects (moderately impaired)								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	15 (3 to 27) [a]	77 (69 to 85) [a]	17 (3 to 31) [a]	74 (66 to 83) [a]	0.7 (0.3 to 1.6) [a]	1.1 (0.9 to 1.3) [a]	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Reaching for objects (severely impaired)								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	30 (15 to 46) [a]	90 (85 to 96) [a]	50 (28 to 72) [a]	81 (73 to 88) [a]	3.2 (1.5 to 7.0) [a]	0.8 (0.6 to 1.0) [a]	Low
Lethargy								
For detecting serious bacterial infection								
1 (Shin, 2009)	221	17 (6 to 29) [a]	72 (65 to 79) [a]	13 (4 to 22) [a]	78 (72 to 85) [a]	0.6 (0.3 to 1.2) [a]	1.2 (1.0 to 1.4) [a]	Low
1 (Ghotbi, 2009)	254	42 (14 to 70) [a]	95 (92 to 97) [a]	28 (7 to 48) [a]	97 (95 to 99) [a]	7.8 (3.3 to 18.2) [a]	0.6 (0.4 to 1.0) [a]	Very low
For detecting bacteraemia								
1 (Crocker, 1985)	201	14 (1 to 27) [a]	78 (72 to 84) [a]	10 (1 to 18) [a]	85 (79 to 90) [a]	0.7 (0.3 to 1.7) [a]	1.1 (0.9 to 1.3) [a]	Low

NC Not calculable

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.8 GRADE profile for evaluation of no smile and/or abnormal cry

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Cry								
For detecting serious bacterial infection								
1 (Craig, 2010)	15781	43 (40 to 45) [a]	68 (67 to 68) [a]	9 (9 to 10) [a]	94 (93 to 94) [a]	1.3 (1.2 to 1.4) [a]	0.9 (0.8 to 0.9) [a]	Low
For detecting bacteraemia								
1 (Crain, 1982)	175	NR [b]	NR [b]	NR [b]	NR [b]	NR [b]	NR [b]	Low
Abnormal cry								
For detecting bacteraemia								
1 (Pantell, 2004)	3066	NR [c]	NR [c]	NR [c]	NR [c]	NR [c]	NR [c]	Low

CI confidence interval, NR not reported, OR odds ratio, P probability

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

[b] Text in the study paper stated that crying is not significantly associated with bacteraemia

[c] Adjusted OR 2.23 (95% CI 1.16 to 4.29), $P < 0.02$

Table 5.9 GRADE profile for evaluation of irritability

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Irritability								
For detecting serious bacterial infection								
1 (Shin, 2009)	221	34 (20 to 49) [a]	63 (56 to 70) [a]	18 (10 to 27) [a]	80 (73 to 87) [a]	0.9 (0.6 to 1.5) [a]	1.0 (0.8 to 1.3) [a]	Low
For detecting bacteraemia								
1 (Crain, 1982)	175	NR [b]	NR [b]	NR [b]	NR [b]	NR [b]	NR [b]	Low
1 (Crocker, 1985)	201	64 (47 to 82) [c]	55 (48 to 62) [c]	19 (11 to 27) [c]	91 (85 to 96) [c]	1.4 (1.0 to 2.0) [c]	0.7 (0.4 to 1.1) [c]	Low
For detecting urinary tract infection								
1 (Morris, 2007)	98	NR [d]	NR [d]	NR [d]	NR [d]	NR [d]	NR [d]	Low
For detecting bacterial meningitis								
1 (Ghotbi, 2009)	254	58 (30 to 86) [c]	86 (82 to 90) [c]	17 (6 to 29) [c]	98 (96 to 100) [c]	4.2 (2.3 to 7.3) [c]	0.5 (0.2 to 0.9) [c]	Low
For detecting viral meningitis or non-specific meningitis								
1 (Gomez, 2012)	309	24 (15 to 32) [c]	78 (72 to 84) [c]	34 (23 to 45) [c]	68 (62 to 74) [c]	1.1 (0.7 to 1.7) [c]	1.0 (0.9 to 1.1) [c]	Low

NR not reported

[a] The selection criteria for including children in the study were not clearly described.

[b] Text in the paper stated that irritability is not significantly associated with bacteraemia.

[c] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study.

[d] Text in the paper stated that irritability is not predictive of urinary tract infection.

Table 5.10 GRADE profile for evaluation of decreased consciousness/coma

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Decreased consciousness								
For detecting serious bacterial infection								
1 (Bleeker, 2001)	231	3 (1 to 6) [a]	91 (84 to 99) [a]	55 (25 to 84) [a]	24 (18 to 30) [a]	0.4 (0.1 to 1.3) [a]	1.1 (1.0 to 1.1) [a]	Very low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Coma								
For detecting bacterial meningitis								
1 (Ghotbi, 2009)	254	8 (0 to 24) [a]	100 (100 to 100) [a]	100 (100 to 100) [a]	96 (93 to 98) [a]	NC	0.9 (0.8 to 1.1) [a]	Moderate
1 (Offringa, 1992)	92	26 (8 to 44) [a]	100 (100 to 100) [a]	100 (100 to 100) [a]	80 (72 to 89) [a]	NC	0.7 (0.6 to 0.9) [a]	Very low
Unrousable coma								
For detecting bacterial meningitis								
1 (Akpede, 1992)	522	22 (5 to 40) [a]	94 (92 to 96) [a]	15 (3 to 27) [a]	97 (95 to 98) [a]	3.9 (1.7 to 9.1) [a]	0.8 (0.7 to 1.0) [a]	Low

NC Not calculable

[a] Calculated by the NCC-WCH based on data reported in the study

Table 5.11 GRADE profile for evaluation of restlessness

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Restlessness								
For detecting serious illness								
1 (Nademi, 2001)	141	76 (62 to 88)	43 (33 to 52)	35 (25 to 45)	81 (70 to 91)	1.3 (1.0 to 1.7) [a]	0.6 (0.3 to 1.0) [a]	Very low

[a] Calculated by the NCC based on data reported in the study

Table 5.12 GRADE profile for evaluation of tachypnoea

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Tachypnea								
For detecting pneumonia								
1 (Taylor, 1995)	572	74 (70 to 77)	77 (77 to 80)	20 (17 to 23)	97 (96 to 99)	3.2 (2.5 to 4.0) [a]	0.3 (0.2 to 0.6) [a]	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Elevated respiratory rate								
For detecting bacteraemia								
1 (Craig, 2010)	12807	11 (3 to 19) [a]	85 (84 to 86) [a]	1 (0 to 1) [a]	100 (99 to 100) [a]	0.7 (0.4 to 1.5) [a]	1.0 (1.0 to 1.1) [a]	Low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.13 GRADE profile for evaluation of crackles

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Chest crackles								
For detecting pneumonia, urinary tract infection and bacteraemia								
1 (Craig, 2010)	12807	19 (17 to 22) [a]	93 (92 to 93) [a]	17 (15 to 19) [a]	1 (1 to 1) [a]	2.6 (2.3 to 2.9) [a]	0.9 (0.8 to 0.9) [a]	Low
Abnormal chest sounds								
For detecting pneumonia, urinary tract infection and bacteraemia								
1 (Craig, 2010)	12807	29 (27 to 32) [a]	85 (85 to 86) [a]	13 (12 to 15) [a]	94 (94 to 94) [a]	2.0 (1.8 to 2.2) [a]	0.8 (0.8 to 0.9) [a]	Low
Crepitations								
For detecting serious bacterial illness								
1 (Bleeker, 2001)	231	2 (0 to 5) [a]	93 (87 to 100) [a]	50 (15 to 85) [a]	24 (19 to 30) [a]	0.3 (0.1 to 1.3) [a]	1.0 (1.0 to 1.1) [a]	Very low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.14 GRADE profile for evaluation of respiratory symptoms

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Respiratory distress								
For detecting urinary tract infection								
1 (Newman, 2002)	3066	4 (1 to 8) [a]	92 (90 to 93) [a]	5 (1 to 9) [a]	90 (88 to 91) [a]	0.5 (0.2 to 1.1) [a]	1.0 (1.0 to 1.1) [a]	Moderate

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Breathing difficulty								
For detecting pneumonia, urinary tract infection, or bacteraemia								
1 (Craig, 2010)	12,807	26 (23 to 28) [a]	87 (87 to 88) [a]	13 (12 to 15) [a]	94 (93 to 94) [a]	2.0 (1.8 to 2.2) [a]	0.9 (0.8 to 0.9) [a]	Moder-ate
Breathing difficulty or chest wall recession								
For detecting serious illness								
1 (Hewson, 2000)	313	NR/NC	65 (NR/NC)	41 (NR/NC)	82 (NR/NC)	NR/NC	NR/NC	Moder-ate
Shortness of breath								
For detecting serious bacterial infection								
1 (Njiman, 2012)	1255	27 (20 to 35) [a]	88 (86 to 90) [a]	21 (15 to 27) [a]	91 (90 to 93) [a]	2.2 (1.6 to 3.1) [a]	0.8 (0.7 to 0.9) [a]	Very low
Respiratory symptoms								
For detecting pneumonia, urinary tract infection, or bacteraemia								
1 (Craig, 2010)	12,807	70 (67 to 72) [a]	28 (27 to 28) [a]	7 (7 to 7) [a]	92 (91 to 93) [a]	1.0 (0.9 to 1.0) [a]	1.1 (1.0 to 1.2) [a]	Moder-ate

NR/NC Not reported/not calculable

[a] Results calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.15 GRADE profile for evaluation of nasal symptoms

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Purulent nasal discharge								
For detecting serious bacterial illness								
1 (Bleeker, 2001)	231	20 (14 to 26) [a]	53 (41 to 66) [a]	56 (44 to 69) [a]	18 (13 to 24) [a]	0.4 (0.3 to 0.7) [a]	1.5 (1.2 to 1.9) [a]	Very low
Upper respiratory tract infection or runny nose								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	5 (2 to 8) [a]	90 (88 to 91) [a]	5 (2 to 8) [a]	90 (88 to 91) [a]	0.5 (0.2 to 1.0) [a]	1.1 (1.0 to 1.1) [a]	Moder-ate

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Mild upper respiratory tract infection symptoms								
For detecting serious bacterial infection								
1 (Shin, 2009)	221	5 (0 to 11) [a]	72 (65 to 79) [a]	4 (0 to 9) [a]	76 (69 to 82) [a]	0.2 (0.0 to 0.7) [a]	1.3 (1.2 to 1.5) [a]	Low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.16 GRADE profile for evaluation of wheeze

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Audible wheeze								
For detecting pneumonia, urinary tract infection, or bacteraemia								
1 (Craig, 2010)	12,807	8 (7 to 10) [a]	94 (93 to 94) [a]	9 (7 to 11) [a]	93 (92 to 93) [a]	1.3 (1.1 to 1.6) [a]	1.0 (1.0 to 1.0) [a]	Moder-ate
Stridor								
For detecting pneumonia, urinary tract infection, or bacteraemia								
1 (Craig, 2010)	12,807	1 (1 to 2) [a]	98 (98 to 98) [a]	5 (2 to 7) [a]	93 (92 to 93) [a]	0.6 (0.4 to 1.1) [a]	1.0 (1.0 to 1.0) [a]	Moder-ate

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.17 GRADE profile for evaluation of chest findings/abnormal chest sounds

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Abnormal chest sounds								
For detecting pneumonia, urinary tract infection, or bacteraemia								
1 (Craig, 2010)	12,807	8 (7 to 10) [a]	94 (93 to 94) [a]	9 (7 to 11) [a]	93 (92 to 93) [a]	1.3 (1.1 to 1.6) [a]	1.0 (1.0 to 1.0) [a]	Moder-ate
Chest findings								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	2 (0 to 4)	95 (94 to 96)	4 (0 to 8)	90 (89 to 92)	0.4 (0.1 to 1.2)	1.0 (1.0 to 1.1)	Moder-ate

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.18 GRADE profile for evaluation of cough

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Cough								
For detecting pneumonia, urinary tract infection, or bacteraemia								
1 (Craig, 2010)	12,807	58 (55 to 61) [a]	46 (46 to 47) [a]	8 (7 to 8) [a]	93 (93 to 94) [a]	1.1 (1.0 to 1.1) [a]	0.9 (0.9 to 1.0) [a]	Moder-ate
For detecting urinary tract infection								
1 (Newman, 2002)	1666	1 (0 to 2) [a]	98 (98 to 99) [a]	4 (0 to 11) [a]	90 (89 to 92) [a]	0.4 (0.1 to 2.7) [a]	1.0 (1.0 to 1.0) [a]	Moder-ate
For detecting meningococcal disease								
1 (Nielsen, 2001)	208	15 (4 to 27) [a]	63 (55 to 70) [a]	9 (2 to 15) [a]	76 (69 to 83) [a]	0.4 (0.2 to 0.9) [a]	1.3 (1.1 to 1.6) [a]	Very low

NA Not applicable

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.19 GRADE profile for evaluation of poor feeding

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Poor intake								
For detecting serious bacterial infection								
1 (Bleeker, 2001)	231	36 (29 to 44) [a]	74 (63 to 85) [a]	81 (72 to 90) [a]	28 (21 to 35) [a]	1.4 (0.9 to 2.3) [a]	0.9 (0.7 to 1.0) [a]	Very low
Poor feeding								
For detecting serious disease								
1 (Nademi, 2001)	141	78 (65 to 90)	43 (33 to 52)	36 (25 to 45)	83 (72 to 92)	1.4 (1.1 to 1.7) [a]	0.5 (0.3 to 0.9) [a]	Very low
For detecting serious bacterial infection								
1 (Shin, 2009)	221	27 (13 to 40) [a]	63 (56 to 70) [a]	15 (7 to 23) [a]	78 (71 to 85) [a]	0.7 (0.4 to 1.2) [a]	1.2 (0.9 to 1.4) [a]	Low
Decreased feeding								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	37 (29 to 44) [a]	63 (60 to 65) [a]	9 (7 to 12) [a]	90 (88 to 92) [a]	1.0 (0.8 to 1.2) [a]	1.0 (0.9 to 1.1) [a]	Low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.20 GRADE profile for evaluation of capillary refill time

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Capillary refill time of 2 to 3 seconds								
For detecting pneumonia, urinary tract infection and bacteraemia								
1 (Craig, 2010)	12807	10 (8 to 11) [a]	96 (96 to 96) [a]	17 (14 to 19) [a]	93 (93 to 94) [a]	2.6 (2.1 to 3.1) [a]	0.9 (0.9 to 1.0) [a]	Low
Capillary refill time of > 3 seconds								
For detecting pneumonia, urinary tract infection and bacteraemia								
1 (Craig, 2010)	12807	1 (1 to 2) [a]	100 (100 to 100) [a]	35 (22 to 49) [a]	93 (92 to 93) [a]	7.0 (3.9 to 12.7) [a]	1.0 (1.0 to 1.0) [a]	Low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.21 GRADE profile for evaluation of reduced urine output

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Reduced urine output								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	17 (11 to 23) [a]	86 (85 to 88) [a]	12 (8 to16) [a]	91 (89 to 92) [a]	1.2 (0.8 to 1.8) [a]	1.0 (0.9 to 1.0) [a]	Low
Poor micturition								
For detecting serious bacterial infection								
1 (Bleeker, 2001)	231	33 (26 to 40) [a]	79 (69 to 90) [a]	83 (74 to 92) [a]	28 (22 to 35) [a]	1.6 (0.9 to 2.8) [a]	0.8 (0.7 to 1.0) [a]	Very low

[a] Calculated by NCC-WCH based on results reported in the study

Table 5.22 GRADE profile for evaluation of duration of fever

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Fever duration > 12 hours								
For detecting serious bacterial infection								
1 (Pratt, 2007)	119	44 (22 to 67) [a]	19 (15 to 22) [a]	2 (1 to 3) [a]	90 (85 to 96) [a]	0.5 (0.3 to 0.9) [a]	3.0 (1.9 to 4.7) [a]	Very low
For detecting bacteraemia								
1 (Haddon, 1999)	534	65 (42 to 87) [a]	38 (29 to 48) [a]	15 (7 to 23) [a]	87 (77 to 97) [a]	1.0 (0.7 to 1.5) [a]	0.9 (0.5 to 1.8) [a]	Very low
Fever duration ≥ 24 hours								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	19 (13 to 25) [a]	90 (89 to 92) [a]	17 (11 to 22) [a]	91 (90 to 93) [a]	1.9 (1.3 to 2.7) [a]	0.9 (0.8 to 1.0) [a]	Low
For detecting bacteraemia								
1 (Teach, 1997)	6619	60 (53 to 67) [a]	28 (27 to 30) [a]	2 (2.00 to 3) [a]	96 (95 to 97) [a]	0.8 (0.7 to 0.9) [a]	1.4 (1.2 to 1.7) [a]	Very low
Fever duration >24 hours								
For detecting serious bacterial infection								
1 (Andreola, 2007)	408	52 (42 to 62) [a]	31 (26 to 36) [a]	18 (14 to 23) [a]	69 (61 to 76) [a]	0.8 (0.6 to 0.9) [a]	1.5 (1.2 to 2.0) [a]	Low
Fever duration ≥ 2 days								
For detecting bacteraemia								
1 (Teach, 1997)	6619	18 (12 to 23) [a]	74 (73 to 75) [a]	2 (1 to 3) [a]	97 (96 to 97) [a]	0.7 (0.5 to 0.9) [a]	1.1 (1.0 to 1.2) [a]	Very low
Fever duration > 48 hours								
For detecting serious bacterial infection								
1 (Berger, 1996)	138	39 (23 to 56) [a]	82 (75 to 89) [a]	41 (24 to 58) [a]	81 (74 to 89) [a]	2.2 (1.2 to 3.9) [a]	0.7 (0.6 to 1.0) [a]	Low
1 (Trautner, 2006)	103	NR/NC [b]	NR/NC [b]	NR/NC [b]	NR/NC [b]	NR/NC [b]	NR/NC [b]	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Fever duration ≥ 72 hours								
For detecting urinary tract infection								
1 (Salleeh, 2010)	818	NR/NC [c]	NR/NC [c]	NR/NC [c]	NR/NC [c]	NR/NC [c]	NR/NC [c]	Very low
Fever duration > 3 days								
For detecting serious bacterial infection								
1 (Factor, 2001)	669	25 (21 to 30)[a]	85 (81 to 89) [a]	69 (61 to 76)[a]	47 (42 to 51)[a]	1.7 (1.2 to 2.3)[a]	0.9 (0.8 to 0.9)[a]	Low

CI confidence interval, NR/NC not reported/not calculable, OR odds ratio, P probability, RR risk ratio

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

[b] OR 1.04 (95% CI 0.35 to 3.12)

[c] RR 1.6 (95% 1.2 to 2.1), *P* = 0.002

Table 5.23 GRADE profile for comparison of duration of fever

Number of studies	Duration of fever		Effect	Quality
	With serious bacterial illness/infection (SBI) (Mean)	Without SBI (Mean)	*P* value	
Duration of fever				
For detecting serious bacterial infection				
1 (Hsiao, 2006)	26.5 hours (SD 41.5)	18.6 hours (SD 21.7)	*P* < 0.001	High
1 (Bleeker, 2007)	2.5 days (SD 2.6)	2.6 days (SD 2.3)	NR	High
1 (Lacour, 2001)	Median 27 hours (range 2 to 140)	Median 24 hours (range 2 to 140)	*P* = 0.02	High
1 (Galetto-Lacour, 2003)	Median 48 hours (range 6 to 140)	Median 24 hours (range 1 to 140)	*P* = 0.026	High
1 (Olaciregui, 2009)	18.62 hours (SD 35.8)	13.81 hours (SD 26)	*P* = 0.26	Moder-ate
1 (Bleeker, 2001)	2.6 days (SD 2.2)	3.2 days (SD 2.8)	*P* < 0.15	Moder-ate
1 (Fouzas, 2010)	Median 14 hours (IQR 6 to 29)	Median 14 hours (IQR 6 to 27)	*P* = 0.49	Moder-ate
For detecting meningococcal disease				
1 (Nielsen, 2001)	Median 21 hours (IQR/range NR)	Median 24 hours (IQR/range NR)	*P* not significant	Low

IQR interquartile range, NR not reported, P probability, SBI serious bacterial illness/infection, SD standard deviation

Table 5.24 GRADE profile for evaluation of height of fever in children younger than 3 months

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Temperature ≥ 38.0°C								
For detecting serious bacterial infection								
1 (Stanley, 2005)	5279	100 (100 to 100) [a]	0 (0 to 0) [a]	9 (8 to 10) [a]	NC	1.0 (1.0 to 1.0) [a]	NC	Very low
For detecting sepsis								
1 (Weber, 2003)	3303	NR/NC [b]	NR/NC[b]	NR/NC[b]	NR/NC[b]	NR/NC[b]	NR/NC[b]	Low
For detecting bacterial meningitis								
1 (Weber, 2003)	3303	NR/NC[c]	NR/NC[c]	NR/NC[c]	NR/NC[c]	NR/NC[c]	NR/NC[c]	Low
Temperature > 39.0°C								
For detecting urinary tract infection								
1 (Zorc, 2005)	1025	37 (27 to 47) [a]	81 (78 to 83) [a]	16 (11 to 21) [a]	93 (91 to 95) [a]	2.0 (1.4 to 2.6) [a]	0.8 (0.7 to 0.9) [a]	Moder-ate
Temperature ≥ 39.5°C								
For detecting serious bacterial infection								
1 (Zarkesh, 2011)	202	24 (10 to 37) [a]	76 (70 to 83) [a]	19 (8 to 30) [a]	81 (75 to 87) [a]	1.0 (0.5 to 1.9) [a]	1.0 (0.8 to 1.2) [a]	Low
For detecting occult bacteraemia, urinary tract infection, or bacteraemia								
1 (Gomez, 2010)	1018	26 (8 to 44) [a]	91 (89 to 93) [a]	6 (1 to 11) [a]	98 (97 to 99) [a]	2.8 (1.4 to 5.8) [a]	0.8 (0.6 to 1.0) [a]	Very low
Temperature ≥ 40.0°C								
For detecting bacterial meningitis, bacteraemia, urinary tract infection, or salmonella enteritis								
1 (Bonadio, 1994)	356	21 (7 to 35) [a]	96 (94 to 98) [a]	35 (14 to 56) [a]	92 (89 to 95) [a]	5.3 (2.3 to 12.3) [a]	0.8 (0.7 to 1.0) [a]	Very low
1 (Stanley, 2005)	5279	7 (5 to 10) [a]	99 (99 to 99) [a]	38 (28 to 48) [a]	91 (91 to 92) [a]	6.1 (4.1 to 9.3) [a]	0.9 (0.9 to 1.0) [a]	Very low

CI confidence interval, NC not calculable, NR/NC not reported/not calculable, OR odds ratio

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

[b] OR 3.6 (95% CI 2.6 to 5.1)

[c] OR 11.8 (95% CI 5.7 to 24.6)

Table 5.25 GRADE profile for evaluation of height of fever in all ages up to 5 years, including those less than 3 months

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Temperature ≥ 37.4°C								
For detecting urinary tract infection								
1 (Shettigar, 2011)	334	100 (100 to 100) [a]	0 (0 to 0) [a]	8 (5 to 11) [a]	NC	1.0 (1.0 to 1.0) [a]	NC	Moderate
Temperature ≥ 37.5°C								
For detecting serious bacterial infection								
1 (Brent, 2011)	1716	61 (49 to 72)	65 (62 to 67)	7 (5 to 9)	2 (2 to 3)	1.7 (0.7 to 4.5)	0.6 (0.2 to 1.6)	Very low
For detecting malaria or bacterial meningitis								
1 (Owusu-Ofori, 2004)	608	75 (67 to 83) [a]	21 (8 to 34) [a]	74 (66 to 82) [a]	22 (8 to 35) [a]	0.9 (0.8 to 1.1) [a]	1.2 (0.6 to 2.4) [a]	Low
For detecting serious illness								
1 (Yeboah-Antwi, 2008)	685	NR/NC [b]	NR/NC [b]	NR/NC [b]	NR/NC [b]	NR/NC [b]	NR/NC [b]	Low
1 (Yeboah-Antwi, 2008)	685	NR/NC [c]	NR/NC [c]	NR/NC [c]	NR/NC [c]	NR/NC [c]	NR/NC [c]	Low
1 (Yeboah-Antwi, 2008)	685	NR/NC [d]	NR/NC [d]	NR/NC [d]	NR/NC [d]	NR/NC [d]	NR/NC [d]	Low
For detecting severe illness requiring hospitalisation								
1 (YICSSG, 2008)	8889	NR/NC [e]	NR/NC [e]	NR/NC [e]	NR/NC [e]	NR/NC [e]	NR/NC [e]	Low
1 (YICSSG, 2008)	8889	NR/NC [f]	NR/NC [f]	NR/NC [f]	NR/NC [f]	NR/NC [f]	NR/NC [f]	Low
1 (YICSSG, 2008)	8889	NR/NC [g]	NR/NC [g]	NR/NC [g]	NR/NC [g]	NR/NC [g]	NR/NC [g]	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Temperature > 37.5°C								
For detecting bacterial meningitis								
1 (Wells, 2001)	218	79 (63 to 95)	55 (48 to 62)	18 (11 to 25)	95 (88 to 100)	1.7 (1.3 to 2.3) [a]	0.4 (0.2 to 0.8) [a]	Low
Temperature ≥ 38.0°C								
For detecting pneumonia, urinary tract infection or bacteraemia								
1 (Craig, 2010)	12807	85 (83 to 87) [a]	22 (22 to 23) [a]	8 (7 to 8) [a]	95 (94 to 96) [a]	1.1 (1.1 to 1.1) [a]	0.7 (0.6 to 0.8) [a]	Low
For detecting urinary tract infection								
1 (Newman, 2002)	1666	100 (100 to 100) [a]	0 (0 to 0) [a]	10 (8 to 11) [a]	NC	1.0 (1.0 to 1.0) [a]	NC	Moderate
For detecting bacteraemia or bacterial meningitis								
1 (Pantell, 2004)	3066	90 (83 to 98) [a]	29 (28 to 31) [a]	3 (2 to 3) [a]	99 (99 to 100) [a]	1.3 (1.2 to 1.4) [a]	0.3 (0.2 to 0.7) [a]	Low
For detecting serious bacterial infection								
1 (Stanley, 2005)	5279	100 (100 to 100) [a]	0 (0 to 0) [a]	9 (8 to 10) [a]	NC	1.0 (1.0 to 1.0) [a]	NC	Very low
For detecting sepsis								
1 (Weber, 2003)	3303	NR/NC [h]	NR/NC [h]	NR/NC [h]	NR/NC [h]	NR/NC [h]	NR/NC [h]	Low
For detecting bacterial meningitis								
1 (Weber, 2003)	3303	NR/NC [i]	NR/NC [i]	NR/NC [i]	NR/NC [i]	NR/NC [i]	NR/NC [i]	Low
Temperature ≥ 38.4°C								
For detecting urinary tract infection								
1 (Shettigar, 2011)	334	78 (62 to 93)	41 (36 to 47)	10 (6 to 15)	95 (92 to 99)	1.3 (1.1 to 1.6)	0.5 (0.3 to 1.1)	Moderate
Temperature ≥ 38.5°C								
For detecting serious bacterial infection								
1 (Brent, 2011)	1716	38 (27 to 50)	85 (83 to 86)	9 (6 to 13)	97 (96 to 98)	2.5 (1.1 to 5.7)	0.7 (0.3 to 1.7)	Very low
For detecting urinary tract infection								
1 (Newman, 2002)	1666	77 (71 to 84) [a]	38 (35 to 40) [a]	12 (10 to 14) [a]	94 (92 to 96) [a]	1.2 (1.1 to 1.4) [a]	0.6 (0.5 to 0.8) [a]	Moderate

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
For detecting bacteraemia or bacterial meningitis								
1 (Pantell, 2004)	3066	61 (48 to 73) [a]	69 (67 to 71) [a]	4 (3 to 5) [a]	99 (98 to 99) [a]	2.0 (1.6 to 2.4) [a]	0.6 (0.4 to 0.8) [a]	Low
Temperature > 38.5°C								
For detecting bacterial meningitis								
1 (Wells, 2001)	218	58 (39 to 78)	81 (75 to 86)	27 (15 to 40)	94 (88 to 100)	3.1 (2.0 to 4.8) [a]	0.5 (0.3 to 0.8) [a]	Low
Temperature ≥ 39.0°C								
For detecting urinary tract infection								
1 (Shaw, 1998)	2411	79 (70 to 88) [a]	33 (31 to 35) [a]	9 (3 to 5) [a]	98 (97 to 99) [a]	1.2 (1.0 to 1.3) [a]	0.6 (0.4 to 1.0) [a]	Moderate
1 (Newman, 2002)	1666	40 (32 to 47) [a]	74 (72 to 76) [a]	14 (11 to 17) [a]	92 (90 to 94) [a]	1.5 (1.2 to 1.9) [a]	0.8 (0.7 to 0.9)	Moderate
For detecting serious infection								
1 (Thompson, 2009)	700	27 (22 to 32)	87 (84 to 91)	41 (30 to 51) [a]	82 (78 to 85) [a]	2.1 (1.5 to 2.9)	0.8 (0.8 to 0.9)	Low
For detecting pneumonia, urinary tract infection or bacteraemia								
1 (Craig, 2010)	12807	54 (51 to 57) [a]	58 (58 to 59) [a]	9 (9 to 10) [a]	94 (94 to 95) [a]	1.3 (1.2 to 1.4) [a]	0.8 (0.7 to 0.8) [a]	Low
For detecting bacteraemia or bacterial meningitis								
1 (Pantell, 2004)	3066	16 (7 to 26) [a]	90 (88 to 91) [a]	3 (1 to 5) [a]	98 (97 to 99) [a]	1.6 (0.9 to 2.8) [a]	0.9 (0.8 to 1.0) [a]	Low
Temperature > 39.0°C								
For detecting serious disease								
1 (Nademi, 2001)	141	14 (3 to 25)	82 (74 to 89)	25 (7 to 42)	70 (61 to 78)	0.8 (0.3 to 1.9) [a]	1.0 (0.9 to 1.2) [a]	Very low
For detecting bacteraemia or bacterial meningitis								
1 (Pantell, 2004)	3066	43 (31 to 55) [a]	81 (79 to 82) [a]	4 (3 to 6) [a]	99 (98 to 99) [a]	2.2 (1.7 to 3.0) [a]	0.7 (0.6 to 0.9) [a]	Low
For detecting urinary tract infection								
1 (Zorc, 2005)	1025	37 (27 to 47) [a]	81 (78 to 83) [a]	16 (11 to 21) [a]	93 (91 to 95) [a]	2.0 (1.4 to 2.6) [a]	0.8 (0.7 to 0.9) [a]	Moderate

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Temperature ≥ 39.1°C								
For detecting pneumonia, urinary tract infection, bacterial meningitis, or bacteraemia								
1 (Rudinsky, 2009)	985	83 (75 to 88)	18 (16 to 21)	13 (11 to 15) [a]	88 (83 to 92) [a]	1.0 (0.9 to 1.1)	0.9 (0.6 to 1.4)	Low
For detecting bacteraemia, bacterial meningitis, urinary tract infection, or pneumonia								
1 (Alpert, 1990)	152	100 (100 to 100) [a]	0 (0 to 0) [a]	14 (9 to 18) [a]	NC	1.0 (1.0 to 1.0) [a]	NC	Very low
Temperature > 39.3°C								
For detecting urinary tract infection								
1 (Shettigar, 2011)	334	33 (16 to 51) [a]	85 (81 to 89) [a]	16 (6 to 26) [a]	93 (91 to 96) [a]	2.2 (1.2 to 3.9) [a]	0.8 (0.6 to 1.0) [a]	Moderate
Temperature ≥ 39.4°C								
For detecting pneumonia, urinary tract infection, bacterial meningitis, or bacteraemia								
1 (Rudinsky, 2009)	985	67 (59 to 75)	36 (33 to 39)	14 (11 to 16) [a]	88 (85 to 91) [a]	1.1 (0.9 to 1.2)	0.9 (0.7 to 1.2)	Low
Temperature ≥ 39.5°C								
For detecting serious bacterial infection								
1 (Zarkesh, 2011)	202	24 (10 to 37) [a]	76 (70 to 83) [a]	19 (8 to 30) [a]	81 (75 to 87) [a]	1.0 (0.5 to 1.9) [a]	1.0 (0.8 to 1.2) [a]	Low
For detecting occult bacteraemia, urinary tract infection, or bacteraemia								
1 (Gomez, 2010)	1018	26 (8 to 44) [a]	91 (89 to 93) [a]	6 (1 to 11) [a]	98 (97 to 99) [a]	2.8 (1.4 to 5.8) [a]	0.8 (0.6 to 1.0) [a]	Very low
For detecting urinary tract infection								
1 (Newman, 2002)	1666	19 (13 to 25) [a]	92 (91 to 94) [a]	21 (15 to 28) [a]	91 (90 to 93) [a]	2.5 (1.8 to 3.6) [a]	0.9 (0.8 to 0.9) [a]	Moderate
For detecting bacteraemia or bacterial meningitis								
1 (Pantell, 2004)	3066	NR/NC [l]	NR/NC [l]	NR/NC [l]	NR/NC [l]	NR/NC [l]	NR/NC [l]	Low
Temperature > 39.5°C								
For detecting serious disease								
1 (Nademi, 2001)	141	7 (0 to 15)	93 (87 to 98)	30 (1 to 58)	71 (63 to 78)	1.0 (0.3 to 3.8) [a]	1.0 (0.9 to 1.1) [a]	Very low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Temperature ≥ 40.0°C								
For detecting bacterial meningitis, bacteraemia, urinary tract infection, or salmonella enteritis								
1 (Bonadio, 1994)	356	21 (7 to 35) [a]	96 (94 to 98) [a]	35 (14 to 56) [a]	92 (89 to 95) [a]	5.3 (2.3 to 12.3) [a]	0.8 (0.7 to 1.0) [a]	Very low
For detecting pneumonia, urinary tract infection, bacterial meningitis, or bacteraemia								
1 (Rudinsky, 2009)	985	29 (22 to 38)	70 (67 to 73)	13 (9 to 16) [a]	87 (84 to 89) [a]	1.0 (0.8 to 1.3)	1.0 (0.9 to 1.1)	Low
For detecting pneumonia, urinary tract infection, or bacteraemia								
1 (Craig, 2010)	12807	15 (13 to 17) [a]	89 (89 to 90) [a]	10 (8 to 11) [a]	93 (93 to 94) [a]	1.4 (1.2 to 1.6) [a]	1.0 (0.9 to 1.0) [a]	Low
Temperature > 40.0°C								
For detecting serious bacterial infection								
1 (Stanley, 2005)	5279	7 (5 to 10) [a]	99 (99 to 99) [a]	38 (28 to 48) [a]	91 (91 to 92) [a]	6.1 (4.1 to 9.3) [a]	0.9 (0.9 to 1.0) [a]	Very low
Temperature ≥ 40.1°C								
For detecting bacteraemia, bacterial meningitis, urinary tract infection, or pneumonia								
1 (Alpert, 1990)	152	71 (55 to 87) [a]	34 (27 to 41) [a]	14 (9 to 20) [a]	88 (81 to 95) [a]	1.1 (0.8 to 1.4) [a]	0.9 (0.5 to 1.5) [a]	Very low
Temperature ≥ 41.1°C								
For detecting bacteraemia, bacterial meningitis, urinary tract infection, or pneumonia								
1 (Alpert, 1990)	152	45 (28 to 63) [a]	69 (62 to 75) [a]	18 (10 to 27) [a]	89 (84 to 94) [a]	1.4 (0.9 to 2.2) [a]	0.8 (0.6 to 1.1) [a]	Very low

CI confidence interval, NA not applicable, NR/NC not reported/not calculable, OR odds ratio

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

[b] OR 7.4 (95% CI 3.0 to 18.5)

[c] OR 11.1 (95% CI 5.2 to 24.1)

[d] OR 7.4 (95% CI 2.8 to 19.5)

[e] OR 4.7 (95% CI 2.8 to 8.0)

[f] OR 7.5 (95% CI 5.0 to 11.4)

[g] OR 3.4 (95% CI 2.4 to 4.9)

[h] OR 3.6 (95% CI 2.6 to 5.1)

[i] OR 11.8 (95% CI 5.7 to 24.6)

[l] Adjusted OR 3.61 (95% CI 1.40 to 9.25)

Table 5.26 GRADE profile for comparison of height of fever in children with and without serious illness – all ages up to 5 years

Number of studies	Height of fever		Effect	Quality
	With serious bacterial illness/infection (SBI) (°C, mean)	Without SBI (°C, mean)	*P* value	
Height of fever				
For detecting serious bacterial infection				
1 (Baskin, 1992)	39.0 (SD 0.6)	38.9 (SD 0.6)	*P* = 0.01	High
1 (Galetto-Lacour, 2003)	Median 39.4 (38.3 to 41)	Median 39.5 (38 to 40.8)	*P* value 'not significant'	High
1 (Hsiao, 2006)	38.4 (SD 0.8)	38.5 (SD 1.0)	*P* = 0.178	High
1 (Lacour, 2001)	39.1 (SD 0.2)	39.0 (SD 0.1)	*P* value 'not significant'	High
1 (Shin, 2009)	38.7 (SD 0.5)	38.6 (SD 0.4)	*P* = 0.34	High
1 (Andreola, 2007)	39.2 (SD 0.8)	39.0 (SD 0.8)	*P* = 0.004	Moderate
1 (Fouzas, 2010)	Median 38.5 (IQR 38.1 to 39.0)	Median 38.5 (IQR 38.1 to 38.8)	*P* = 0.22	Moderate
1 (Nijman, 2011)	Median 39.3 (IQR 38.6 to 39.8)	Median 38.9 (IQR 38.1 to 39.6)	*P* < 0.000	Moderate
1 (Olaciregui, 2009)	38.23 (SD 0.82)	38.23 (SD 0.64)	*P* = 0.58	Moderate
1 (Maniaci, 2008)	38.9 (SD 0.72)	38.6 (SD 0.45)	*P* = 0.003	Low
1 (Nguyen, 1984)	39.9 (SD 0.96)	39.1 (SD 3.0)	*P* > 0.2	Low
For detecting bacteraemia				
1 (Crocker, 1985)	40.0 (SD 0.4)	40.1 (SD 0.3)	*P* value 'not significant'	High
1 (Haddon, 1999)	39.7 (SD 0.39)	39.7 (SD 0.55)	*P* = 0.91	High
1 (Singhi, 1992)	38.8 (SD 0.3)	38.8 (SD 0.15)	NR	High
1 (Singhi, 1992)	38.7 (SD 0.2)	38.8 (SD 0.15)	NR	High
1 (Teach, 1997)	40.0 (SD 0.61)	39.8 (SD 0.55)	*P* < 0.001	High

Number of studies	Height of fever		Effect	Quality
	With serious bacterial illness/infection (SBI) (°C, mean)	Without SBI (°C, mean)	*P* value	
1 (Stathakis, 2007)	39.0 (SD 0.9)	38.8 (SD 1.0)	*P* = 0.80	Moderate
For detecting meningococcal disease				
1 (Nielsen, 2001)	Median 40 (IQR/range not reported)	Median 39 (IQR/range not reported)	*P* < 0.01	High
For detecting pneumonia, urinary tract infection, bacterial meningitis, or bacteraemia				
1 (Rudinsky, 2009)	103.3°F (SD 1.2)	103.2°F (SD 1.2)	*P* = 0.26	Moderate
For detecting urinary tract infection				
1 (Singhi, 1992)	38.8 (SD 0.1)	38.8 (SD 0.15)	NR	High

IQR interquartile range, NR not reported, P probability, SBI serious bacterial illness/infection, SD standard deviation

Table 5.27 GRADE profile for evaluation of bulging fontanelle

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Bulging fontanelle								
For detecting serious bacterial illness								
1 (Bleeker, 2001)	231	5 (2 to 9) [a]	90 (82 to 97) [a]	60 (35 to 85) [a]	24 (18 to 30) [a]	0.5 (0.2 to 1.4) [a]	1.1 (1.0 to 1.2) [a]	Very low
For detecting bacterial meningitis								
1 (Ghotbi, 2009)	254	8 (0 to 24) [a]	100 (100 to 100) [a]	100 (100 to 100) [a]	96 (93 to 98) [a]	NC	0.9 (0.8 to 1.1) [a]	Low
For detecting pneumonia, urinary tract infection, or bacteraemia								
1 (Craig, 2010)	12807	1 (0 to 1) [a]	100 (100 to 100) [a]	19 (7 to 31) [a]	93 (92 to 93) [a]	3.0 (1.4 to 6.5) [a]	1.0 (1.0 to 1.0) [a]	Low

NC Not calculable

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.28 GRADE profile for evaluation of neck stiffness

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Nuchal rigidity								
For detecting bacterial meningitis								
1 (Ghotbi, 2009)	254	8 (0 to 24) [a]	100 (100 to 100) [a]	100 (100 to 100) [a]	96 (93 to 98) [a]	NC	0.9 (0.8 to 1.1) [a]	Low
1 (Offringa, 1992)	92	48 (27 to 68) [a]	100 (100 to 100) [a]	100 (100 to 100) [a]	85 (77 to 93) [a]	NC	0.5 (0.4 to 0.8) [a]	Very low
For detecting meningococcal disease								
1 (Nielsen, 2001)	208	41 (26 to 56) [a]	97 (94 to 100) [a]	76 (58 to 94) [a]	88 (83 to 92) [a]	13.9 (5.4 to 35.6) [a]	0.6 (0.5 to 0.8) [a]	Very low

NC Not calculable

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.29 GRADE profile for evaluation of focal seizures

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Focal seizures								
For detecting bacterial meningitis								
1 (Akpede, 1992)	522	41 (20 to 61) [a]	92 (90 to 94) [a]	18 (8 to 29) [a]	97 (96 to 99) [a]	5.1 (2.9 to 9.2) [a]	0.6 (0.5 to 0.9) [a]	Very low
1 (Joffe, 1983)	241	38 (12 to 65) [b]	91 (87 to 95) [b]	20 (4 to 34) [b]	96 (94 to 99) [b]	4.2 (1.9 to 9.3) [a]	0.7 (0.4 to 1.0) [a]	Very low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study.

[b] Confidence intervals were calculated by the NCC-WCH based on results reported in the study.

Table 5.30 GRADE profile for evaluation of non-blanching rash

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Rash								
For detecting serious bacterial infection								
1 (Nijman, 2011)	1255	3 (0 to 6) [a]	97 (96 to 98) [a]	12 (1 to 23) [a]	90 (88 to 91) [a]	1.1 (0.4 to 3.1) [a]	1.0 (1.0 to 1.0) [a]	Very low
For detecting pneumonia, UTI or bacteraemia								
1 (Craig, 2010)	12,807	12 (10 to 14) [a]	82 (81 to 83) [a]	5 (4 to 6) [a]	92 (92 to 93) [a]	0.7 (0.6 to 0.8) [a]	1.1 (1.1 to 1.1) [a]	Low
Purpura								
For detecting bacteraemia								
1 Mandl (1997)	411	83 (40 to 99)	97 (95 to 98)	31 (5 to 57)	99 (99 to 100)	28.1 (14.5 to 54.5) [a]	0.2 (0.0 to 1.0) [a]	Very low
1 (Nademi , 2001)	141	29 (15 to 43)	98 (95 to 100)	86 (67 to 100)	77 (69 to 84)	8.9 (2.6 to 30.4) [a]	0.8 (0.6 to 0.9) [a]	Very low
1 (Baker, 1989)	190	40 (15 to 65) [a]	89 (80 to 98) [a]	55 (25 to 84) [a]	82 (71 to 92) [a]	3.6 (1.3 to 10.1) [a]	0.7 (0.4 to 1.0) [a]	Very low
1 (Offringa , 1992)	401	13 (0 to 27) [a]	100 (100 to 100) [a]	100 (100 to 100) [a]	78 (69 to 86) [a]	NC	0.9 (0.7 to 1.0) [a]	Very low
1 Mandl (1997)	411	83 (54 to 100)[a]	97 (95 to 99) [a]	31 (9 to 54) [a]	100 (99 to 100) [a]	28.5 (14.4 to 56.4)[a]	0.2 (0.0 to 1.0) [a]	Very low
1 (Nielsen, 2001)	208	74 (61 to 88) [a]	49 (42 to 57) [a]	25 (17 to 33) [a]	89 (83 to 96) [a]	1.5 (1.2 to 1.9) [a]	0.5 (0.3 to 0.9) [a]	Very low
1 (Nielsen, 2001)	208	95 (88 to 100) [a]	78 (72 to 84) [a]	50 (39 to 61) [a]	99 (96 to 100) [a]	4.3 (3.2 to 5.8) [a]	0.1 (0.0 to 0.3) [a]	Very low
1 (Nielsen, 2001)	208	74 (61 to 88) [a]	92 (88 to 96) [a]	67 (53 to 81) [a]	94 (90 to 98) [a]	9.0 (5.3 to 15.3) [a]	0.3 (0.2 to 0.5) [a]	Very low

NC Not calculable

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.31 GRADE profile for evaluation of diarrhoea

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Diarrhoea								
For detecting serious bacterial infection								
1 (Craig, 2010)	15781	21 (19 to 24) [a]	74 (73 to 75) [a]	6 (5 to 7) [a]	92 (92 to 93) [a]	0.8 (0.7 to 0.9) [a]	1.1 (1.0 to 1.1) [a]	Low
1 (Berger, 1996)	138	55 (38 to 72) [a]	20 (12 to 28) [a]	18 (10 to 25) [a]	58 (42 to 74) [a]	0.7 (0.5 to 0.9) [a]	2.3 (1.3 to 3.9) [a]	Low
For detecting urinary tract infection								
1 (Morris, 2007)	98	NC[b]	NC[b]	NC[b]	NC[b]	NC[b]	NC[b]	Low
For detecting bacterial illness								
1 (Trautner, 2006)	103	NC[c]	NC[c]	NC[c]	NC[c]	NC[c]	NC[c]	Very low
Diarrhoea and vomiting								
For detecting serious bacterial infection								
1 (Nijman, 2011)	1255	6 (2 to 10) [a]	91 (89 to 92) [a]	7 (2 to 12) [a]	89 (87 to 91) [a]	0.6 (0.3 to 1.3) [a]	1.0 (1.0 to 1.1) [a]	Very low
Mild gastrointestinal symptoms								
For detecting serious bacterial infection								
1 (Shin, 2009)	221	15 (4 to 25) [a]	89 (84 to 94) [a]	24 (7 to 41) [a]	81 (76 to 87) [a]	1.3 (0.6 to 3.1) [a]	1.0 (0.8 to 1.1) [a]	Low

CI confidence interval, NC non-calculable, OR odds ratio

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study.

[b] Text in the paper stated that diarrhoea is not predictive of urinary tract infection

[c] The paper reported: OR 3.93 (95% CI 1.27 to 12.19)

Table 5.32 GRADE profile for evaluation of vomiting

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Vomiting								
For detecting serious bacterial infection								
1 (Bleeker, 2007)	381	49 (40 to 59) [a]	69 (64 to 75) [a]	36 (28 to 44) [a]	80 (75 to 85) [a]	1.6 (1.2 to 2.1) [a]	0.7 (0.6 to 0.9) [a]	Low
1 (Bleeker, 2001)	231	37 (30 to 44) [a]	43 (30 to 56) [a]	66 (57 to 75) [a]	19 (12 to 25) [a]	0.7 (0.5 to 0.9) [a]	1.5 (1.1 to 2.0) [a]	Very low
For detecting serious disease								
1 (Nademi , 2001)	141	59 (43 to 73)	60 (50 to 69)	38 (25 to 49)	78 (68 to 87)	1.5 (1.0 to 2.1) [a]	0.7 (0.5 to 1.0) [a]	Very low
For detecting bacterial illness								
1 (Trautne r, 2006)	103	NR [b]	NR [b]	NR [b]	NR [b]	NR [b]	NR [b]	Very low
For detecting urinary tract infection								
1 (Morris, 2007)	98	NR [c]	NR [c]	NR [c]	NR [c]	NR [c]	NR [c]	Low
1 (Rabasa Al, 2009)	145	60 (39 to 81) [a]	60 (51 to 69) [a]	19 (10 to 29) [a]	90 (84 to 97) [a]	1.5 (1.0 to 2.3) [a]	0.7 (0.4 to 1.2) [a]	Low
For detecting bacterial meningitis								
1 (Ghotbi, 2009)	254	67 (40 to 93) [a]	100 (100 to 100) [a]	100 (100 to 100) [a]	98 (97 to 100) [a]	NC	0.3 (0.1 to 0.7) [a]	Very low
1 (Offringa , 1992)	92	48 (27 to 68) [a]	81 (72 to 90) [a]	46 (26 to 66) [a]	82 (73 to 91) [a]	2.5 (1.3 to 4.9) [a]	0.6 (0.4 to 1.0) [a]	Very low
For detecting meningococcal disease								
1 (Nielsen, 2001)	208	44 (28 to 59) [a]	60 (52 to 67) [a]	20 (12 to 29) [a]	82 (75 to 89) [a]	1.1 (0.7 to 1.6) [a]	0.9 (0.7 to 1.3) [a]	Very low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Increased vomiting								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	15 (9 to 20) [a]	82 (80 to 84) [a]	8 (5 to 11) [a]	90 (88 to 92) [a]	0.8 (0.6 to 1.2) [a]	1.0 (1.0 to 1.1) [a]	Low
Diarrhoea and vomiting								
For detecting serious bacterial infection								
1 (Nijman, 2011)	1255	6 (2 to 10) [a]	91 (89 to 92) [a]	7 (2 to 12) [a]	89 (87 to 91) [a]	0.6 (0.3 to 1.3) [a]	1.0 (1.0 to 1.1) [a]	Very low
Mild gastrointestinal symptoms								
For detecting serious bacterial infection								
1 (Shin, 2009)	221	15 (4 to 25) [a]	89 (84 to 94) [a]	24 (7 to 41) [a]	81 (76 to 87) [a]	1.3 (0.6 to 3.1) [a]	1.0 (0.8 to 1.1) [a]	Low

Ci confidence interval, NR not reported, odds ratio

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

[b] The paper reported: OR 0.76 (95% CI 0.26 to 2.18)

[c] The text in the paper stated that vomiting is not predictive of urinary tract infection

Table 5.33 GRADE profile for evaluation of abdominal pain

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Abdominal pain								
For detecting serious illness								
1 (Nijman, 2011)	1255	5 (1 to 8) [a]	97 (95 to 98) [a]	13 (3 to 23) [a]	90 (88 to 91) [a]	1.3 (0.6 to 3.1) [a]	1.0 (1.0 to 1.0) [a]	Very low
For detecting urinary tract infection								
1 (Morris, 2007)	98	NR [b]	NR [b]	NR [b]	NR [b]	NR [b]	NR [b]	Low

NR Not reported

[a] Calculated by the NCC-WCH based on data reported in the study

[b] The text in the paper stated that abdominal pain is not predictive or urinary tract infection

Table 5.34 GRADE profile for evaluation of crying on micturition/dysuria

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Crying on micturition/dysuria								
For detecting urinary tract infection								
1 (Rabasa Al, 2009)	145	10 (0 to 23) [a]	86 (79 to 92) [a]	10 (0 to 23) [a]	86 (79 to 92) [a]	0.7 (0.2 to 2.8) [a]	1.1 (0.9 to 1.2) [a]	Low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.35 GRADE profile for evaluation of headache

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Headache								
For detecting bacterial meningitis								
1 (Ghotbi, 2009)	254	17 (0 to 38) [a]	100 (99 to 100) [a]	67 (13 to 100) [a]	96 (94 to 98) [a]	40.3 (3.9 to 414.3) [a]	0.8 (0.6 to 1.1) [a]	Very low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.36 GRADE profile for evaluation of conjunctivitis

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Conjunctivitis								
For detecting urinary tract infection								
1 (Newman, 2002)	1666	1 (1 to 2) [a]	99 (99 to 100) [a]	7 (6 to 21) [a]	90 (89 to 92) [a]	0.7 (0.1 to 5.5) [a]	1.0 (1.0 to 1.0) [a]	Low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.37 GRADE profile for evaluation of poor peripheral circulation

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Poor peripheral circulation								
For detecting serious bacterial infection								
1 (Bleeker, 2001)	231	11 (6 to 16) [a]	78 (69 to 88) [a]	59 (42 to 76) [a]	23 (17 to 28) [a]	0.5 (0.3 to 0.9) [a]	1.1 (1.0 to 1.3) [a]	Very low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 3.38 GRADE profile for evaluation of bulging abdomen

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Bulging abdomen								
For detecting serious bacterial infection								
1 (Bleeker, 2001)	231	6 (2 to 9) [a]	88 (80 to 96) [a]	59 (35 to 82) [a]	24 (18 to 30) [a]	0.5 (0.2 to 1.2) [a]	1.1 (1.0 to 1.2) [a]	Very low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.39 GRADE profile for evaluation of paresis or paralysis

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Paresis or paralysis								
For detecting bacterial meningitis								
1 (Offringa, 1992)	92	30 (12 to 49) [a]	91 (85 to 98) [a]	54 (27 to 81) [a]	80 (71 to 89) [a]	3.5 (1.3 to 9.4) [a]	0.8 (0.6 to 1.0) [a]	Very low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.40 GRADE profile for evaluation of abnormal neurological findings

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Abnormal neurological findings								
For detecting bacterial meningitis								
1 (Joffe, 1983)	241	92 (78 to 100) [a]	84 (79 to 89) [a]	25 (13 to 37) [a]	99 (98 to 100) [a]	5.8 (4.2 to 8.2) [a]	0.1 (0.0 to 0.6) [a]	Very low
1 (Offringa, 1992)	92	64 (44 to 84) [a]	91 (88 to 94) [a]	35 (20 to 50) [a]	97 (95 to 99) [a]	7.0 (4.3 to 11.4) [a]	0.4 (0.2 to 0.7) [a]	Very low
Neurological deficit								
For detecting bacterial meningitis								
1 (Batra, 2011)	199	80 (45 to 100) [a]	99 (98 to 100) [a]	80 (45 to 100) [a]	99 (98 to 100) [a]	155.2 (20.9 to 1150.8) [a]	0.2 (0.0 to 1.2) [a]	Very low

[a] Calculated by the NCC-WCH based on data reported in the study

Table 5.41 GRADE profile for evaluation of impression of tone

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Impression of tone								
For detecting bacteraemia								
1 (Crain, 1982)	175	NR [a]	NR [a]	NR [a]	NR [a]	NR [a]	NR [a]	Low

NR Not reported

[a] Text in the paper stated that impression of tone is not significantly associated with bacteraemia

Table 5.42 GRADE profile for evaluation of tenderness on examination

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Tenderness on examination								
For detecting urinary tract infection								
1 (Shaw, 1998)	2411	5 (0 to 10) [a]	99 (98 to 99) [a]	13 (1 to 26) [a]	97 (96 to 98) [a]	4.5 (1.6 to 12.5) [a]	1.0 (0.9 to 1.0) [a]	Low

[a] Calculated by the NCC-WCH based on data reported in the study

Table 5.43 GRADE profile for evaluation of urinary symptoms

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Urinary symptoms								
For detecting serious bacterial infection								
1 (Craig, 2010)	15781	5 (4 to 6) [a]	98 (98 to 98) [a]	17 (13 to 21) [a]	93 (93 to 93) [a]	2.7 (2.0 to 3.6) [a]	1.0 (1.0 to 1.0) [a]	Low
1 (Nijman, 2011)	1255	8 (4 to 13) [a]	99 (98 to 99) [a]	41 (22 to 59) [a]	90 (89 to 92) [a]	5.9 (2.8 to 12.4) [a]	0.9 (0.9 to 1.0) [a]	

[a] Calculated by the NCC-WCH based on data reported in the study

Table 5.44 GRADE profile for evaluation of abnormal ear, nose and throat signs

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Abnormal ear, nose and throat signs								
For detecting serious bacterial infection								
1 (Craig, 2010)	15781	42 (39 to 45) [a]	45 (44 to 46) [a]	6 (5 to 6) [a]	91 (90 to 92) [a]	0.8 (0.7 to 0.8) [a]	1.3 (1.2 to 1.4) [a]	Low
Ear problems								
Serious bacterial infection								
1 (Nijman, 2011)	1255	4 (1 to 7) [a]	99 (98 to 99) [a]	17 (3 to 31) [a]	94 (93 to 95) [a]	3.2 (1.2 to 8.3) [a]	1.0 (0.9 to 1.0) [a]	Very low

[a] Calculated by the NCC-WCH based on data reported in the study

Table 5.45 GRADE profile for evaluation of rigor and/or chills

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Rigors								
For detecting confirmed or presumed bacterial illness								
1 (Tal, 1997)	434	28 (23 to 34) [a]	83 (78 to 89) [a]	67 (78 to 95) [a]	49 (44 to 55) [a]	1.7 (1.2 to 2.5) [a]	0.9 (0.8 to 1.0) [a]	Very low

[a] Calculated by a member of the technical team at the NCC-WCH based on results reported in the study

Table 5.46 GRADE profile for evaluation of Yale Observation Scale

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Score of 3 or 4								
For detecting serious illness								
1 (McCarthy, 1981)	312	67 (45 to 88) [a]	79 (74 to 84) [a]	19 (9 to 29) [a]	97 (95 to 99) [a]	3.2 (2.1 to 4.8) [a]	0.4 (0.2 to 0.8) [a]	Very low
1 (McCarthy, 1981)	312	56 (33 to 79) [a]	89 (85 to 93) [a]	27 (13 to 41) [a]	96 (94 to 99) [a]	5.0 (2.9 to 8.7) [a]	0.5 (0.3 to 0.8) [a]	Very low
1 (McCarthy, 1981)	312	72 (52 to 93) [a]	79 (74 to 84) [a]	20 (10 to 30) [a]	97 (95 to 100) [a]	3.5 (2.4 to 5.0) [a]	0.4 (0.2 to 0.7) [a]	Very low
Score of 4 or 5								
For detecting bacteraemia								
1 (Haddon, 1999)	534	6 (0 to 16) [a]	95 (92 to 97) [a]	5 (0 to 15) [a]	95 (93 to 97) [a]	1.0 (0.1 to 7.4) [a]	1.0 (0.9 to 1.1) [a]	Low
Score of 5, 6, or 7								
For detecting bacterial illness or pneumonia								
1 (McCarthy, 1980)	219	60 (35 to 85) [a]	76 (70 to 82) [a]	16 (6 to 25) [a]	96 (93 to 99) [a]	2.5 (1.5 to 4.0) [a]	0.5 (0.3 to 1.0) [a]	Very low
1 (McCarthy, 1980)	219	27 (4 to 49) [a]	94 (91 to 97) [a]	25 (4 to 46) [a]	95 (91 to 98) [a]	4.5 (1.7 to 12.4) [a]	0.8 (0.6 to 1.1) [a]	Very low
Score > 6								
For detecting bacteraemia								
1 (Teach, 1995)	6680	29 (22 to 35) [b]	83 (82 to 83) [b]	5 (3 to 6) [b]	97 (97 to 98) [b]	1.6 (1.3 to 2.1) [a]	0.9 (0.8 to 0.9) [a]	Very low
Score > 8								
For detecting bacteraemia								
1 (Teach, 1995)	6680	17 (11 to 22)[b]	92 (91. to 93) [b]	6 (4 to 8) [b]	97 (97 to 98) [b]	2.0 (1.5 to 2.8) [a]	0.9 (0.9 to 1.0) [a]	Very low
1 (Bang, 2009)	219	97 (79 to 99)	66 (55 to 72)	52 (43 to 62)	98 (93 to 100)	2.8 (2.2 to 3.5)	0.1 (0.0 to 0.2)	Moderate

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Score > 9								
For detecting serious bacterial infection								
1 (Thayyil, 2005)	72	13 (0 to 35) [a]	33 (21 to 44) [a]	2 (0 to 7) [a]	75 (59 to 91) [a]	0.2 (0.0 to 1.2) [a]	2.7 (1.7 to 4.1) [a]	Low
Score of > 10								
For detecting serious illness (including aseptic bacterial meningitis)								
1 (Baker, 1990)	126	46 (30 to 62) [b]	80 (71 to 88) [b]	49 (32 to 65) [b]	78 (70 to 87) [b]	2.3 (1.3 to 3.9) [a]	0.7 (0.5 to 0.9) [a]	Moderate
For detecting serious bacterial infection								
1 (Galetto-Lacour, 2003)	110	23 (5 to 54)	82 (67 to 92)	32 (12 to 51)	75 (66 to 84)	1.3 (0.6 to 2.9) [b]	0.9 (0.8 to 1.2) [b]	Very low
1 (Andreola, 2007)	408	38 (28 to 48) [b]	68 (63 to 73) [b]	26 (19 to 34) [b]	79 (74 to 83) [b]	1.2 (0.9 to 1.6) [b]	0.9 (0.8 to 1.1) [b]	Low
For detecting bacteraemia								
1 (Teach, 1995)	6680	5 (2 to 8) [b]	97 (96 to 97) [b]	5 (2 to 7) [b]	97 (97 to 98) [b]	1.6 (0.9 to 3.0) [a]	1.0 (0.9 to 1.0) [a]	Very low
1 (Bang, 2009)	219	88 (71 to 93)	84 (73 to 87)	68 (56 to 78)	95 (89 to 98)	5.4 (3.7 to 7.9)	0.1 (0.1 to 0.3)	Moderate
For detecting bacterial disease								
1 (Baker, 1990)	126	33 (7 to 60) [b]	73 (65 to 81) [b]	11 (1 to 22) [b]	91 (85 to 97) [b]	1.2 (0.5 to 2.9) [a]	0.9 (0.6 to 1.4) [a]	Moderate
For detecting urinary tract infection								
1 (Zorc, 1995)	1025	4 (0 to 9) [a]	93 (91 to 94) [a]	6 (0 to 11) [a]	91 (89 to 93) [a]	0.6 (0.2 to 1.6) [a]	1.0 (1.0 to 1.1) [a]	Moderate
Score of 10 to 16								
For detecting serious bacterial infection								
1 (Andreola, 2007)	408	43 (33 to 53) [a]	74 (69 to 79) [a]	33 (24 to 41) [a]	81 (77 to 86) [a]	1.6 (1.2 to 2.2) [a]	0.8 (0.6 to 0.9) [a]	Low
Score of 11 to 15								
For detecting serious illness								
1 (McCarthy, 1982)	312	31 (16 to 46) [a]	84 (79 to 89) [a]	26 (13 to 39) [a]	87 (82 to 2) [a]	1.9 (1.1 to 3.4) [a]	0.8 (0.7 to 1.0) [a]	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
1 (Teach, 1995)	6680	1 (0 to 2) [b]	99 (99 to 99) [b]	1 (0 to 4) [b]	97. (97 to 97) [b]	0.4 (0.1 to 3.2) [a]	1.0 (1.0 to 1.0) [a]	Very low
1 (Bang, 2009)	219	48 (27 to 56)	91 (67 to 90)	68 (52 to 82)	82 (75 to 87)	5.5 (3.0 to 9.8)	0.6 (0.4 to 0.7)	Moderate
Score of ≥ 16								
For detecting serious illness								
1 (McCarthy, 1982)	312	33 (18 to 49)[a]	99 (98 to 100) [a]	92 (78 to 100) [a]	89 (85 to 93) [a]	64.7 (8.7 to 482.0) [a]	0.7 (0.5 to 0.8) [a]	Low
For detecting serious bacterial infection								
1 (Andreola, 2007)	408	9 (3 to 14) [a]	98 (96 to 99) [a]	53 (28 to 79) [a]	78 (74 to 82) [a]	3.8 (1.4 to 10.3) [a]	0.9 (0.9 to 1.0) [a]	Low

NA Not applicable

[a] Calculated by the NCC-WCH from results reported in the study

[b] Confidence intervals calculated by the NCC-WCH from data reported in the study

Table 5.47 GRADE profile for comparison of Yale Observation Scores

Number of studies	Duration of fever		Effect	Quality
	With serious bacterial illness/infection (SBI) (Mean, SD)	Without SBI (Mean, SD)	*P* value	
Yale Observation Score				
For detecting serious bacterial infection				
1 (Hsiao, 2006)	9.4 (SD 4.6)	8.1 (SD 3.6)	P < 0.05	High
For detecting bacteraemia				
1 (Haddon, 1999)	7.0 (SD 1.5)	7.4 (SD 1.9)	*P* = 0.45	High

P probability, SBI serious bacterial illness/infection, SD standard deviation

Evidence statements

The following definitions have been used when summarising the likelihood ratios:

- Convincing: positive likelihood ratio (LR+) 10 or higher, negative likelihood ratio (LR-) 0.1 or lower
- Strong: LR+ 5 or higher (but less than 10), LR- 0.2 or lower (but higher than 0.1)
- Not strong: LR+ 4.9 or lower, LR- higher than 0.2

The following definitions have been used when summarising the levels of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV):

- High: 90% and above
- Moderate: 75% to 89%
- Low: 74% or below

The symptoms and signs were grouped by the categories used in the 2007 traffic light table, namely 'colour', 'activity', 'hydration', 'respiratory' and 'other'.

Colour

Pallor reported by parent/carer or pale/mottled/ashen/blue (in 2007 traffic light table)

Colour was reported in two studies where the definitions used included cyanotic, pale or mottled colour. The sensitivity was low for detecting serious bacterial infection or urinary tract infection. The specificity ranged from high to low. The positive predictive values were low and the negative predictive values ranged from high to low. The likelihood ratios were not strong.

Activity

Not responding normally to social cues or no response to social cues (in 2007 traffic light table)

Social cues were reported in one study. Decreased social interaction had low sensitivity, specificity and positive predictive value for detecting urinary tract infection. The negative predictive value was high. The likelihood ratios were not strong.

Appears ill to a healthcare professional (in 2007 traffic light table) or parent carer

'Appears ill to a healthcare professional' was reported in 15 studies; definitions of this included 'appears unwell', 'poor appearance', 'mildly unwell', 'not well-appearing', 'moderately ill', 'moderately unwell', 'moderately ill appearance', 'ill appearance', 'ill general appearance', 'very unwell', 'very ill appearance', 'severely ill', 'toxicity' and 'suspicious physical findings'. Some of these studies included included parent/carer reports of 'appears ill', but did not present this data separately.

Sensitivity, specificity, positive predictive values and negative predictive values ranged from high to low for detecting urinary tract infection, pneumonia, bacteraemia, serious bacterial infection, serious illness, occult infections, invasive bacterial disease, serious invasive bacteraemia or bacterial meningitis. Positive likelihood ratios ranged from not strong to convincing. Convincing positive likelihood ratios were reported for using 'very unwell' to detect urinary tract infection, pneumonia or bacteraemia, and for using 'toxicity' to detect bacterial meningitis. The negative likelihood ratios were not strong.

Wakes only with prolonged stimulation or does not wake, or if roused, does not stay awake (in 2007 traffic light table)

'Wakes only with prolonged simulation' was reported in five studies, in which the definitions used included drowsy being reported in case notes or on examination, increased sleepiness, drowsiness, drowsiness at home and postictal drowsiness. Sensitivity was low for detecting serious illness, urinary tract infection or bacterial meningitis. Specificity and positive predictive values ranged from high to low. The negative predictive values ranged from high to moderate. The positive likelihood ratio ranged from not strong to strong. The negative likelihood ratios were not strong.

Decreased activity (in 2007 traffic light table)

Decreased activity was reported in six studies, including 'looking around the room', 'moving arms and legs spontaneously', 'reaching for objects' and lethargy. Sensitivity was low for detecting urinary tract infection, bacteraemia, serious bacterial infection or bacterial meningitis. Specificity, positive predictive values and negative predictive values ranged from high to low. Positive likelihood ratios ranged from not strong to strong. The negative likelihood ratios were not strong.

No smile (in 2007 traffic light table)

No evidence was reported for no smile.

Weak, high-pitched or continuous cry (in 2007 traffic light table)

An abnormal cry was reported in three studies. One study stated that crying was not significantly associated with bacteraemia, whilst another reported a significant odds ratio for bacteraemia in children with an abnormal cry compared to those without an abnormal cry. One study reported low sensitivity, specificity and positive predictive value for detecting serious bacterial infection. The negative predictive value was high. The likelihood ratios were not strong.

Irritability (identified in 2013 review)

Irritability was reported in six studies. In two of these studies irritability was not found to be significantly associated with bacteraemia or urinary tract infection. The other four studies reported diagnostic accuracy data, with two of these finding a high negative predictive value for detecting bacteraemia or bacterial meningitis. However, in the other two studies the negative predictive value was low to moderate for detecting serious bacterial infection and viral or non-specific meningitis. The sensitivity, specificity and positive predictive values were not high in any of the studies, and the likelihood ratios were not strong.

Decreased consciousness and/or coma (identified in 2013 review)

Decreased consciousness and/or coma were reported in four studies, including unrousable coma. Decreased consciousness or coma has a low sensitivity for detecting serious bacterial infection or bacterial meningitis; however, the specificity was high. The positive and negative predictive values ranged from high to low, and the likelihood ratios were not strong.

Restlessness (identified in 2013 review)

Restlessness was reported in one study. The sensitivity, specificity, positive predictive value and negative predictive value were not high and the likelihood ratios were not strong for detecting serious illness.

Respiratory

Nasal flaring (in 2007 traffic light table)

No evidence was reported on the use of nasal flaring for detecting serious illness.

Grunting (in 2007 traffic light table)

No evidence was reported on the use of grunting for detecting serious illness.

Tachypnoea (in 2007 traffic light table)

Tachypnoea was reported in two studies, including ‘elevated respiratory rate’. The sensitivity was low and specificity was moderate for detecting pneumonia or bacteraemia. The positive predictive values were low and the negative predictive values were high. The positive and negative likelihood ratios were not strong.

Oxygen saturation (in 2007 traffic light table)

No evidence was reported on the use of oxygen saturation for detecting serious illness.

Moderate or severe chest indrawing (in 2007 traffic light table)

No evidence was reported on the use of chest indrawing for detecting serious illness.

Crackles (in 2007 traffic light table)

The presence of crackles was reported in two studies, including ‘abnormal chest sounds’ and crepitation. The sensitivity was low for detecting pneumonia, urinary tract infection, bacteraemia or serious bacterial illness. The specificity ranged from high to moderate. The positive predictive values were low, and the negative predictive values ranged from high to low. The positive and negative likelihood ratios were not strong.

Respiratory symptoms (identified in 2013 review)

Respiratory symptoms were reported in four studies, with definitions including respiratory distress, breathing difficulty, shortness of breath and breathing difficulty or chest wall recession. The sensitivity was not high or not reported for detecting urinary tract infection, pneumonia, bacteraemia, serious illness or serious bacterial infection. The specificity ranged from high to low. The positive predictive value was not high, and the negative predictive value ranged from high to moderate. The likelihood ratios were either not strong or not reported.

Nasal symptoms (identified in 2013 review)

Nasal symptoms were reported in three studies, with definitions including purulent nasal discharge, upper respiratory tract infection or runny nose, and symptoms of mild upper respiratory tract infection. The sensitivity was not high for detecting serious bacterial illness or urinary tract infection. The specificity ranged from high to low. The positive predictive value was not high and the negative predictive value ranged from high to low. The likelihood ratios were not strong.

Wheeze (identified in 2013 review)

Wheeze was reported in one study, including audible wheeze and stridor. The sensitivity was low for detecting pneumonia, urinary tract infection or bacteraemia. The specificity was high. The positive predictive values were low and the negative predictive values were high. The likelihood ratios were not strong.

Chest findings/abnormal chest sounds (identified in 2013 review)

Chest findings/abnormal chest sounds were reported in two studies. The sensitivity was low for detecting pneumonia, urinary tract infection or bacteraemia. The specificity was high. The positive predictive values were low and the negative predictive values were high. The likelihood ratios were not strong.

Cough (identified in 2013 review)

Cough was reported in three studies. The sensitivity was low for detecting pneumonia, urinary tract infection, bacteraemia or meningococcal disease. The specificity ranged from high to low. The positive predictive values were low and the negative predictive values ranged from high to moderate. The likelihood ratios were not strong.

Hydration

Dry mucous membranes (in 2007 traffic light table)

No evidence was reported on the use of dry mucous membranes for detecting serious illness.

Reduced skin turgor (in 2007 traffic light table)

No evidence was reported on the use of reduced skin turgor for detecting serious illness.

Poor feeding (in 2007 traffic light table)

Poor feeding was reported in four studies; definitions used included poor intake and decreased feeding. The sensitivity, specificity and positive predictive values ranged from moderate to low for detecting serious bacterial infection, serious illness or urinary tract infection. The negative predictive values ranged from high to low. The positive and negative likelihood ratios were not strong.

Capillary refill time of 3 seconds or more (in 2007 traffic light table)

Capillary refill time was reported in one study, using a time of 2 to 3 seconds or more than 3 seconds. The sensitivity was low for detecting pneumonia, urinary tract infection or bacteraemia. The specificity was high. The positive predictive value was low. The negative predictive value was high. The positive likelihood ratio ranged from not strong to strong. The negative likelihood ratios were not strong.

Reduced urine output (in 2007 traffic light table)

Reduced urine output was reported by two studies, including poor micturition. The sensitivity was low and the specificity was moderate for detecting urinary tract infection or serious bacterial infection. The positive predictive value ranged from moderate to low and the negative predictive value ranged from high to low. The likelihood ratios were not strong.

Other

Fever for 5 days or more (in 2007 traffic light table)

As shown in Table 5.22, duration of fever was reported in 17 studies, at the following time points: 12 hours, 24 hours, 48 hours, 2 days and 72 hours for detecting serious bacterial infection, bacteraemia, meningococcal disease and urinary tract infection. All of the time points resulted in a low sensitivity. The specificities ranged from high to low; however, the expected correlation between increasing specificities and increasing fever duration was not found. Positive predictive values were mainly low. Negative predictive values ranged from high to low, although again this was not in the expected pattern. The positive and negative likelihood ratios were not strong for any cutoffs. There was no

significant difference in the odds of serious bacterial infection when comparing children who had had fever for longer than 48 hours with those who had had fever for less than 24 hours; however, a fever duration of 72 hours or longer was significantly associated with serious illness.

As shown in Table 5.23, there were mixed results when comparing the duration of fever in children with and without serious illness. Some studies reported that children with serious illness had had fever for significantly longer than those without, whilst other studies reported that there was no significant difference in the duration of fever.

Temperature of 38°C or higher at age 0–3 months, temperature of 39°C or higher at age 3–6 months (in 2007 traffic light table)

Thirty-six studies reported on the height of fever in children aged less than 5 years. As shown in Table 5.25, various cut-offs were reported, including 37.4°C or higher, 37.5°C or higher, higher than 37.5°C, 38°C or higher, 38.4°C or higher, 38.5°C or higher, higher than 38.5°C, 39°C or higher, higher than 39°C, 39.1°C or higher, higher than 39.3°C, 39.4°C or higher, 39.5°C or higher, higher than 39.5°C, 40°C or higher, higher than 40°C, 40.1°C or higher, and 41.1°C or higher. These were used to try to detect urinary tract infection, serious bacterial infection, malaria or meningitis, serious illness, severe illness requiring hospitalisation, bacteraemia, bacterial meningitis, bacterial infection, pneumonia, sepsis and serious disease. Sensitivity and specificity ranged from high to low but were not correlated with temperature. All of the positive predictive values were low. The negative predictive values ranged from high to low, although also not in the expected pattern. Positive likelihood ratios were strong for 40°C or higher and for higher than 40°C, but were not strong for any other cutoffs. Negative likelihood ratios were not strong for any cutoffs. As shown in Table 5.26, when comparing the mean or median height of fever in those with and without serious illness, there were mixed results as to whether the difference was significant or not.

Six of the 36 studies reported on the height of fever exclusively in children aged less than 3 months, including 38°C or higher, higher than 39°C, 39.5°C or higher, 40°C or higher, and higher than 40°C for detecting serious bacterial infection, urinary tract infection, occult bacteraemia, bacteraemia, meningitis, bacterial meningitis, salmonella enteritis, sepsis and serious bacterial illness. Sensitivity was high for 38°C or higher and low for all other cutoffs. Specificity was low for 38°C or higher, moderate for higher than 39°C, moderate to high for 39.5°C or higher, and high for 40°C or higher and higher than 40°C. Positive predictive values were low for all cutoffs, and negative predictive values were high for all cutoffs. Positive likelihood ratios were not strong for 38°C or higher, higher than 39°C and 39.5°C or higher. They were strong for 40°C or higher and higher than 40°C. Negative likelihood ratios were not reported for 38°C or higher and were not strong for the other cutoffs.

No studies reported on the height of fever solely in children aged from 3 to 6 months.

Swelling of a limb or joint (in 2007 traffic light table)

No evidence was reported on the use of swelling of a limb or joint to detect serious illness.

Non-weight bearing limb/not using an extremity (in 2007 traffic light table)

No evidence was reported on the use of non-weight bearing limb or not using an extremity to detect serious illness.

Non-blanching rash (in 2007 traffic light table)

Non-blanching rash was reported in seven studies, including ‘rash’, purpura, petechiae, purpura with petechiae, more than 20 haemorrhages, haemorrhages greater than 1 mm in diameter, and haemorrhages greater than 2 mm in diameter. The sensitivity, specificity and positive predictive values ranged from high to low for detecting pneumonia, urinary tract infection, bacteraemia, serious disease, serious bacterial infection, invasive disease, bacterial meningitis or meningococcal disease. The negative predictive values ranged from high to moderate. The positive and negative likelihood ratios ranged from not strong to convincing.

Bulging fontanelle (in 2007 traffic light table)

Bulging fontanelle was reported in three studies. The sensitivity was low for detecting serious bacterial illness, bacterial meningitis, pneumonia, urinary tract infection or bacteraemia. The specificity was high. The positive and negative predictive values ranged from high to low. The likelihood ratios were not strong.

Neck stiffness (in 2007 traffic light table)

Nuchal rigidity was reported in three studies. The sensitivity was low for detecting bacterial meningitis or meningococcal disease. The specificity was high. The positive and negatives predictive values ranged from high to moderate. The positive likelihood ratio was either not calculable or convincing. The negative likelihood ratios were not strong.

Status epilepticus (in 2007 traffic light table)

No evidence was reported on the use of status epilepticus to detect serious illness.

Focal neurological signs (in 2007 traffic light table)

No evidence was reported on the use of focal neurological signs to detect serious illness.

Focal seizures (in 2007 traffic light table)

Focal seizures were reported in two studies. The sensitivity was low for detecting bacterial meningitis. The specificity was high. The positive predictive values were low and the negative predictive values were high. The positive likelihood ratios ranged from not strong to strong and the negative likelihood ratios were not strong.

A new lump larger than 2 cm (in 2007 traffic light table)

No evidence was reported on the use of a new lump larger than 2 cm to detect serious illness.

Bile-stained vomiting (in 2007 traffic light table)

No evidence was reported on the use of bile-stained vomiting to detect serious illness.

Diarrhoea (identified in 2013 review)

Diarrhoea was reported in five studies including diarrhoea alone, diarrhoea with vomiting, and ‘mild gastrointestinal symptoms’. The sensitivity, specificity and positive predictive value for using diarrhoea to detect serious bacterial infection was low. The negative predictive value ranged from high to low. The likelihood ratios were not strong. The specificity for using diarrhoea and vomiting or mild gastrointestinal symptoms to detect serious bacterial infection was moderate to high, although the sensitivity, positive predictive value and negative predictive value were not high and the likelihood ratios were not strong. One study reported that the odds of having a bacterial illness are 3.9 times greater in those with diarrhoea compared to those without diarrhoea. One study reported that diarrhoea was not predictive of urinary tract infection.

Vomiting (identified in 2013 review)

Vomiting was reported in 11 studies, including increased vomiting, vomiting reported with diarrhoea, and ‘mild gastrointestinal symptoms’. The sensitivity for detecting serious bacterial infection, serious disease, bacterial illness, urinary tract infection, bacterial meningitis or meningococcal disease was low. The specificity, positive predictive value and negative predictive value ranged from high to low. The likelihood ratios were not strong. One study reported that the odds of having a bacterial illness were not significantly different in those with vomiting compared with those without. One of the studies reported that vomiting was not predictive of urinary tract infection.

Abdominal pain (identified in 2013 review)

Abdominal pain was reported in two studies. One study reported a low sensitivity and positive predictive value and high specificity and negative predictive value for detecting serious illness. The likelihood ratios were not strong. The other study reported no significant association between abdominal pain and urinary tract infection.

Crying on micturition/dysuria (identified in 2013 review)

Crying on micturition/dysuria was reported in one study. The sensitivity, specificity, positive predictive value and negative predictive value were not high, and the likelihood ratios were not strong, for detecting urinary tract infection.

Headache (identified in 2013 review)

Headache was reported in one study. The sensitivity for detecting bacterial meningitis was low. The specificity was high. The positive predictive value was low and the negative predictive value was high. The positive likelihood ratio was convincing, although the negative likelihood ratio was not strong.

Conjunctivitis (identified in 2013 review)

Conjunctivitis was reported in one study. The sensitivity for detecting urinary tract infection was low. The specificity was high. The positive predictive value was low and the negative predictive value was high. The likelihood ratios were not strong.

Poor peripheral circulation (identified in 2013 review)

Poor peripheral circulation was reported in one study. The sensitivity, specificity, positive predictive value and negative predictive value were not high, and the likelihood ratios were not strong, for detecting serious bacterial infection.

Bulging abdomen (identified in 2013 review)

Bulging abdomen was reported in one study. The sensitivity, specificity, positive predictive value and negative predictive value were not high, and the likelihood ratios were not strong, for detecting serious bacterial infection.

Paresis or paralysis (identified in 2013 review)

Paresis or paralysis was reported in one study. The sensitivity for detecting bacterial meningitis was low. The specificity was high, the positive and negative predictive values were not high, and the likelihood ratios were not strong.

Abnormal neurological findings (identified in 2013 review)

Abnormal neurological findings were reported in three studies, including neurological deficit. The sensitivity for detecting bacterial meningitis ranged from high to low. The specificity ranged from high to moderate. The positive predictive values were not high and the negative predictive values were high. The positive likelihood ratio ranged from convincing to strong, and the negative likelihood ratio from convincing to not strong.

Impression of tone (identified in 2013 review)

Impression of tone was reported in one study. The study reported no significant association between the symptom/sign and bacteraemia.

Tenderness on examination (identified in 2013 review)

Tenderness on examination was reported in one study. The sensitivity for detecting urinary tract infection was low. The specificity was high. The positive predictive value was low and the negative predictive value was high. The likelihood ratios were not strong.

Urinary symptoms (identified in 2013 review)

Urinary symptoms were reported in two studies. The sensitivity for detecting serious bacterial infection was low. The specificity was high. The positive predictive value was low and the negative predictive value was high. The likelihood ratios were not strong.

Abnormal ear, nose and throat signs (identified in 2013 review)

Abnormal ear, nose and throat signs were reported in two studies, including ‘ear problems’ reported in one study. The sensitivity and positive predictive values for detecting serious bacterial infection were not high. The specificity was high in one study and low in the other study. The negative predictive value was high in both studies. The likelihood ratios were not strong in either study.

Rigor and/or chills (identified in 2013 review)

Rigor and/or chills were reported in one study. The sensitivity, specificity, positive predictive value and negative predictive value for detecting bacterial illness were not high. The likelihood ratios were not strong.

Cold hands and feet (identified in 2013 review)

No studies were found that looked specifically at cold hands and feet for detecting serious illness in febrile children.

Yale Observation Scale

The Yale Observation Scale was reported in 12 studies. The sensitivity, specificity and positive predictive value ranged from high to low for detecting serious illness, bacteraemia, pneumonia, serious bacterial infection, bacterial disease and/or urinary tract infection, but were not correlated with

the YOS score. The negative predictive value ranged from moderate to high and was also not correlated with the YOS score. The positive and negative likelihood ratios ranged from not strong to convincing.

Health economic evidence statements

No health economic studies were identified and no health economic evaluation was undertaken for this question.

Evidence to recommendations

Relative value placed on the outcomes considered

The overarching aim of the guideline is to provide a framework for healthcare professionals to enable the early and accurate detection of serious illness in children with fever. This allows suitable treatment to begin, thereby reducing subsequent potential mortality and morbidity.

The GDG considered the likelihood ratios, sensitivity, specificity and predictive values of each symptom or sign when discussing the evidence. However, particular emphasis was given to likelihood ratios, with a positive likelihood ratio of 5 or higher being used as a good indicator that a symptom or sign should be presented in the red column of the traffic light table. In addition, the expert opinion and experience of the GDG members also informed the final decision about whether to include, remove or move a symptom or sign in the traffic light table.

Consideration of clinical benefits and harms

The traffic light table was created in order to encourage healthcare professionals to consider signs or symptoms in their totality and not in isolation. Therefore, the evidence for any individual symptom or sign had to be balanced by its contribution to the overall clinical picture and practical clinical application. Furthermore, the GDG highlighted that studies assessing the use of combinations of signs and symptoms show they have better predictive values than symptoms in isolation (for example Van Den Bruel et al, 2007 and Thompson et al., 2012). This concept was incorporated into the recommendation of 'none of the amber or red symptoms or signs' in the green column, and the 'appears ill to a healthcare professional' in the red column, without the need to specify the absence of particular symptoms or signs.

For each symptom and sign presented below, the GDG has stated:

- why the symptom or sign was included in the 2007 traffic light table (if applicable)
- the GDG's interpretation of the diagnostic outcome measures presented in the evidence statements for the symptom or sign
- the GDG's expert opinion on the inclusion of the symptom or sign in the traffic light table, and
- whether the symptom or sign was included in the 2013 update of the traffic light table.

Colour

Pallor reported by parent/carer or pale/mottled/ashen/blue (included in 2007 traffic light table)

'Colour' had been included in the 2007 traffic light as part of the YOS.

Low quality evidence from two studies was identified in the 2013 review. The reported evidence showed that children with cyanotic, pale or flushed/mottled skin were not more likely to have a serious illness than children with normal colour skin (not a strong positive likelihood ratio). Children with a serious illness did not usually have cyanotic, pale or flushed/mottled skin (low sensitivity). However, the evidence for children without serious illness was mixed, with one study showing they did not usually have cyanotic, pale or flushed/mottled skin (high specificity) and one study showing that they usually did have cyanotic, pale or flushed/mottled skin (low specificity). One of the studies used colour to detect urinary tract infection and the GDG members were not convinced of the relevance of colour to this diagnosis.

Given the quality of the evidence, the GDG members were of the clinical opinion that children with pale/mottled/ashen/blue skin were not being incorrectly categorised as in the 'red' category. The GDG decided that there was no reason to change or remove this sign from the traffic light table.

Activity

Not responding normally to social cues or no response to social cues (included in 2007 traffic light table)

'Activity' was included in the 2007 traffic light as part of the YOS.

The 2013 review found evidence to support assessing activity level at presentation. The reported evidence showed that children with decreased social interaction were not more likely to have a urinary tract infection than children with normal social interaction (not a strong positive likelihood ratio). Children without a urinary tract infection often showed decreased social interaction (low specificity). Children with a urinary tract infection did not usually show decreased social interaction (low sensitivity). The evidence was of low quality.

The GDG acknowledged that it would be helpful to define 'social cues' for parents, caregivers or less experienced healthcare professionals. The glossary of the guideline has been updated to outline that this can include the parents' perception of a baby behaving differently, response to their name, smiling and/or giggling.

The GDG chose to keep decreased activity in the 'amber' column as the evidence did not support movement into the 'red' category based on definitions used in the study. If the decreased activity is severe, healthcare professionals may use their clinical judgement of 'appears ill to a healthcare professional' to manage the child appropriately. Therefore, no change was made to the traffic light table.

Appears ill to a healthcare professional (included in 2007 traffic light table) and parents/carers

'Appears ill to a healthcare professional' was included in the 2007 traffic light table as part of the YOS.

The 2013 review supported the results of the 2007 review. The results of the studies were mixed, with some studies showing that children who appeared unwell were not more likely to have a serious illness than those who appeared well (not a strong positive likelihood ratio), and other studies showing that children who appeared unwell were more likely to have a serious illness (convincing positive likelihood ratio).

The GDG members acknowledged that being 'very ill' was more predictive than 'appears ill'. However, they were aware that it is hard to distinguish between the two terms. The majority of studies reporting this sign did not define 'appears ill', and those that did used a combination of symptoms and signs that are presented elsewhere in the traffic light table. As there was no separate data available on parent/carer reports of 'appears ill', the GDG decided that that no recommendation could be made specifically on parent/carer reports of 'appears ill'. However, the GDG highlighted that parent/carer reports of fever and other specific symptoms were covered by other recommendations in the guideline.

Based on their expert opinion, the GDG members noted that 'appears ill to a healthcare professional' can be subjective and difficult to define. Therefore, the GDG concluded that for this sign, the definition of 'healthcare professional' should be restricted to those who are trained in assessing children, for example GPs, specialist nurses and paediatricians. The GDG members concluded, based on their clinical experience, that there was not a strain on resources for children who are inappropriately referred because of this sign.

Given the mixed quality of the evidence, the GDG did not change 'appears ill to a healthcare professional' in the traffic light table.

Wakes only with prolonged stimulation or does not wake, or if roused, does not stay awake (included in 2007 traffic light table)

'Wakes only with prolonged stimulation or does not wake, or if roused, does not stay awake' was included in the 2007 traffic light table as part of the YOS.

The evidence was mixed, with some studies showing that children who were difficult to rouse were more likely to have a serious illness than those who were not difficult to rouse (convincing and strong positive likelihood ratio), and some studies showing that children who were difficult to rouse were not more likely to have a serious illness than those who were not difficult to rouse (not a strong positive likelihood ratio). The evidence was of low and very low quality, and most of the studies focused on detecting bacterial meningitis rather than serious illness in general.

The GDG did not believe the evidence was strong enough to move or remove this from the traffic light table, and therefore no changes were made to the traffic light table for this sign.

Decreased activity (included in 2007 traffic light table)

'Decreased activity' was included in the 2007 traffic light table as part of the YOS.

Some studies showed that children with decreased activity were more likely to have a serious illness than children with normal levels of activity (strong positive likelihood ratio); however, other studies showed that children with decreased activity were not more likely to have a serious illness than children with normal levels of activity (not strong positive likelihood ratio). The evidence was mainly of low to very low quality.

The 2007 recommendation referred to decreased activity by parental report, but the 2013 review shows that decreased activity at presentation to a healthcare professional was also a useful symptom or sign of serious illness. The GDG acknowledged that decreased activity was difficult to define, and that it was difficult to distinguish between 'moderate' and 'severe' impairment, as reported in one of the included studies.

Given the varied definitions and quality of the evidence the GDG decided to keep decreased activity in the 'amber' column, and so no changes were made to the traffic light table.

No smile (included in 2007 traffic light table)

The GDG stated this was included in the 2007 traffic light table as part of the YOS.

No new evidence was identified in the 2013 review. Therefore, the GDG agreed that this sign would not be changed or removed.

Weak, high-pitched or continuous cry (included in 2007 traffic light table)

This feature was included in the 2007 traffic light table as part of the YOS.

The evidence from the 2013 review was low in quality, and only one study reported diagnostic data or data that allowed diagnostic data to be calculated. The study showed that children with an abnormal cry were not more likely to have serious illness than children without an abnormal cry (not a strong positive likelihood ratio).

The GDG therefore stated that the 2013 data was not strong enough to change or remove 'weak, high-pitched or continuous cry' from the traffic light table.

Irritability (identified in 2013 review)

The evidence showed that children who were irritable were not more likely to have a serious illness than children who were not irritable (not a strong positive likelihood ratio).

The sign 'content/smiles' is already included in the 'green' column of the traffic light table. The GDG believed that this is in line with the evidence that shows children without irritability usually do not have a serious illness. The GDG believed there was a general consensus in clinical practice that irritability can be defined as when an infant or child is uncomfortable when picked up or moved; however, none of the studies adequately defined irritability.

As 'content/smiles' is already included in the 'green' column of the table, the GDG did not add irritability to the traffic light table.

Decreased consciousness and/or coma (identified in 2013 review)

The evidence implied that children with decreased consciousness were not more likely to have a serious illness than children with a normal level of consciousness (not a strong positive likelihood ratio). In addition, the evidence showed that children without a serious bacterial infection usually did

not have decreased consciousness (high specificity). However, children with a serious bacterial infection did not usually present with decreased consciousness (low sensitivity).

The reviewed evidence was based on a population outside the intended guideline population; that is, children older than 5 years or those with febrile convulsions. Furthermore, the GDG believed that this sign was already included in the traffic light table as 'does not wake, or if roused, does not stay awake'.

Based on the quality of the available evidence and its discussion, the GDG decided that no changes relating to decreased consciousness and/or coma were needed to the traffic light table.

Restlessness (identified in 2013 review)

The evidence regarding restlessness was reported in one study. Children who were restless were not more likely to have a serious illness than children who were not restless (not a strong positive likelihood ratio). Children with a serious illness were often restless (moderate sensitivity); however, children without a serious illness were also often restless (low specificity).

Based on the limited evidence, the GDG did not believe restlessness was a useful symptom to detect serious illness. Therefore, restlessness was not added to the traffic light table.

Respiratory

The majority of respiratory symptoms were originally included in the traffic light table as indicators of pneumonia.

Nasal flaring and grunting (included in 2007 traffic light table)

No new evidence was found for nasal flaring or grunting in the 2013 review.

The GDG emphasised that clinical judgment should be used to distinguish between nasal flaring (amber symptom/sign) and grunting (red symptom/sign).

Based on the available evidence and its discussion, the GDG decided that no changes relating to nasal flaring and grunting were needed to the traffic light table.

Tachypnoea (included in 2007 traffic light table)

Abnormal respiratory rate was included in the 2007 traffic light table as a non-specific marker of serious illness, a specific feature of pneumonia and required for the assessment of dehydration. A statement about measuring respiratory rate was combined with the statement about the physiological parameters which should be documented as part of the assessment.

The 2013 review of the evidence showed that children who had tachypnoea were not more likely to have a serious illness than children who did not have tachypnoea (not a strong positive likelihood ratio). In addition, the evidence showed that children without a serious illness often did not have tachypnoea (moderate specificity). However, the evidence showed that children with a serious illness also did not usually have tachypnoea (low sensitivity). The available evidence was of low quality.

The cut-offs proposed by Fleming et al. (2011) and Nijman et al. (2012) were reviewed, but there was no significantly clear evidence on specific rates to alter the categories.

The GDG members concluded from their experience that respiratory rate is an important physiological parameter which needs to be assessed by healthcare professionals.

Given the low quality of the evidence, the GDG did not believe the evidence was strong enough to change or remove an existing recommendation. Therefore, no changes relating to tachypnoea were made to the traffic light table.

Oxygen saturation (included in 2007 traffic light table)

Oxygen saturation was included in the original traffic light table as a specific sign of pneumonia.

The current review did not find any evidence regarding oxygen saturation for detecting serious illness.

However, the GDG members were aware that the measurement of oxygen saturation is becoming more common amongst GPs and non-paediatric accident and emergency departments. Using their expert opinions, the GDG members believed that oxygen saturation should be retained in the traffic light table.

Based on the available evidence and its discussion, the GDG decided that no changes relating to oxygen saturation were needed to the traffic light table.

Moderate or severe chest indrawing (included in 2007 traffic light table)

Chest indrawing was included in the original traffic light table as a specific sign of pneumonia.

The current review did not find any further evidence regarding chest indrawing for detecting serious illness.

The GDG decided it should be retained in the traffic light table. A definition of chest indrawing is provided in the glossary.

Crackles (included in 2007 traffic light table)

The evidence relating to crackles in the 2013 review was of low and very low quality. The evidence showed that children with crackles were not more likely to have a serious illness than children who did not have crackles (not a strong positive likelihood ratio). In addition, the evidence suggested children without a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, usually did not have crackles (high specificity). However, children with a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, also did not usually have crackles (low sensitivity). The evidence was of low to very low quality.

Given the quality of the evidence, the GDG did not believe the evidence was strong enough to change or remove an existing recommendation. Therefore, no changes relating to crackles were made to the traffic light table.

Respiratory symptoms (identified in 2013 review)

The 2013 review highlighted that the evidence supports existing symptoms and signs in the original traffic light table.

The GDG believed that the new evidence was not defined well enough to add anything further to the assessment of respiratory symptoms.

Therefore, no changes relating to respiratory symptoms were made to the traffic light table.

Nasal symptoms (identified in 2013 review)

The evidence shows that serious illness is not ruled out by a lack of nasal symptoms (low sensitivity).

The GDG members were aware from their clinical experience that less serious complaints, such as upper respiratory tract infections, are often used to rule out the presence of a serious illness. However, the GDG stated that nasal symptoms were too common to be of practical use.

The GDG, therefore, did not add nasal symptoms to the ‘green’ column of the traffic light table.

Wheeze (or stridor) (identified in 2013 review)

The evidence shows that children who had wheeze were not more likely to have a serious illness than children who did not have wheeze (not a strong positive likelihood ratio). In addition, the evidence showed that children without a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, usually did not have wheeze (high specificity). However, children with a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, also did not usually have wheeze (low sensitivity).

The GDG agreed that wheeze was too common a symptom to be moved into the ‘amber’ or ‘red’ columns. Therefore, no changes relating to wheeze were made to the traffic light table.

Chest findings/abnormal chest sounds (identified in 2013 review)

The evidence showed that children who had abnormal chest sounds were not more likely to have a serious illness than children who did not have abnormal chest sounds (not a strong positive likelihood ratio). In addition, the evidence showed that children without pneumonia, a urinary tract infection or bacteraemia usually did not have abnormal chest sounds (high specificity). However, it also showed that children with pneumonia, a urinary tract infection or bacteraemia also did not usually have abnormal chest sounds (low sensitivity).

The GDG highlighted that ‘crackles’ was already included in the traffic light table, which was a better defined sign than ‘chest findings’ or ‘abnormal chest sounds’. One of the studies included in the

review was for detecting urinary tract infection, and the GDG was unsure how relevant chest findings or abnormal chest sounds would be to this diagnosis.

Given the quality of the evidence and the fact that an item already covering this feature was already included in the traffic light table, the GDG decided not make any changes relating to chest findgins/abnormal chest sounds to the traffic light table.

Cough (identified in 2013 review)

The available evidence showed that children who had a cough were not more likely to have a serious illness than children who did not have a cough (not a strong positive likelihood ratio).There was some evidence that children without a urinary tract infection usually did not have a cough (high specificity), but other evidence showed that children without a urinary tract infection, pneumonia, bacteraemia or meningococcal disease often had a cough (low specificity). In addition, children with a urinary tract infection, pneumonia, bacteraemia or meningococcal disease did not usually have a cough (low sensitivity).

The evidence suggests cough was not a useful predictor of serious illness, although the GDG highlighted that two of the studies were on detecting urinary tract infection and it was not clear how relevant cough was to this diagnosis. There was not enough evidence for the GDG to determine that cough was a useful symptom or sign in the detection of serious illness. Furthermore, the GDG stated 'cough' was too common to be of practical use.

Based on the available evidence and the results of its discussion, the GDG decided not make any changes relating to cough to the traffic light table.

Circulation and hydration

In the 2007 guideline the GDG recognised that dehydration was a marker of serious illness but there was a lack of evidence to determine the difference between mild, moderate and severe dehydration. The most specific symptoms and signs of dehydration have been highlighted for healthcare professionals to assess in order to ensure a low false positive rate and are included in the guideline Diarrhoea and vomiting in children under 5 (NICE, 2009). As evidence was found relating to the use of heart rate in the diagnosis of serious illness, the 'hydration' category was changed to 'circulation and hydration' for greater clarity.

Dry mucous membranes and reduced skin turgor (included in 2007 traffic light table)

The GDG acknowledged that dry mucous membranes and reduced skin turgor were included in the 2007 traffic light table based on a study that reviewed signs and symptoms of dehydration, rather than a study of serious illness associated with fever. However, the GDG members stated that, in their experience, dehydration was a marker for serious illness and therefore should be included in the traffic light table.

No new evidence was found for dry mucous membranes and/or reduced skin turgor in the 2013 review.

The GDG acknowledged that the recommendations regarding signs of dehydration in the 2007 Fever guideline were intended for use primarily in children who had been sent home after seeing a healthcare professional. Since the publication of the 2007 Feverish Illness in Children guideline, a clinical guideline on diarrhoea and vomiting has been published (Diarrhoea and vomiting in children under 5, NICE 2009). The Diarrhoea and vomiting guideline concluded that looking at physical signs of dehydration was an inaccurate way of determining whether a child was moderately or severely ill, as it is difficult to distinguish between different severities of dehydration. However, the two guidelines consider different populations, and if a child exhibits diarrhoea and/or vomiting they are treated in accordance with that guideline rather than the Fever guideline. The GDG also emphasised that the purpose of the traffic light table is to raise awareness rather than to make clear definitive diagnosis.

In the absence of evidence to challenge the 2007 recommendation, the GDG did not change it.

Poor feeding (included in 2007 traffic light table)

The 2013 review did not find clear evidence relating poor feeding to an increased risk of serious illness. Children who showed poor feeding were not more likely to have a serious illness than children who showed normal feeding (not a strong positive likelihood ratio). The evidence was of low to very low quality.

However, the GDG members stated that, in their clinical experience, poor feeding was a key reason that parents or caregivers bring their child to a healthcare professional. In recognition that poor feeding was a worrying feature, but not an immediate alarm feature, its position was in the amber column in the 2007 traffic light table. The GDG acknowledged that it was hard to define poor feeding. Depending on the age of the child, it can be difficult to assess how much the child is feeding, for example if the child is being breastfed. Furthermore, the GDG also acknowledged that the Nademi et al. (2001) study includes children up to age 16 years, who have more control over their own feeding habits, and therefore the data may not be applicable to the population covered by this guideline who are under aged 5 years. In addition, the Newman et al. (2002) study investigates urinary tract infection, which is not relevant to this sign.

The GDG's decision was that the new data was not strong enough to support changing the 2007 recommendation, and so no changes were made to it.

Capillary refill time of 3 seconds or more (included in 2007 traffic light table)

In the 2007 guideline the GDG noted that capillary refill time is quick to carry out and exhibits moderate reproducibility. A statement about measuring capillary refill time was combined with the statement about the physiological parameters which should be documented as part of the assessment (see the end of Respiratory rate section). The GDG considered that a capillary refill time of 3 seconds or more was an 'amber' sign (see the recommendations at the end of Respiratory rate section).

For the 2013 review the evidence showed that children with a capillary refill time of more than 3 seconds were more likely to have a serious illness than children with a capillary refill time of 3 seconds or less (strong positive likelihood ratio). In addition, evidence showed that children without a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, usually did not have an increased capillary refill time (high specificity). However, children with a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, did not usually have an increased capillary refill time either (low sensitivity). The evidence was of low to very low quality.

The GDG acknowledged that in the cut-offs reported in the Craig et al. (2010) study the capillary refill time is measured in whole seconds, and so greater than 3 seconds would be 4 seconds or more. However, the other cut-off reported in the study is 2 to 3 seconds. It was not clear whether the data for children with a capillary refill time of 3 seconds were included in the results.

In the 2013 review the GDG acknowledged that there is a difference in central and peripheral capillary refill time. The GDG was aware that peripheral capillary refill time can be affected without indicating a serious illness, and that taking peripheral measurements can be inaccurate and lead to false positives. The GDG emphasised that it is not a sign that should be used in isolation. For further details, please refer to the guideline Bacterial meningitis and meningococcal septicaemia (NICE, 2010).

The GDG stated that the data identified in the 2013 review was of limited quality and not strong enough to change the 2007 recommendations. Therefore, no changes relating to capillary refill time of 3 seconds or more were made to the 2007 recommendations.

Reduced urine output (included in 2007 traffic light table)

The evidence in the 2013 review showed that children with a reduced urine output were not more likely to have a serious illness than children with a normal urine output (not a strong positive likelihood ratio). In addition, the evidence showed children without a serious bacterial infection or a urinary tract infection often did not have reduced urine output (moderate specificity). However, children with a serious bacterial infection or a urinary tract infection also did not usually have reduced urine output (low sensitivity). The evidence was of low to very low quality.

The GDG members stated that in their experience reduced urine output is commonly reported by parents and caregivers as a marker of dehydration and its position in the amber column reflected its relevance.

Based on the quality of the evidence and its discussion, the GDG decided to keep the existing recommendation.

Other

Fever for 5 days or more (included in 2007 traffic light table)

This sign was included in the 2007 guideline as it was indicative of Kawasaki disease; however, the new review found only two studies that reported on Kawasaki disease, neither of which reported on duration of fever. There was evidence that those with a serious bacterial illness had had fever for longer than children without serious illness (significant *P* values), and children who had had fever for three days or more were significantly more likely to have a urinary tract infection than those who had not (significant relative risk). No evidence was reported that examined fever duration of longer than 5 days.

Based on their clinical experience, the GDG members argued that most non-serious illnesses will resolve themselves after 5 days, and therefore a fever of more than 5 days duration is a good indicator of serious illness. The GDG acknowledged that in the evidence there is a weak correlation between duration of fever and severity of illness. However, it believed this may be in part to relying on parental/caregiver recall of when the fever started. Also, the evidence was limited as many studies excluded children who had had fever for 5 days or longer and none of the studies used 5 days as a cut-off.

The GDG concluded that the evidence in the current review was not strong enough to change the 2007 recommendations and therefore no such changes were made.

Temperature of 38°C or more in children age under 3 months, temperature of 39°C or more in children age 3–6 months (included in 2007 traffic light table)

In the 2007 guideline the GDG concluded that healthcare professionals should be aware that there is an association between height of body temperature and risk of serious bacterial illness. However, this association was not sufficiently robust to recommend immediate action or referral based on body temperature alone. An exception was made for children aged less than 6 months with a body temperature of 39°C or higher because the evidence was strongest for this age group.

In the 2013 review, the GDG acknowledged the ambiguity of the age groups in the 2007 recommendation regarding height of fever, and altered the text of the recommendation to reflect the intended meaning of less than 3 months for one group, and age 3 to 6 months (inclusive) for the other group. No studies were identified for the 3 to 6 month age group specifically, although most studies included this age group in their sample. The studies often did not report how the temperature was measured, and the studies tended to look at one or two cut-offs rather than a range of temperatures, making it hard to compare data from different temperature cut-offs. Despite these limitations in the data, the GDG highlighted that there is a correlation between high temperature and serious bacterial infection in general, but that, on an individual basis, high temperature was not useful for detecting serious illness. The current review suggests that there is a plateau in positive predictive values, negative predictive values and likelihood ratios around 39°C and 40°C, suggesting that a temperature above this does not provide a better indication of serious illness. The GDG therefore decided to move the recommendation regarding height of fever in the 3 to 6 month age group from the red column to the amber column. The GDG acknowledged that any fever in a child under 3 months is a risk factor for serious illness in itself, and so the recommendation for this age group remained in the red column.

The GDG made it clear that use of height of fever alone should not be used to diagnosis a serious illness. In addition, the GDG noted that children aged less than 3 months with fever are generally at a higher risk of serious illness (see Section 8.2). The incidence of serious illness in this group, for instance, was over ten times higher than that in older children. The clinical studies that provide the evidence for this age group used a body temperature of 38°C or higher as the definition of fever.

The GDG was also aware that infants in England and Wales have their first immunisations at age 2 months and that most of these infants experience post-immunisation fever. There was a discussion about what impact a recommendation on height of fever in this age group would have on health services, with a potential for health services to be overwhelmed. However, it was highlighted that parents and carers were routinely advised to expect their child to have a fever within 48 hours of immunisation and that there was no evidence of an increase in consultations due to this.

The GDG therefore decided that children aged less than 3 months with a body temperature of 38°C or higher should be included in the recommendation about risk of serious illness.

Non-blanching rash, bulging fontanelle and neck stiffness (included in 2007 traffic light table)

In the 2007 traffic light table there were several symptoms and signs that were included because they are indicative of meningococcal septicaemia or bacterial meningitis, including non-blanching rash, bulging fontanelle and neck stiffness. The evidence was of low to very low quality.

The 2013 review reported that there was some evidence that children with a non-blanching rash were more likely to have a serious illness than children who did not have a non-blanching rash (convincing positive likelihood ratio); however, there was also evidence that children with a non-blanching rash were not more likely to have a serious illness than children who did not have a non-blanching rash (not a strong positive likelihood ratio). In addition, children without a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, usually did not have a non-blanching rash (high specificity). Children with a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, also did not usually have a non-blanching rash (low sensitivity).

The evidence for the 2013 review showed that children with a bulging fontanelle were not more likely to have a serious illness than children without a bulging fontanelle (not a strong positive likelihood ratio). Children without a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, usually did not have bulging fontanelle (high specificity). Children with a serious illness, such as pneumonia, a urinary tract infection or bacteraemia, also did not usually have a bulging fontanelle (low sensitivity).

The 2013 review reported that there was some evidence that children with neck stiffness were more likely to have meningococcal disease than children who did not have neck stiffness (convincing positive likelihood ratio). In addition, children without a serious illness, such as meningitis, pneumonia, a urinary tract infection or bacteraemia, usually did not have neck stiffness (high specificity). Children with a serious illness, such as meningitis, pneumonia, a urinary tract infection or bacteraemia, also did not usually have neck stiffness (low sensitivity).

The 2013 review also found that there was evidence that children with focal seizures were more likely to have bacterial meningitis than children who did not have focal seizures (strong positive likelihood ratio). In addition, children without a serious illness, such as meningitis, pneumonia, a urinary tract infection or bacteraemia, usually did not have focal seizures (high specificity). However, children with a serious illness, such as meningitis, pneumonia, a urinary tract infection or bacteraemia, also did not usually have focal seizures (low sensitivity).

Since the 2007 Fever guideline, a guideline on bacterial meningitis in children and young people has been published. The guideline Bacterial meningitis and meningococcal septicaemia (NICE, 2010) includes a comprehensive list of symptoms and signs of bacterial meningitis and meningococcal septicaemia. However, it is worth noting that the bacterial meningitis guideline is relevant when bacterial meningitis or meningococcal septicaemia is suspected, whereas the Fever guideline is relevant for children that do not have a known source of fever. The GDG stated that the most relevant symptoms and signs of bacterial meningitis and meningococcal septicaemia were included in the 2007 traffic light table, and the 2013 review found no strong evidence to move or remove these from the traffic light table. The GDG was aware that the symptoms of cold hands and feet and limb pain are included in the list of clinical features found in meningococcal disease and meningitis in the 2010 guideline.

Although it was of low quality, the available evidence supported the existing recommendation and matched the opinion of the GDG. Therefore, it was decided that the traffic light table did not need to be changed.

Status epilepticus (included in 2007 traffic light table)

No evidence was identified in the 2013 review for status epilepticus.

Based on their clinical experience, the GDG members stated that status epilepticus should remain in the ‘red’ column, as it is a serious condition and a child with status epilepticus needs urgent referral. Therefore, no changes were made to the recommendation on status epilepticus.

Focal neurological signs and focal seizures (included in 2007 traffic light table)

The GDG highlighted that focal neurological and focal seizures were included in the traffic light table as they may be indicative of *Herpes simplex encephalitis*.

There was no evidence identified in the 2013 review that reported on neurological signs or focal seizures for identifying serious illness.

Based on their clinical experience, the GDG members did not know of any clinical reason to move these signs from the 'red' column of the traffic light table and therefore no changes were made.

Swelling of a limb or joint, and non-weight bearing limb/not using an extremity (included in 2007 traffic light table)

The GDG highlighted that both swelling of a limb or joint and non-weight bearing limb/not using an extremity were included in the 2007 traffic light table as they are indicative of septic arthritis.

No evidence was identified in the 2013 review regarding swelling of a limb or joint and/or non-weight bearing limb for detecting serious illness. The GDG acknowledged that the consequences of missing the diagnosis of septic arthritis in a child are serious. However, it was also aware that this is not a common illness. The GDG also acknowledged that many children with swelling and/or non-weight bearing will recover from these symptoms in a few days, and so they do not require immediate referral.

Based on the available evidence and its discussion, the GDG decided that no changes were needed and these two symptoms should remained in the amber category of the traffic light table.

A new lump greater than 2 cm (included in 2007 traffic light table)

There was no evidence in the 2013 review to support including 'new lump greater than 2 cm' in the traffic light table. The study on which the 2007 recommendation was based was excluded as it included non-febrile surgical patients.

The GDG highlighted that 'new lump greater than 2 cm' was originally included in the traffic light table based on one study that was excluded from the update as the population included a high proportion of children without fever. A significant number of children in this study were diagnosed with hernias and other surgical conditions. Moreover, in a subset analysis of children with fever from this study, a new lump larger than 2 cm did not feature in a set of risk factors for serious illness. The GDG stated that a new lump larger than 2 cm most likely indicated a hernia or an abscess requiring surgical intervention, and was not associated with fever.

The GDG therefore decided to remove the existing recommendation, and so removed 'new lump greater than 2 cm' from the traffic light table.

Bile-stained vomiting (included in 2007 traffic light table)

There was no evidence in the 2013 review to support including 'bile-stained vomiting' in the traffic light table. The study on which the 2007 recommendation was based was excluded as it included non- febrile surgical patients.

The GDG was aware that bile-stained vomiting is more likely to indicate a surgical problem, rather than a serious bacterial illness. It was included in 2007 based on one study that was excluded from the updated review, as it included a high proportion of children without fever. A significant number of children in this study were diagnosed with hernias and other surgical conditions. Moreover, in a subset analysis of children with fever from this study, bile-stained vomiting did not feature in a set of risk factors for serious illness.

The GDG therefore decided to remove the existing recommendation, and hence removed 'bile- stained vomiting' from the traffic light table.

Diarrhoea (identified in 2013 review)

The evidence relating to diarrhoea was mixed, with some studies showing that children without a serious bacterial infection, a urinary tract infection or a bacterial illness usually did not have diarrhoea (high and moderate specificity) and some showing that children without serious bacterial infection often had diarrhoea (low specificity). However, children with a serious bacterial infection, a urinary tract infection or a bacterial illness did not usually have diarrhoea (low sensitivity). Children with diarrhoea were not more likely to have a serious illness than children without diarrhoea (not strong positive likelihood ratio).

The GDG stated that the evidence was not consistent enough to add diarrhoea to the traffic light table. The GDG highlighted that dehydration was already included in the traffic light table. The GDG

also highlighted that a child presenting with diarrhoea and/or vomiting should be managed as outlined in the guideline Diarrhoea and vomiting in children under 5 (NICE, 2009).

Based on the available evidence and its discussion, the GDG decided that no changes relating to diarrhoea were needed to the traffic light table.

Vomiting (identified in 2013 review)

The evidence showed children with vomiting were not more likely to have a serious illness than children without vomiting (not a strong positive likelihood ratio). In addition, some studies showed that children without a serious bacterial infection, a urinary tract infection or a bacterial illness usually did not have vomiting and some showed that the children without bacterial meningitis or urinary tract infection often had vomiting (moderate to high specificity). However, children with a serious bacterial infection, a urinary tract infection or a bacterial illness did not usually have vomiting (low sensitivity). The evidence was of low to very low quality.

The GDG stated that the evidence was not consistent enough to add vomiting to the traffic light table and highlighted that dehydration was already included in the traffic light table. The GDG also highlighted that a child presenting with diarrhoea and/or vomiting should be managed as outlined in the guideline Diarrhoea and vomiting in children under 5 (NICE, 2009).

Based on the available evidence and its discussion, the GDG decided that no changes relating to vomiting were needed to the traffic light table.

Abdominal pain (identified in 2013 review)

The evidence showed that children with abdominal pain were not more likely to have a serious illness than children without abdominal pain (not a strong positive likelihood ratio). In addition, the evidence showed that children without a serious illness usually did not have abdominal pain (high specificity). However, children with a serious illness also did not usually have abdominal pain (low sensitivity). The evidence was of low to very low quality.

The GDG stated that the evidence was of low and very low quality, and evidence on diagnostic accuracy was limited to that of one study. The other included study did not report diagnostic data or data that would allow diagnostic data to be calculated. It is worth noting that the temperature used as an inclusion criterion for this study was lower than other studies in the review.

The non-diagnostic accuracy evidence stated that abdominal pain is not predictive of urinary tract infection. Therefore, the GDG concluded that abdominal pain should not be added to the traffic light table.

Crying on micturition/dysuria (identified in 2013 review)

The evidence showed that children who cried on micturition were not more likely to have a urinary tract infection than children who did not cry on micturition (not a strong positive likelihood ratio). In addition, the evidence showed that children without a urinary tract infection often did not cry on micturition (moderate specificity); however, children with a urinary tract infection also did not usually cry on micturition (low sensitivity).

The GDG highlighted that the evidence was of low quality and limited to that of one study. Furthermore, the GDG stated that a child presenting with crying during micturition or dysuria would clearly be indicative of a urinary tract infection and should be managed as outlined in the guideline Urinary tract infection in children (NICE, 2007).

Based on the available evidence and its discussion, the GDG decided that no changes relating to crying on micturition/dysuria were needed to the traffic light table and this symptom was not added.

Headache (identified in 2013 review)

The evidence showed that children with a headache were more likely to have bacterial meningitis than children without a headache (convincing positive likelihood ratio). Evidence also showed that children without bacterial meningitis usually did not have a headache (high specificity) and that children with bacterial meningitis also did not usually have headache (low sensitivity).

The evidence for headache was of very low quality and limited to that of one study. The study included children from 6 months to 5 years, and it was not clear to the GDG how pre-verbal children

would communicate that they had a headache. The GDG concluded that the evidence was not strong enough to add headache to the traffic light table.

Based on the quality of the available evidence and its discussion, the GDG decided that no changes relating to headache were needed to the traffic light table.

Conjunctivitis (identified in 2013 review)

The evidence showed that children with conjunctivitis were not more likely to have a urinary tract infection than children without conjunctivitis (not a strong positive likelihood ratio). In addition, the evidence showed that children without a urinary tract infection usually did not have conjunctivitis (high specificity). However, children with a urinary tract infection also did not usually have conjunctivitis (low sensitivity).

The evidence for conjunctivitis was in relation to detecting urinary tract infection, and the GDG was not convinced of a clinical link between the two conditions. Therefore, the GDG did not add conjunctivitis to the traffic light table.

Poor peripheral circulation (identified in 2013 review)

The evidence showed that children with poor peripheral circulation were not more likely to have a serious illness than children with normal peripheral circulation (not a strong positive likelihood ratio). In addition, the evidence showed that children without a serious bacterial infection often had normal peripheral circulation (moderate specificity); however, children with a serious bacterial infection also usually had normal peripheral circulation (low sensitivity).

The GDG highlighted that capillary refill time, which acts as an indicator of poor peripheral circulation with a recognised definition, is already included in the traffic light table. Furthermore, the evidence was of very low quality and was limited to that of one study. In addition, poor peripheral circulation was not defined in the study, and the evidence shows that it was not a good detector of serious illness.

Based on the available evidence and its discussion, the GDG decided that no changes relating to poor peripheral circulation were needed to the traffic light table.

Bulging abdomen (identified in 2013 review)

The evidence showed that children with a bulging abdomen were not more likely to have a serious illness than children without a bulging abdomen (not a strong positive likelihood ratio). In addition, the evidence showed that children without a serious bacterial infection often did not have a bulging abdomen (moderate specificity); however, children with a serious bacterial infection also usually did not have a bulging abdomen (low sensitivity).

Evidence was of very low quality and was limited to that of one study. The GDG found that the evidence that bulging abdomen was a useful predictor of serious illness was not convincing. Therefore, no changes relating to bulging abdomen were made to the traffic light table.

Paresis or paralysis (identified in 2013 review)

The evidence showed that children with paresis or paralysis were not more likely to have bacterial meningitis than children without paresis or paralysis (not a strong positive likelihood ratio). In addition, the evidence showed that children without bacterial meningitis usually did not have paresis or paralysis (high specificity). However, children with bacterial meningitis also did not usually have paresis or paralysis (low sensitivity).

The evidence for paresis or paralysis for detecting serious illness was of very low quality and was limited to that of one study. The included children had all had a febrile convulsion prior to inclusion in the studies. The GDG stated that a child with paresis or paralysis is likely to be identified using the traffic light table under ‘appears ill to a healthcare professional’ and ‘focal neurological signs’. The evidence was not convincing to add paresis or paralysis as an additional symptom or sign.

The GDG decided that paresis or paralysis should not be added to the traffic light table.

Abnormal neurological findings (identified in 2013 review)

The GDG stated that ‘abnormal neurological findings’ is already covered in the traffic light table under ‘focal neurological signs’ and ‘appears ill to a healthcare professional’. The new evidence was not

strong enough to add abnormal neurological findings to the traffic light table as a separate symptom or sign. All of the included studies used abnormal neurological findings to detect bacterial meningitis, and a child presenting with bacterial meningitis should be managed as outlined in the guideline Bacterial meningitis and meningococcal septicaemia (NICE, 2010).Therefore, no changes relating to abnormal neurological findings were made to the traffic light table.

Impression of tone (identified in 2013 review)

The evidence for impression of tone was limited to one study, which did not report diagnostic accuracy data or data that would allow diagnostic accuracy data to be calculated. The evidence stated that tone was not significantly associated with bacteraemia.

Therefore, the GDD decided that impression of tone should not be added to the traffic light table.

Tenderness on examination (identified in 2013 review)

The review results showed that children who showed tenderness on examination were not more likely to have a urinary tract infection than children who did not show tenderness on examination (not a strong positive likelihood ratio). In addition, the evidence showed that children without a urinary tract infection usually did not have tenderness on examination (high specificity). However, children with a urinary tract infection also did not usually have tenderness on examination (low sensitivity).

The GDG stated that tenderness on examination was not described in enough detail in the study to be used, although the GDG acknowledged that it was likely to refer to abdominal tenderness, as the study reports on urinary tract infection. In addition, the evidence was not strong enough for it to be added to the traffic light table.

Therefore, the GDG decided that tenderness on examination should not be added to the traffic light table.

Urinary symptoms (identified in 2013 review)

The evidence showed that children with urinary symptoms were not more likely to have a serious bacterial infection than children without urinary symptoms (not a strong positive likelihood ratio). In addition, the evidence showed that children without a serious bacterial infection usually did not have urinary symptoms (high specificity). However, children with a serious bacterial infection also did not usually have urinary symptoms (low sensitivity). There was some evidence that children with urinary symptoms were more likely to have a serious bacterial infection than children without urinary symptoms (strong positive likelihood ratio).

'Urinary symptoms' was not defined in the studies, although the GDG acknowledged that the term is likely to refer to symptoms and signs of urinary tract infection. This suggests a definite source cause of fever, and was not a helpful symptom or sign to add to the traffic light table. A child presenting with urinary symptoms should be managed as outlined in the guideline Urinary tract infection in children (NICE, 2007).

The GDG stated that two of the symptoms described in the 2007 guideline – offensive urine and haematuria – were rare, and if present would refer to a urinary condition. Therefore, these were removed from the recommendation. This did not result in any changes to the traffic light table.

Abnormal ear, nose and throat signs (identified in 2013 review)

The evidence was mixed for 'abnormal ear, nose and throat signs'. One study showed that children with abnormal ear, nose and throat signs were not more likely to have a serious illness than children with no signs (not a strong positive likelihood ratio). In addition, the evidence showed that children without a serious bacterial infection often had abnormal ear, nose and throat signs (low specificity), while another study showed that children without a serious bacterial infection usually did not have ear problems (high specificity). Both studies showed that children with a serious bacterial infection did not usually have abnormal ear, nose and throat signs or ear problems (low sensitivity).

The GDG highlighted that the evidence was of low and very low quality, and symptoms were too common to add 'abnormal ear, nose and throat signs' to the traffic light table.

The GDG therefore did not add 'abnormal ear, nose and throat signs' to the traffic light table.

Rigor and/or chills (identified in 2013 review)

The evidence suggested that children with rigors were not more likely to have a bacterial illness than children who did not have rigors (not strong positive likelihood ratio). The evidence showed that children without bacterial illness often did not have rigors (moderate specificity); however, children with a bacterial illness also usually did not have rigors (low sensitivity). The evidence was of very low quality.

The GDG highlighted that rigors are caused by a high body temperature, and are therefore associated with high temperatures in children. The GDG acknowledged that there was evidence of a link between higher temperatures in children and serious illness, and therefore rigors could be an indicator of serious illness. The GDG was aware that rigors are an uncommon symptom/sign in children under 5 years, but there was insufficient evidence that rigors alone signal the need for urgent attention.

The GDG stated that the quality of the evidence and positive likelihood ratio meant that rigors could not be added to the red column of the traffic light table. However, the GDG did feel it was an important feature and the decision was therefore made to add rigors to the amber category of the traffic light table.

Cold hands and feet (identified in 2013 review)

No evidence regarding cold hands and feet was reported in the 2013 review.

The GDG noted clinical overlap with poor peripheral circulation but that the NICE Bacterial meningitis and meningococcal septicaemia guideline (NICE, 2010) had identified cold hands and feet as a relevant sign when considering a diagnosis of meningitis.

The GDG was aware that the symptoms of cold hands and feet are included in the list of clinical features found in meningococcal disease and meningitis in the NICE 2010 meningitis guideline. However these symptoms were taken from uncontrolled studies and did not therefore fulfil the inclusion criteria of the updated Feverish illness guideline. Moreover, a study of these symptoms and signs in children with self-limiting viral illness found that cold hands and feet were reported in 20% to 24% of young children. The specificity of this symptom for detecting meningococcal disease would therefore be low. The GDG emphasised that in isolation, for undifferentiated children with fever, other features of the traffic light table were sufficient to identify high risk children and therefore did not add this symptom or sign to the traffic light table.

A child presenting with cold hands and feet should be diagnosed as outlined in the guideline Bacterial meningitis and meningococcal septicaemia (NICE, 2010).

Based on the available evidence and its discussion, the GDG decided that no changes relating to cold hands and feet were needed to the traffic light table.

Yale Observation Score

The evidence suggests that the Yale Observation Score was good at identifying children who do not have a serious illness. However, it was less good at identifying children who do have a serious illness. This was in line with the evidence found for the 2007 review that the YOS alone was not a good detector of serious illness. As highlighted in the 2007 review, the GDG acknowledged that the usefulness of the YOS was increased when it was used in combination with a history taken by a physician and examination.

Consideration of health benefits and resource uses

The GDG highlighted that the traffic light system would improve the initial management of examinations and reduce variation in practice. This would ensure that resources are focused on those who need further investigations and treatment, and not wasted on investigations or treatments that are not needed. It will also prevent unnecessary stress and anxiety for the child and their caregivers.

The GDG stated that the traffic light system was a quick and non-invasive method of identifying children with fever who may have a serious illness. Therefore, very little additional cost was associated with its use over and above a standard clinical examination, but its value was in the accuracy of the signs and symptoms that it contains.

Quality of evidence

The evidence ranged from high to very low in quality. There were a number of common issues which influenced the quality of the evidence, including lack of blinding of the clinicians and the use of different tests to confirm serious illness. However, the GDG highlighted that while much of the evidence was low quality, it was the best that is available on signs and symptoms.

The number of studies for most of the symptoms or signs was limited and not all of the reported evidence was directly relevant to the review question. This affected how applicable the data was to changing the traffic light table and meant that, for some symptoms and signs, the GDG did not have enough relevant data to make a decision on recommendations. In addition, the included studies varied in their approach, including which illnesses were being detected, the definition and measurement of symptoms and signs, the temperature cut-off for inclusion into the trial, the way in which inclusion temperature was measured (such as tympanic, rectal, axillary), the age of the included children, and the setting of the study (for example GP offices, hospital). These variations in the studies meant that data could not be pooled and made it difficult for the GDG to compare evidence from multiple studies for a symptom or sign. These variations also made it difficult for the GDG to compare the efficacy of different symptoms and signs with each other to inform decisions about whether a symptom or sign should be in the green, amber or red column of the traffic light table.

Some symptoms and signs were not well defined and the GDG did not believe it could add them to the traffic light table. In these cases, the GDG concluded that the details in the traffic light table provided a better definition of the symptoms or signs than the new evidence in the studies.

Some studies only included children who had experienced a febrile convulsion prior to presentation to a healthcare professional. These were included as there was a lack of data for the majority of symptoms and signs; however, the GDG emphasised that these children do not necessarily represent every child presenting to a healthcare professional with fever.

Due to these limitations with the studies, and without a sound clinical reason to alter the traffic light table, the majority of recommendations remained as they were in the 2007 guideline.

Other considerations

There were no other considerations specific to this section.

Equalities

The GDG acknowledged that special consideration needs to be made when assessing children with learning disabilities. Healthcare professionals should be aware that it may not be possible to apply all parts of the traffic light table to these children, and that care should be taken in interpreting the table when assessing these children.

The GDG also highlighted that care should be taken in interpreting the traffic light table when a complete history is not available, for example when a child presents without parents or caregivers. This may happen if the child is brought to a healthcare professional by a teacher or child minder, for example. It does not prevent the traffic light table from being used, but healthcare professionals should exercise caution in their approach.

The GDG stated that it can be difficult to assess pallor or a pale/mottled/ashen/blue appearance in children who have darker skin. Therefore, the GDG altered the wording of the existing recommendation to clarify that a pale/mottled/ashen/blue appearance can be identified on the lips or tongue of a child, as well as their skin. The wording of the green column heading and criteria was then edited to avoid repetition.

Similarly, capillary refill time may be a less useful test in children with darker skin tones. Peripheral measures may have to be used rather than central measures, for example in the beds of nails. Non-blanching rash may also be harder to detect, and clinicians should be aware of where a rash can be more easily identified, such as palms of hands, conjunctivae and soles of feet. For further details, please refer to the guideline Bacterial meningitis and meningococcal septicaemia (NICE, 2010).

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

.

Heart rate

Introduction

A specific review question was outlined for heart rate because no evidence was found for the 2007 guideline and it was known that new evidence had become available.

Heart rate is often assumed to be a useful marker of serious illness. For example, it is widely taught to use heart rate as a marker of circulatory insufficiency in shock.[110] However, heart rate is affected by a variety of factors (such as age, activity, anxiety, pain, body temperature) as well as the presence or absence of serious illness. A specific search was thus undertaken to look at heart rate in the context of serious illness.

Review question

The clinical question outlined in the scope was 'What is the predictive value of heart rate, including:

- how heart rate changes with temperature
- whether heart rate outside the normal range is a sign of serious illness.'

This translates into the following review question 'What is the predictive value of heart rate, including:

- how heart rate changes with temperature?
- whether heart rate outside the normal range detects serious illness?
- whether heart rate and temperature outside normal range detects serious illness?'

Description of excluded studies

Only one study was reviewed for the 2007 guideline and this was included in the updated review. No other studies were excluded.

Description of included studies

Six studies were identified for inclusion in this review (Brent et al., 2011; Davies et al., 2009; Hanna et al., 2004; Thompson et al., 2009; Thompson et al, 2008; Craig et al, 2010).

Three studies were included that evaluated how heart rate changes with temperature (Davies et al., 2009; Hanna et al., 2004; Thompson et al., 2008). The first study was a retrospective observational study (Davies et al., 2009) that included 21,033 children. The second was a prospective study (Hanna et al., 2004) that included 490 children who attended paediatric emergency departments, but who were not consequently admitted to hospital. The third study was a prospective cross-sectional study (Thompson et al., 2008) that included 1589 children who presented to primary care with a suspected acute infection.

Three studies were included that evaluated if heart rate alone could detect serious illness (Brent et al., 2010, Thompson et al., 2009; Craig, 2010). The Brent (2010) study included two datasets. The first was from a cross-sectional prospective study of 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection and the second was from a case–control study including 325 children with confirmed meningitis. The Thompson (2009) study examined 700 children attending a paediatric assessment unit for suspected infection. The Craig (2010) study examined 12,807 children presenting at a children's emergency department in a hospital in Australia. The study used an elevated heart rate to detect pneumonia, urinary tract infection or bacteraemia.

One study was included that examined heart rate in conjunction with temperature (Brent et al., 2010).

Evidence profile

The evidence is presented in both narrative and GRADE format.

How heart rate changes with temperature

Three studies are reviewed in this section (see Table 5.51 for the GRADE evidence profile).

The first study was a multi-centre, retrospective observational study (Davies et al., 2009) of 21,033 children which aimed to assess the effect of body temperature on heart rate in children attending a paediatric emergency department.

The authors of the paper analysed the data using a quantile regression and a statistical model to develop the following best fit equation:

$$\text{Expected parameter value (heart rate)} = (\text{Temperature } [^{o}C] \times a) + (\text{Age [months]} \times b) + (\text{Age}^2\ [\text{months}^2] \times c) + \text{constant}$$

In the equation, the temperature multiplier **a** has a mean increase of 10.52 beats per minute (bpm) through the centile, resulting in a heart rate increase of approximately 10 bpm with each 1°C increment in temperature. The results are shown in Table 5.48.

Table 5.48 Heart rate calculations for the 5th, 25th, 50th, 75th and 95th centiles

Percentile	a	b	c	constant
5th	9.468	-0.6543	0.001998	230.2
25th	10.99	-0.7040	0.002198	270.1
50th	11.44	-0.7393	0.002374	274.9
75th	11.35	-0.7615	0.002474	258.8
95th	9.397	-0.8494	0.002848	163.3

A number of limitations were identified including variation in how the measurements of pulse and temperature were taken, and the study including children older than 5 years.

The second study was a cross-sectional prospective study (Thompson et al., 2008) of 1589 children attending a paediatric emergency department that aimed to produce centile charts for heart rates in febrile children.

Centile charts of heart rate plotted against temperature in febrile children were produced. The incremental increases of heart rate for each increment of 1°C in temperature are shown in Table 5.49. Heart rate was negatively correlated with age (r = -0.62) and positively correlated with temperature (r = 0.49).

Table 5.49 The incremental increases of heart rate for each increment of 1 °C in temperature

Population	Mean increase in pulse rate per 1°C (1.8°F) Increase in temperature (95% CI)
Combined group of 1589 children	13.7
Age 3–12 months	12.1
Age 1–2 years	9.9
Age 2–5 years	14.1

CI confidence interval

This study showed that, in the study population, the heart rate increases by 9.9 to 14.1 bpm with each 1°C increment in temperature. The mean values of heart rate grouped by age at the 50th, 75th, 90th and 97th centiles are displayed in Table 5.49.

A number of limitations were identified, including: the children recruited were not a representative sample from primary care; and the study included children older than 5 years.

The third study was a prospective observational study (Hanna et al., 2004) which evaluated the effect on heart rate of fever in a cohort of 490 children attending a paediatric emergency department.

Centiles charts of pulse rate plotted against temperature in febrile children younger than 1 year were produced. The linear regression analysis of the relation between pulse rate and temperature is shown in Table 5.50.

Table 5.50 Linear regression analysis of the relation between pulse rate and temperature

Age (months)	Adjusted R^2	Mean increase in pulse rate (bpm) per 1°C increase in temperature (95% CI)
0–1	0.004	2.2 (-1.3 to 5.6)
2–3	0.16	10.0 (5.1–14.8)
4–5	0.25	10.6 (6.4–14.8)
6–7	0.22	9.2 (4.9–13.4)
8–9	0.10	6.8 (1.8–11.7)
10–11	0.38	10.9 (6.9–14.9)

bpm beats per minute, CI confidence interval

This study found that for every 1°C rise in body temperature, the resting heart rate rose by 9.6 bpm.

A number of limitations were identified: baseline figures were not controlled in analysis; there was limited reporting on exclusion criteria; and inconsistency was observed in the data from children with very low or very high temperature.

The GRADE evidence profiles for this review question are presented in Table 5.51.

Table 5.51 GRADE profile of study quality for change in heart rate with change in body temperature

Number of studies	Number of children	Quality
Change in heart rate (with increasing body temperature)		
1 study (Davies, 2009)	21,033[a]	Very low
Change in heart rate (with increasing body temperature)[e]		
1 study (Thompson, 2009)	1,589[b]	Low
Change in heart rate (with increasing body temperature)		
1 study (Hanna, 2004)	490[c]	Very low

[a] The data were analysed using a quantile regression and a statistical model to develop a best fit equation:

Expected parameter value = (Temperature (°C) × **a**) + (Age (months) × **b**) + (Age^2 (months^2) × **c**) + constant

The temperature multiplier **a** has a mean increase of 10.52 beats per minute (bpm) through the centile, resulting in a heart rate increase of approximately 10 bpm with each 1°C increment in temperature.

[b] Children were not truly representative of a primary care population due to problems with recruiting. Recruitment was not systematic, the proportion of children consulting out-of-hours care was high, and the researcher set the minimum recruitment targets for each age–temperature combination.

[c] Mean increase in pulse rate per 1°C increase in temperature was calculated using linear regression analysis of the relation between pulse rate and temperature. The authors report that for every 1°C rise in body temperature, the resting heart rate rose by 9.6 bpm.

Heart rate alone in the clinical assessment of serious illness

Three studies were considered that examined the use of heart rate for detecting serious illness.

The study by Brent (2011) found a positive association between the risk of serious bacterial infection and heart rates (probability [*P*] = 0.0005) (see Table 5.53 for GRADE profile). A correlation between tachycardia and serious bacterial infection was also found in this dataset (odds ratio [OR] 2.90, confidence interval [CI] 1.60 to 5.29; *P* = 0.0002). Table 5.52 shows diagnostic usefulness was high

for specificity at a cut-off above the 90th centile and moderately useful for sensitivity above a cut-off of 50% but low for everything else and the test was not useful in terms of LR+ or LR−.

In the second part of the Brent study, the usefulness of heart rate for detecting serious illness was assessed (see Tables 5.52 and 5.55).

A limitation in the first part of the study was the lack of a clear, gold standard for the definition of severe bacterial illness. The main limitation in the second part of the study was that the study included only children with meningococcal disease.

Table 5.52 Percentage sensitivity cut-offs defined by temperature heart rate centile, heart rate and tachycardia to distinguish between children with meningococcal septicaemia and those with severe disease

Age-specific pulse centiles	All children with meningococcal septicaemia (95%CI)	Children with severe disease on admission (95%CI)
Above 97th centile	11.0 (7.7 to 15.1)	17.9 (10.2 to 28.3)
Above 90th centile	27.8 (22.8 to 33.2)	38.5 (27.7 to 50.2)
Above 75th centile	49. (43.4 to 55.0)	61.5 (49.8 to 72.3)
Above 50th centile	73.9 (68.5 to 78.8)	84.6 (74.7 to 91.8)
Below 50th centile	26.1 (21.2 to 31.5)	15.4 (8.2 to 25.3)
Tachycardia	68.9 (63.3 to 74.1)	78.2 (67.4 to 86.8)

The study by Thompson (2009) examined tachycardia alongside other potential markers of serious illness (see Tables 5.56 and 5.57) for the GRADE profile). The study found a statistical relationship between children presenting with tachycardia and those found to have serious or intermediate infections ($P < 0.001$). However, the diagnostic value of the tachycardia was limited (sensitivity = 62 [95% CI 57 to 68], specificity = 58 [95% CI 53 to 63], positive LR = 1.5 [95% CI 1.3 to 1.7], negative LR = 0.7 [95% CI 0.6 to 0.8]). The study quality was limited due to the observational design that was used and the inclusion of children older than 5 years.

The study by Craig (2010) examined elevated heart rate alongside other potential markers of serious illness (see Table 5.57 for the GRADE profile). The study found a statistically significant relationship between elevated heart rates and serious bacterial illness in febrile children (OR 2.3 [1.7 to 3.1]). However, the diagnostic usefulness of elevated heart rate alone was limited (sensitivity = 58 [95% CI 55 to 61], specificity = 58 [95% CI 57 to 59], positive LR = 1.4 [95% CI 1.3 to 1.5], negative LR = 0.7 [95% CI 0.7 to 0.8]. The study quality was limited due to the observational design that was used and the inclusion of children older than 5 years.

Evidence profile

The GRADE profiles show results of included studies for the review question:

- Table 5.53 − GRADE profile for the distribution of age-specific heart rate data by centile group for 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection for the detection of serious illness
- Table 5.54 − GRADE profile for the sensitivity, specificity and positive and negative likelihood ratios for significant bacterial infection of cut-offs defined by pulse centiles in 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection for the detection of serious illness
- Table 5.55 − GRADE profile for the sensitivity of cut-offs defined by heart rate centiles for detecting meningococcal septicaemia of various degrees of severity in 325 children presenting to hospital with meningitis
- Table 5.56 − GRADE findings for evaluation of elevated heart rate

- Table 5.57 − GRADE findings for evaluation of elevated heart rate
- Table 5.57 − GRADE findings for evaluation of elevated heart rate

Table 5.53 GRADE profile for the distribution of age specific heart rate data by centile group for 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection for the detection of serious illness

Number of studies	Number of children		Effect		Quality
	Total	Children with SBI	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Detection of serious illness using heart rate above 97th centile					
1 (Brent, 2011)	28	1	OR 1.51 (0.19 to 12.0)	-	Very low
Detection of serious illness using heart rate above 90th centile					
1 (Brent, 2011)	91	10	OR 5.04 (2.14 to 11.9)	-	Low
Detection of serious illness using heart rate above 75th centile					
1 (Brent, 2011)	199	12	OR 2.62 (1.19 to 5.79)	-	Low
Detection of serious illness using heart rate above 50th centile					
1 (Brent, 2011)	324	14	OR 1.85 (0.87 to 3.93)	-	Very low
Detection of serious illness using heart rate below equal 50th centile					
1 (Brent, 2011)	586	14	OR 1.00 (Ref)	-	Low
Tachycardia					
1 (Brent, 2011)	514	34	OR 2.90 (1.60 to 5.26)	-	Low

OR odds ratio, SBI serious bacterial infection

Table 5.54 GRADE profile for the sensitivity, specificity and positive and negative likelihood ratios for significant bacterial infection of cut-offs defined by pulse centiles in 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection for the detection of serious illness

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Detection of serious illness using heart rate above 97th centile						
1 (Brent, 2011)	1360	2.0 (0.04 to 10.4)	97.7 (96.7 to 98.5)	2.7 (2.2 to 3.4)	0.96 (0.76 to 1.2)	Low
Detection of serious illness using heart rate above 90th centile						
1 (Brent, 2011)	1360	21.6 (11.3 to 35.3)	90.8 (89.0 to 92.4)	2.4 (1.6 to 3.7)	0.86 (0.57 to 1.3)	Low
Detection of serious illness using heart rate above 75th centile						
1 (Brent, 2011)	1360	45.1 (31.1 to 59.7)	75.7 (73.1 to 78.1)	1.7 (0.84 to 3.3)	0.78 (0.40 to 1.5)	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Detection of serious illness using heart rate above 50th centile						
1 (Brent, 2011)	1360	72.5 (58.3 to 84.1)	48.6 (45.7 to 51.5)	1.3 (0.58 to 3.1)	0.64 (0.28 to 1.5)	Low
Tachycardia						
1 (Brent, 2011)	1360	66.7 (52.1 to 79.2)	59.2 (56.3 to 62.0)	1.5 (0.67 to 3.4)	0.65 (0.29 to 1.46)	Low

Table 5.55 GRADE profile for the sensitivity of cut-offs defined by heart rate centiles for detecting meningococcal septicaemia of various degrees of severity in 325 children presenting to hospital with meningitis

Number of studies	Total number of children	Sensitivity (95% confidence interval)		Quality
		All children with meningococcal septicaemia	Children with severe disease on admission	
Detection of serious illness using heart rate above 97th centile				
1 (Brent, 2011)	325	11.0 (7.7 to 15.1)	17.9 (10.2 to 28.3)	Low
Detection of serious illness using heart above 90th centile				
1 (Brent, 2011)	325	27.8 (22.8 to 33.2)	38.5 (27.7 to 50.2)	Low
Detection of serious illness using heart rate above 75th centile				
1 (Brent, 2011)	325	49. (43.4 to 55.0)	61.5 (49.8 to 72.3)	Low
Detection of serious illness using heart rate above 50th centile				
1 (Brent, 2011)	325	73.9 (68.5 to 78.8)	84.6 (74.7 to 91.8)	Low
Detection of serious illness using heart rate below 50th centile				
1 (Brent, 2011)	325	26.1 (21.2 to 31.5)	15.4 (8.2 to 25.3)	Low

Table 5.56 GRADE findings for evaluation of elevated heart rate

Number of studies	Number of children		Effect		Quality
	Total	Children with serious bacterial infection (SBI)	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Tachycardia					
Thompson, 2009	691	191 of 307 compared to 160 of 384	2.3 (1.7 to 3.1)	-	Low

Table 5.57 GRADE findings for evaluation of elevated heart rate

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Elevated heart rate[a]								
For detecting pneumonia, urinary tract infection, or bacteraemia								
1 (Craig, 2010)	12,807	58 (55 to 61)	58 (57 to 59)	10 (9 to 10)	95 (94 to 95)	1.4 (1.3 to 1.5)	0.7 (0.7 to 0.8)	Low
1 (Thompson, 2009)	691	62 (57 to 68)	58 (53 to 63)	NR	NR	1.5 (1.3 to 1.7)	0.7 (0.6 to 0.8)	Low

NR Not reported

[a] Based on figures: Age (years) and recommended upper limit of normal for FEVER study (source): 0 = 160 (WHO); 1 = 150 (WHO); 2 = 150 (WHO); 3 = 140 (WHO); 4 = 130 (Wallis); 5 = 120 (Wallis). From: 1) Wallis et al, Arch. Dis. Child. 2005;90;1117-1121. 2) WHO. Pocket Book of Hospital Care for Children: Guidelines for the management of common illnesses with limited resources. 2005, page 232.

Heart rate alone and in conjunction with temperature in the clinical assessment of serious illness

Only one study was identified that addressed the review question. This was a cross-sectional prospective study (Brent et al., 2011) that included two datasets which were analysed and reported separately. The first included 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection; the second included 325 children presenting to hospital with meningitis. The study examined whether serious bacterial infection could be identified by heart rate in conjunction with temperature or heart rate alone.

Dataset including 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection

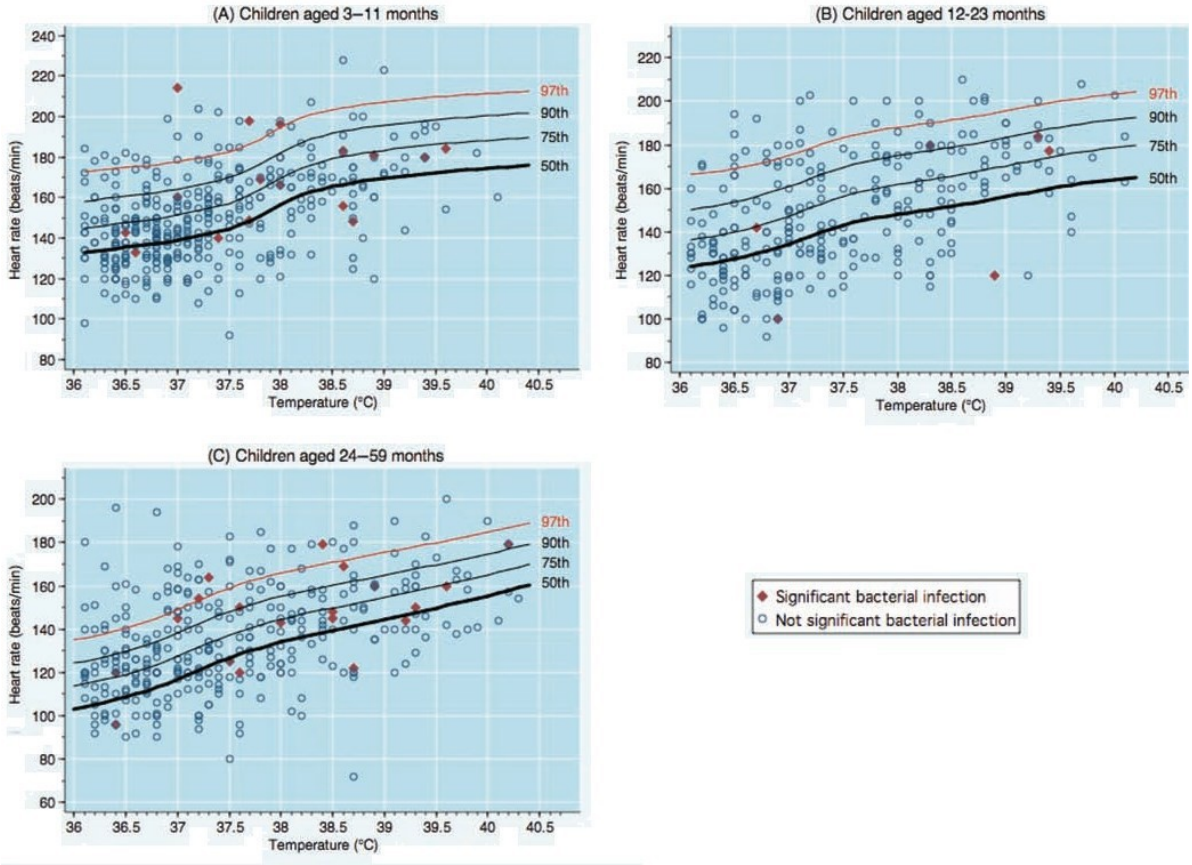
In the first part of the Brent study, age-specific centile charts of heart rate plotted against temperature were produced (see Figure 5.1). The distribution of children with or without serious bacterial infection and the odds ratios (OR) for serious bacterial infection were examined (see Table 5.59) and there was no significant trend across the temperature heart rate charts in the proportion of children with serious bacterial infection (P = 0.288). Table 5.60 shows that diagnostic usefulness was high specificity above 90th centile, but low for sensitivity, PPV and NPV, and the test was not useful in terms of LR+ or LR-.

Dataset including 325 children presenting to hospital with meningitis

In the second part of the Brent study, age-specific centile charts were plotted of heart rate against temperature involving children presenting at hospital with meningitis (see Figure 5.2). The sensitivity cut-offs defined by temperature heart rate centile, heart rate and tachycardia are shown in Table 5.58 (see also Table 5.61). Higher temperature and heart rate centile categories and higher heart rate centile categories showed a higher proportion of children with severe disease (P = 0.041 and P = 0.004, respectively).

A limitation in the first part of the study was the lack of a clear gold standard for the definition of severe bacterial illness. The main limitation in the second part of the study was that the study included only children with meningococcal disease.

Figure 5.1 Temperature and pulse of children presenting to the emergency department with and without significant bacterial infection (Brent et al., 2011) (Reproduced under open access publishing agreements)



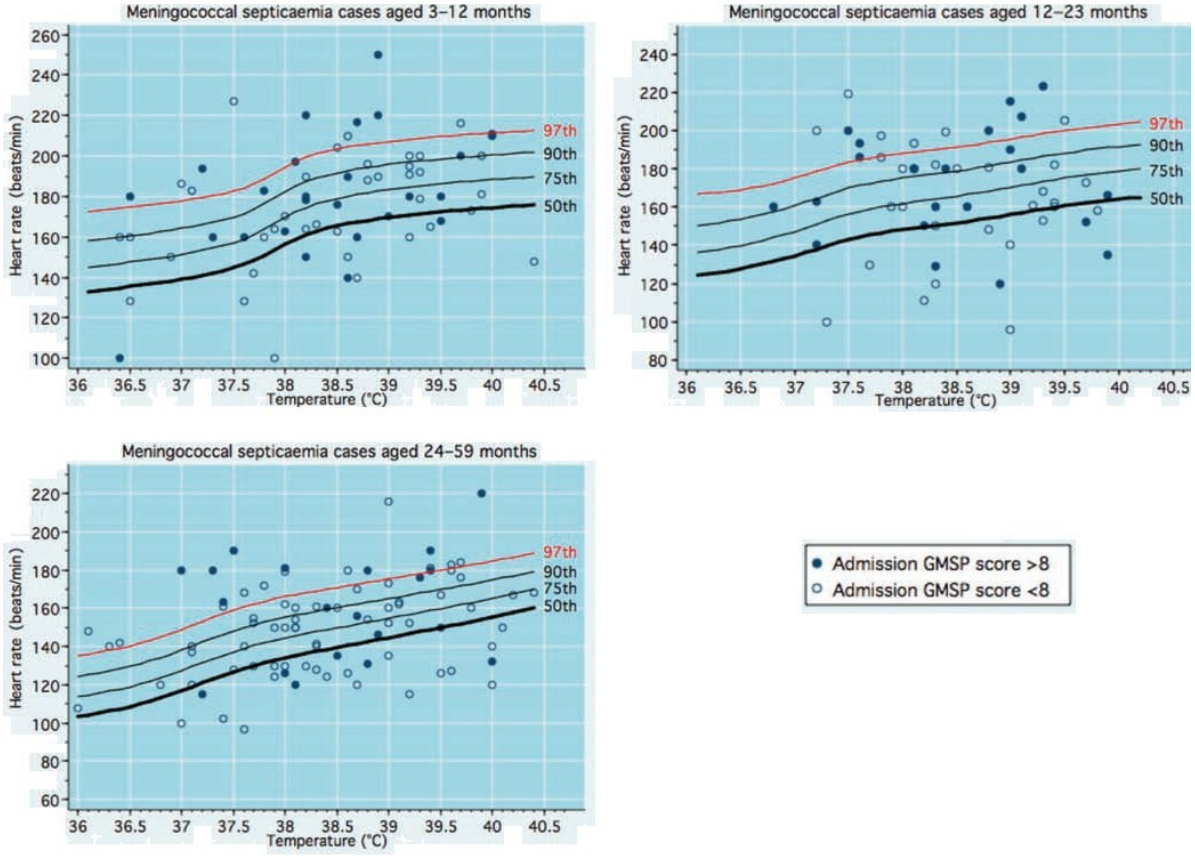
(Brent et al., 2011, Evaluation of temperature–pulse centile charts in identifying serious bacterial illness: observational cohort study Arch Dis Child 2011;96:368-373) (Reproduced under the open access publishing agreements)

Table 5.58 Percentage sensitivity cut-offs defined by temperature heart rate centile, heart rate and tachycardia to distinguish between children with meningococcal septicaemia and those with severe disease

Age-specific temperature–pulse centiles	All children with meningococcal septicaemia (95% CI)	Children with severe disease on admission (95% CI)
Above 97th centile	23.6 (18.5 to 29.3)	33.3 (22.9 to 45.2)
Above 90th centile	37.8 (31.8 to 44.1)	50.7 (38.9 to 62.4)
Above 75th centile	55.5 (49.2 to 61.7)	62.7 (50.7 to 73.6)
Above 50th centile	70.1 (64.0 to 75.6)	74.7 (63.3 to 84.0)
Below 50th centile	29.9 (24.4 to 36.0)	25.3 (16.0 to 36.7)

CI confidence interval

Figure 5.2 Admission temperature and pulse of children with meningococcal septicaemia, superimposed on proposed age-specific temperature–pulse centile charts. GMSP, Glasgow Meningococcal Septicaemia Prognostic score. (Brent et al., 2011) (Reproduced under open access publishing agreements)



(Brent et al., 2011, Evaluation of temperature–pulse centile charts in identifying serious bacterial illness: observational cohort study Arch Dis Child 2011;96:368-373) (Reproduced under the open access publishing agreements)

Evidence profile

The GRADE profiles presented show results of included studies for the review question.

- Table 5.59 − GRADE profile for the distribution of age-specific heart rate temperature data by centile group for 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection for the detection of serious illness.
- Table 5.60 − GRADE profile reporting the sensitivity, specificity, positive and negative likelihood ratio for significant bacterial infection of cut-offs defined by heart rate and body temperature for 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection.
- Table 5.61 − GRADE profile for the sensitivity of cut-offs defined by heart rate and body temperature centiles and tachycardia for detecting children with meningococcal septicaemia of various degrees of severity in 325 children presenting to hospital with meningitis.

Table 5.59 GRADE profile for the distribution of age-specific heart rate temperature data by centile group for 1,360 children presenting at a paediatric emergency department with suspected serious bacterial infection for the detection of serious illness

Number of studies	Number of children		Effect		Quality
	Total	Children with SBI	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Detection of serious illness using heart rate and temperature above 97th centile					
1 (Brent et al., 2011)	135	7	OR 1.84 (95% CI 0.72 to 4.71)	-	Very low
Detection of serious illness using heart rate and temperature above 90th centile					
1 (Brent et al., 2011)	110	4	OR 1.19 (95% CI 0.38 to 3.73)	-	Very low
Detection of serious illness using heart rate and temperature above 75th centile					
1 (Brent et al., 2011)	227	11	OR 1.67 (95% CI 0.73 to 3.79)	-	Very low
Detection of serious illness using heart rate and temperature above 50th centile					
1 (Brent et al., 2011)	316	16	OR 1.75 (95% CI 0.83 to 3.69)		Very low
Detection of serious illness using heart rate and temperature below or equal to 50th centile					
1 (Brent et al., 2011)	439	13	OR 1.00 (NR)		Low

CI confidence interval, OR odds ratio, NR not reported, SBI severe bacterial infection

Table 5.60 GRADE profile reporting the sensitivity, specificity and positive and negative likelihood ratios for significant bacterial infection of cut-offs defined by heart rate and body temperature for 1360 children presenting at a paediatric emergency department with suspected serious bacterial infection

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Detection of serious illness using heart rate and temperature above 97th centile						
1 (Brent et al., 2011)	1360	13.7 (5.7 to 26.3)	89.4 (87.5 to 91.1)	1.4 (0.69 to 2.7)	0.96 (0.48 to 1.9	Low
Detection of serious illness using heart rate and temperature above 90th centile						
1 (Brent et al., 2011)	1360	21.6 (11.3 to 35.3)	80.0 (77.6 to 82.3)	1.2 (0.76 to 1.8)	0.96 (0.63 to 1.5)	Low
Detection of serious illness using heart rate and temperature above 75th centile						
1 (Brent et al., 2011)	1360	43.1 (29.3 to 57.8)	61.7 (58.8 to 64.5)	1.2 (0.58 to 2.3)	0.90 (0.45 to 1.8)	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Detection of serious illness using heart rate and temperature above 50th centile						
1 (Brent et al., 2011)	1360	74.5 (60.4 to 85.7)	36.2 (33.4 to 39.0)	1 (0.50 to 2.6)	0.75 (0.33 to 1.7)	Low
Tachycardia						
1 (Brent et al., 2011)	1360	66.7 (52.1 to 79.2)	59.2 (56.3 to 62.0)	1.5 (0.67 to 3.4)	0.65 (0.29 to 1.46)	Low

Table 5.61 GRADE profile for the sensitivity of cut-offs defined by heart rate and body temperature centiles and tachycardia for detecting children with meningococcal septicaemia of various degrees of severity in 325 children presenting to hospital with meningitis

Number of studies	Total number of children	Sensitivity (95% confidence interval)		Quality
		All children with meningococcal septicaemia	Children with severe disease on admission	
Detection of serious illness using heart rate and temperature above 97th centile				
1 (Brent et al., 2011)	325	23.6 (18.5 to 29.3)	33.3 (22.9 to 45.2)	Low
Detection of serious illness using heart rate and temperature above 90th centile				
1 (Brent et al., 2011)	325	37.8 (31.8 to 44.1)	50.7 (38.9 to 62.4)	Low
Detection of serious illness using heart rate and temperature above 75th centile				
1 (Brent et al., 2011)	325	55.5 (49.2 to 61.7)	62.7 (50.7 to 73.6)	Low
Detection of serious illness using heart rate and temperature above 50th centile				
1 (Brent et al., 2011)	325	70.1 (64.0 to 75.6)	74.7 (63.3 to 84.0)	Low
Detection of serious illness using heart rate and temperature below 50th centile				
1 (Brent et al., 2011)	325	29.9 (24.4 to 36.0)	25.3 (16.0 to 36.7)	Low

Evidence statements

How heart rate changes with temperature

Three studies (one retrospective and two prospective) evaluated how heart rate changes with temperature in children with self-limiting infections. The studies reported that heart rate increased approximately 10 bpm with each 1°C increment in temperature. The studies were of low quality.

Using changes in heart rate alone to detect serious illness

Three prospective observational studies examined if heart rate could be used to identify children with bacterial infection, and to differentiate between serious and non-serious infection. The studies

reported that the risk of serious bacterial infection increased with higher heart rate centile ranges. They also showed a tendency to include a higher proportion of children with severe disease in higher heart rate centile categories. The studies were of low quality.

Using changes in heart rate adjusted for temperature to detect serious illness

One prospective study containing two datasets examined if age-specific centile charts of pulse rate plotted against temperature could be used to identify children with bacterial infection, and to differentiate between serious and non-serious infection. This study reported that there were no significant trends across heart rate/body temperature centiles that enabled identification of children with a severe illness. The study was of low quality.

Health economic evidence statements

No health economic studies were identified and no health economic evaluation was undertaken for this question.

Evidence to recommendations

Relative value placed on the outcomes considered

The GDG stated that the overarching aim of the guideline was the early and accurate detection of serious illness in children with fever. This allows for suitable treatment to begin, which will then reduce mortality and morbidity.

Consideration of clinical benefits and harms

The GDG stated that, to their knowledge, all the relevant studies had been included in the review.

How heart rate changes with temperature

The GDG highlighted that the results of the updated review supported the conclusion of the 2007 guideline, which was that heart rate and temperature are associated with approximately a 10 bpm increase for each 1ºC increase in temperature.

Using changes in heart rate alone to detect serious illness

The GDG reviewed evidence on the association between unadjusted heart rate and serious illness. Based on the papers presented, the GDG concluded that there was sufficient evidence to support the inclusion of tachycardia in the traffic light table.

The GDG emphasised that heart rate would vary by age and this would also have to be taken into account in any assessment, and as a result the GDG wanted to provide reference ranges for elevated heart rate. This led to a discussion on available reference ranges. The figures used for the Brent study were not available, while those used for the Craig study are shown in Table 5.62.

Table 5.62 Reference ranges for elevated heart rate used in the Craig study

Age (years)	Recommended upper limit of normal
0	160 (WHO)
1	150 (WHO)
2	150 (WHO)
3	140 (WHO)
4	130 (Wallis)
5	120 (Wallis)

Source: Wallis, Arch. Dis. Child. 2005;90;1117-1121. WHO. Common surgical problems. Ch9 Pocket Book of Hospital Care for Children. Guidelines for the management of common illnesses with limited resources. 2005, p232.

The GDG members stated that in their experience one of two recognised standards were usually used to assess heart rate in children; these being the Advanced Paediatric Life Support (APLS) and Pediatric Advanced Life Support (PALS) (see Table 5.63). The GDG stated that APLS was the most

commonly used scale in the UK, was simple to apply and closely matched the cut-offs used in the Craig study, which had shown an association between tachycardia and serious illness.

Table 5.63 Normal ranges of heart rate according to Advanced Paediatric Life Support (APLS) and Pediatric Advanced Life Support (PALS)

Age range (years)	APLS	PALS
Neonate	110 – 160	85 – 205
0 – 1	110 – 160	100 – 190
1 – 2	100 – 150	100 – 190
2 – 3	95 – 140	60 – 140
3 – 5	95 – 140	60 – 140
5 – 6	80 – 120	60 – 140

APLS Advanced Paediatric Life Support, PALS Pediatric Advanced Life Support

Source: Fleming et al, 2011, Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies, The Lancet 2011; 377: 1011–18

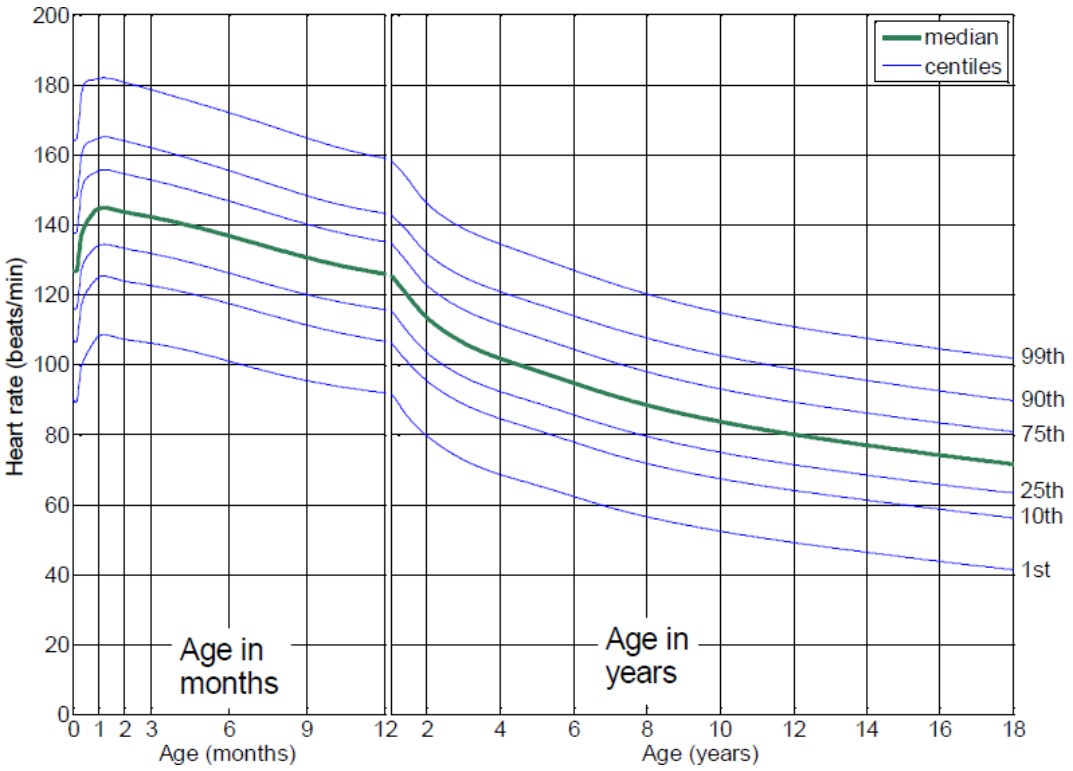
However, the GDG also highlighted the findings of a systematic review of normal heart rates in children (Fleming et al., 2011). This review contained data on heart rate in children from 59 studies that included 143,346 children (see Table 5.64).

Table 5.64 Normal ranges of heart rate according to the Fleming study

Age range	10th centile	25th centile	Median	75th centile	90th centile
Birth	107	116	127	138	148
0 – 3 months	123	133	143	154	164
12 – 18 months	103	112	123	132	140
18 – 24 months	98	106	116	126	135
2 – 3 years	92	100	110	119	128
3 – 4 years	86	94	104	113	123
4 – 6 years	81	89	98	108	117

(Fleming et al, 2011, Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies, The Lancet 2011; 377: 1011–18)

Figure 5.3 Centiles of heart rate for healthy children from birth to 18 yeas of age

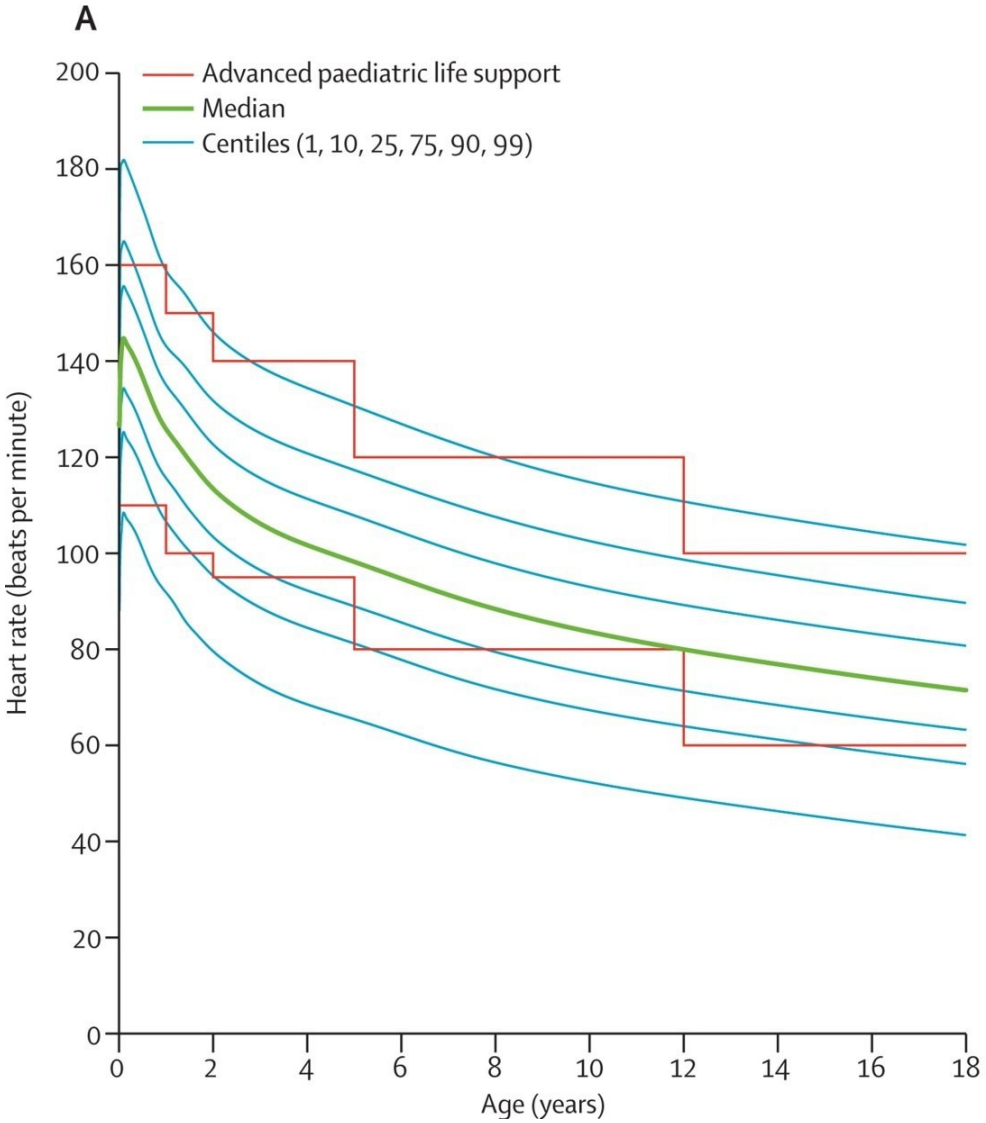


Source: Fleming et al, 2011, Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies, The Lancet 2011; 377: 1011–18 (Reproduced with permission, Elsevier Limited)

Fleming (2011) showed that there are inconsistencies between existing reference ranges and ranges of normal heart rate reported in observational studies. The authors demonstrated that this potentially leads to the misclassification of children as having either normal or abnormal heart rates, and that the use of updated centile heart rate charts could improve the specificity by up to 20%. However, the authors concluded that further research was needed before their centile charts could be adopted in practice.

Given this conclusion, the GDG decided that the APLS reference ranges were still the most practical and relevant cut-offs, and should continue to be used until the new centile charts had been validated. In addition, the GDG noted the APLS reference ranges and centile charts did overlap in children aged under 5 years.

Figure 5.4 Comparison of heart rate centiles from Fleming study with heart rate ranges from the advance paediatric life support.



Source: Fleming et al, 2011, Normal ranges of heart rate and respiratory rate in children from birth to 18 years of age: a systematic review of observational studies, The Lancet 2011; 377: 1011–18 (Reproduced with permission, Elsevier Limited)

The GDG also emphasised the difficulty of accurately measuring resting heart rate in children and that results varied depending on equipment used, so measurement error would also be a significant issue. For these reasons the GDG specified that heart rate should be added to the traffic light table in the 'amber' category, and should not be used in isolation to identify serious illness.

Using changes in heart rate adjusted for temperature to detect serious illness

The GDG concluded that the evidence on use of a combined temperature and heart rate measure did not support its inclusion in the traffic light table as it was shown to have less diagnostic value than either temperature or heart rate alone.

Consideration of health benefits and resource uses

The GDG emphasised that heart rate should be routinely recorded and health professionals should have been how trained in how to do this, so there were no resource implications associated with the implementation of this recommendation.

Quality of evidence

The available evidence was of low or very low quality due to serious illness not being fully defined, not all children receiving the same test and children older than 5 years being included.

Other considerations

No equalities issue were identified in relation to this question.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Blood pressure

Evidence summary

Blood pressure was not identified as an independent risk factor for serious illness in any of the prospective cohort studies and scoring systems. Low blood pressure was identified as one of several risk factors for adverse outcome in children with meningococcal disease.[119]

GDG translation

The GDG agreed with stakeholder comments that blood pressure should be measured in children with fever who are displaying features of possible serious illness. Blood pressure can be a helpful measurement to monitor children with possible sepsis although low blood pressure is a late feature of septic shock. Other markers such as raised heart rate and prolonged capillary refill time are present earlier and require no special equipment to measure. The GDG concluded that blood pressure should be measured when facilities exist to monitor blood pressure and other markers of inadequate organ perfusion (i.e. shock) are detected.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Assessment of dehydration

A number of studies have used degree of dehydration as a marker of serious illness. However, the symptoms and signs used in a number of studies have lacked rigour. The GDG looked for evidence for objective symptoms and signs for dehydration.

Narrative evidence

A recent EL 2+ SR[117] looking at children 1 month to 5 years was found. Although this SR only searched MEDLINE, it was judged to be adequate for inclusion. The authors reviewed 1603 papers, half of which were excluded because of lack of rigour or lack of clarity in outcomes. Of the remainder, only 26 were found to be rigorous enough to meet their criteria. Moreover, in this SR, dehydration was measured using percentage volume lost. They found three studies that evaluated the accuracy of a history of low urine output. A history of low urine output did not increase the likelihood of 5% dehydration (likelihood ratio [LR] 1.3, 95% CI 0.9 to 1.9). The most sensitive signs not requiring particular specialised tests for dehydration were dry mucous membranes, poor overall appearance, and sunken eyes and absent tears (see Table 5.3 for the sensitivities). Prolonged capillary refill time, cool extremities, reduced skin turgor and abnormal respiratory pattern were the most specific individual signs of dehydration.

Evidence summary

It is difficult to detect dehydration in children with fever. Individual symptoms and parental observations are poor predictors of dehydration. Furthermore, history of low urine output does not increase the risk of dehydration. The results showed that prolonged capillary refill time, reduced skin turgor and abnormal respiratory pattern are the most specific individual signs of dehydration.

Table 5.65 Summary characteristics for clinical findings to detect 5% dehydration[117]

Clinical feature	Sensitivity (95% CI)	Specificity (95% CI)
Prolonged capillary refill time	0.60 (0.29 to 0.91)	0.85 (0.72 to 0.98)
Abnormal skin turgor	0.58 (0.40 to 0.75)	0.76 (0.59 to 0.93)
Abnormal respiratory pattern	0.43 (0.31 to 0.55)	0.79 (0.72 to 0.86)

Clinical feature	Sensitivity (95% CI)	Specificity (95% CI)
Sunken eyes	0.75 (0.62 to 0.88)	0.52 (0.22 to 0.81)
Dry mucous membranes	0.86 (0.80 to 0.92)	0.44 (0.13 to 0.74)
Absent tears	0.63 (0.42 to 0.84)	0.68 (0.43 to 0.94)
Increased heart rate	0.52 (0.44 to 0.60)	0.58 (0.33 to 0.82)
Sunken fontanelle	0.49 (0.37 to 0.60)	0.54 (0.22 to 0.87)
Poor overall appearance	0.80 (0.57 to 1.04)	0.45 (−0.1 to 1.02)
Cool extremities	0.10–0.11 (range)	0.93–1.00 (range)

GDG translation

The GDG recognised that dehydration is a marker of serious illness but there was a lack of evidence to determine the difference between mild, moderate and severe dehydration. The most specific symptoms and signs of dehydration have been highlighted for healthcare professionals to assess to ensure a low false positive rate. The most sensitive symptoms and signs have been highlighted for parents to assess to ensure a low false negative rate (see Chapter 10).

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

5.5 Symptoms and signs of specific serious illnesses

This section was partially updated in 2013.

Introduction

The next priority in the assessment of a child with a feverish illness is to determine the underlying source of their illness.

The guideline is not meant to be a textbook on how to examine a child for all possible infections. However, the scope does include ‘identification of signs and symptoms that would help to establish the possible diagnoses and focus for infection’. The GDG focused on those serious illnesses that may have immediate consequences to the child’s life expectancy or long-term quality of life.

The GDG looked at those symptoms and signs that are predictive of specific serious illnesses, which are:

- bacterial meningitis
- septicaemia
- bacteraemia
- pneumonia
- urinary tract infection
- encephalitis (herpes simplex)
- septic arthritis/osteomyelitis
- Kawasaki disease.

The databases were searched and the highest evidence levels, i.e. prospective cohort studies, were used when evidence was available. Retrospective studies were included when there is a lack of better quality studies. The data were appraised, summarised and translated by the GDG members.

Review question

In children with fever, what symptoms and signs or combinations of symptoms and signs are predictive of the specific conditions defined as serious illnesses?

For the summary table for symptoms and signs suggestive of specific diseases see www.nice.org.uk/guidance/ng143

Meningococcal disease

Narrative evidence and summary

Three EL 2+ prospective population-based studies[94,118,132] to determine the clinical predictors of meningococcal disease in children with a haemorrhagic (non-blanching) rash with or without fever were found. The children's ages ranged from > 1 month[94,118,132] to < 16 years[132] and the population varied from Denmark,[132] and the UK[118] to the USA.[94] The features that helped predict the presence of meningococcal disease were:

- distribution of rash below the superior vena cava distribution (OR 5.1[132])
- presence of purpura – lesions > 2 mm (OR 7.0[132]; 37.2[118])
- neck stiffness (OR 6.9[132])
- capillary refill time > 2 seconds (OR 29.4[118])
- ill appearance (OR 16.7[118])
- CRP > 6 mg/litre.[118,132]

One recent UK-based EL 3 retrospective study[133] was also found that aimed to determine the frequency and time of onset of clinical features of meningococcal disease, to enable clinicians to make an early diagnosis before the individual was admitted to hospital. The researchers found that most children had only non-specific symptoms in the first 4–6 hours, but were close to death by 24 hours. The classic features of haemorrhagic rash, meningism and impaired consciousness developed later (median onset 13–22 hours). In contrast, 72% of children had earlier symptoms (leg pains, cold hands and feet, abnormal skin colour) that first developed at a median time of 8 hours.

GDG translation

The GDG considered a non-blanching rash (petechiae or purpura), neck stiffness and ill appearance on clinical examination as being 'red' features.

The feature of rash below the nipple line was not included in the traffic light table. This is because the sign is more useful in ruling out meningococcal disease if the rash is only found in the superior vena cava distribution rather than ruling the diagnosis in.

The GDG decided that they could not make a recommendation based on the possible early features of meningococcal disease[133] because of the retrospective nature of the study, the lack of controls and the possibility of recollection bias. The GDG did appreciate the potential benefit of diagnosing meningococcal disease at an early stage and called for further, prospective, research on this subject.

The updated review for capillary refill time was undertaken as part of the main symptoms and signs review and can be found in section 5.4.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Non-meningococcal septicaemia

No prospective population studies were found which determined the clinical features of non-meningococcal sepsis. Papers on occult pneumococcal bacteraemia were excluded as they only included laboratory screening test data. After searching for retrospective studies in the recent 10 years, there was no study judged to be of good enough quality to base recommendations upon and therefore none have been made.

Bacterial meningitis

Two EL 2+ prospective population studies[134,135] and one EL 2- narrative review[136] on determining the symptoms and signs of bacterial meningitis were found. Neck stiffness and a decreased conscious level are the best predictors of bacterial meningitis. However, neck stiffness is absent in 25% of infants under 12 months.[134] (EL 2+) Infants under 6 months of age have a bulging fontanelle in 55% of bacterial meningitis cases.[134](EL 2+)

A third EL 2+ prospective population study to determine the causes of status epilepticus in children was submitted by the GDG.[137] In this UK study, 17% of children with a first-ever febrile convulsive status epilepticus had bacterial meningitis.

GDG translation

The GDG considered neck stiffness, a bulging fontanelle and a decreased conscious level as being 'red' features. Although the management of febrile convulsions is outside the scope of the guideline the GDG felt it important to highlight the risk of bacterial meningitis in children with a prolonged febrile seizure. The GDG also felt it was important to highlight to healthcare professionals that classical features of bacterial meningitis are often absent in infants.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Herpes simplex encephalitis

Narrative evidence and summary

Only one EL 3 retrospective case series[138] conducted in Scotland was found which looked at the signs of herpes simplex encephalitis (HSE) in children. Focal neurological signs (89%) and seizures (61%), especially focal seizures, were the most frequent signs of HSE, but also neck stiffness (65%) and a decreased conscious level (52%).

GDG translation

Although the evidence was weak, the GDG felt that it was important to highlight these signs because early treatment of HSE improves outcomes.

The GDG considered neck stiffness, focal neurological signs, partial (focal) seizures and a decreased conscious level as being 'red' features.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Pneumonia

Narrative evidence and summary

Six EL 2+ prospective studies[139–144] that looked at clinical features of pneumonia were found. The study sites varied widely, from the USA,[139,140] the Philippines,[141] India[142] and Jordan[143] to Lesotho.[144] The age included also varied from 2 years[140] to < 6 years.[143]

Respiratory rate is a useful marker of pneumonia. Using age-related respiratory rates for tachypnoea (> 59 breaths/minute in the age group 0–5 months, > 52 breaths/minute in the age group 6– 12 months and > 42 breaths/minute in the age group > 12 months) there is a relative risk (RR) of 7.73[140] of having radiological signs of pneumonia. Other overall findings are:

- presence of cough has a sensitivity of 98% and specificity of 70% in children admitted for pneumonia[143]
- crepitations has a RR of 16.2[142]
- cyanosis has a RR of 4.38[142]
- oxygen saturations = 95% have an RR of 3.5[139]
- chest indrawing has an RR of 8.38[142]
- nasal flaring if age <12 months has an adjusted OR of 2.2)[139]

There are difficulties with all the studies in that the gold standard for diagnosing bacterial pneumonia is not specific as viral pneumonia cannot be confidently excluded on chest X-ray.

GDG translation

None of the signs for pneumonia are diagnostic in isolation. Not all of the signs found in the evidence were appropriate to the UK population. The GDG considered a respiratory rate of > 60 breaths/minute, moderate/severe chest indrawing, 'ashen' or 'blue' skin colour and grunting as being 'red' features. The GDG considered tachypnoea, nasal flaring and oxygen saturations ≤ 95% in air as being 'amber' features.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Urinary tract infection

Refer to the NICE *Urinary Tract Infection in Children* (UTIC) guideline for the summary of evidence and translation.

The recommendations below have been adapted from the NICE UTIC guideline as the scope of the two guidelines overlapped. The recommendation for children over 3 months has been altered as the population for whom this guideline applies all have a feverish illness.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Septic arthritis/osteomyelitis

Narrative evidence and summary

One EL 2+ prospective validation US study[145] of a clinical decision rule for a septic hip that recruited 51 children (age not specified) with septic arthritis was found. The study used two clinical features (fever and ability to bear weight on affected limb) and two laboratory features (erythrocyte sedimentation rate (ESR) and white blood cell count (WBC)). These performed well when all the features were available to assess. It was felt that the evidence for using the signs without blood tests was inadequate to base recommendations upon, and thus retrospective studies were searched for. Three EL 3 retrospective studies for osteomyelitis/septic arthritis[146–148] conducted in Taiwan,[146] Malaysia[147] and Nigeria[148] were found. The extra signs detected by retrospective studies were swelling of an affected limb and the limb not being used.

GDG translation

Recommendations have only been made for the clinical features, as definitive diagnosis of septic arthritis and/or osteomyelitis is beyond the scope of the guideline. The GDG considered non-weight bearing, swelling of a limb or joint and not using an extremity as being 'amber' features.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

.

Kawasaki disease

This section has been updated in 2019. See the current evidence review and recommendations at www.nice.org.uk/guidance/ng143

Narrative evidence and summary

No prospective studies looking at clinical features that are predictive of Kawasaki disease were found and thus retrospective studies from the past 10 years were searched for.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Research recommendations

Number	Research recommendation
	Symptoms and signs of serious illness
RR 2	The GDG recommends a UK-based epidemiological study on the symptoms and signs of serious illness. [new 2013]. Why this is important The current recommendations on symptoms and signs in the NICE guideline are based on a series of heterogeneous studies (using different methods, populations, outcomes and of varying quality) and a degree of subjectivity was needed to bring these together in the guideline. Therefore, the GDG recommends that a large prospective UK-wide study (n = 20,000 plus) should be undertaken comparing all of these symptoms and signs covered in the guideline. This would allow for a standardised comparison of each symptom and sign, and for validation of the existing 'traffic light' table. The study should use a standardised data collection protocol. Where possible the study should link with routinely collected data sets, such as Hospital Episode Statistics. The study should include a variety of settings and locations – that is, wherever children present, including primary care. The primary outcome of the study should be the final diagnosis and results of treatment.

5.6 Imported infections

The management of children with imported infections is beyond the scope of this guideline. However, the GDG recognised that significant numbers of children do enter or return to the UK from overseas each year. Some of these children will have been in countries where tropical and sub-tropical infectious diseases such as malaria and typhoid fever are endemic. Accordingly, the GDG decided to make the recommendation below.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

6 Management by remote assessment

Introduction

When a concerned parent or carer decides to make contact with a healthcare professional about a feverish child, the initial contact may be by telephone and in these circumstances a remote assessment may be undertaken. In this context, 'remote' refers to the assessment of the child's symptoms carried out by an assessor who is geographically remote from the child. It is common practice for remote assessment to be carried out during the out-of-hours period and, similarly, remote assessment may be a prerequisite for patients requesting an urgent in-hours appointment with their GP. Specific advice lines also exist, such as the 0845 4647 service offered by NHS Direct.*999 calls to the ambulance service are similarly assessed in order to determine the urgency of the response required.

The purpose of the remote assessment is to identify the level of care the child needs and to refer to the most appropriate location of care to meet those needs within an appropriate time frame. This process will include the identification of those with potentially life-threatening compromise to airway, breathing, circulation and level of consciousness, those with symptoms suggestive of serious illness and also identification of those children who are most likely to have a self-limiting illness and for whom care at home is the most appropriate option.

The skills and experience of the healthcare professional carrying out the remote assessment will vary and their assessment may or may not be supported by decision support software or other paper- based protocols. Remote assessment can be difficult as the assessor has only the symptoms reported by the caller on which to base the assessment. An additional difficulty, particularly when assessing a small child, is that the quality of information reported by the caller is likely to be variable and may be influenced by parental/carer concern. Symptoms which concern one parent/carer may not concern another and similarly symptoms which concern a parent/carer may not be those which most concern a healthcare professional.

It is essential that listening and critical thinking skills are employed throughout the assessment in order to ensure that all cues are identified and interpreted appropriately. This will include taking into account the level of parental/carer concern, the cause of which may not be easy to pinpoint. At times, however, it will be possible to identify a likely cause of the fever and that being the case the appropriate guidance for that condition should be followed.

In some circumstances the child may not be geographically remote from the assessor but physical examination of the child may not fall within the scope of practice for that healthcare professional. The assessor may thus feel it is more appropriate to follow the remote assessment guidance rather than that for face-to-face assessment which takes into account signs found on physical examination.

6.1 Clinical assessment

It is assumed that children with feverish illnesses undergoing a remote assessment will have a clinical assessment as described in Chapter 5. By necessity, the emphasis will be on detecting symptoms rather than physical signs. The first priority is to identify any immediately life-threatening features, including compromise of the airway, breathing, circulation and level of consciousness. Children with feverish illness should then be assessed for the presence or absence of symptoms that predict the

* Please note that this service will be replaced by NHS 111, which is due to be implemented nationally in 2013.

risk of serious illness using the traffic light system (see Table 5.2). Finally, the healthcare professional should seek the presence of symptoms that might suggest a particular diagnosis.

6.2 Management according to risk of serious illness

Evidence summary and GDG statement

The guideline development group (GDG) sought evidence that might refer particularly to the clinical evaluation of risk of serious illness by remote assessment or might direct management in this situation. No additional studies were found to add to the body of evidence which is described in Chapter 5. None of the studies found were specific to remote assessment or gave an indication of the time frame within which interventions should occur. With the exception of studies concerning the subjective detection of fever by parents and carers (section 4.3), no studies were found validating symptoms reported by parents or carers on remote assessment.

In line with the evidence presented in Chapter 5, the GDG concluded that children with immediately life-threatening features should receive emergency care. Children with 'red' features should be referred for an urgent face-to-face assessment, preferably within primary care. Those with 'amber' features would also require a face-to-face assessment although usually there would be less urgency. As described in Chapter 5, children with 'green' features only are at very low risk of serious illness and can be cared for at home. For children requiring an urgent face-to-face assessment, the GDG felt it was important to define the time frame within which an urgent assessment should be carried out because children with 'red' features are at high risk of having a serious illness. The GDG was unable to achieve consensus among themselves about the time limit for an urgent assessment and this question was therefore put out to formal consensus. The GDG used the Delphi panel to establish the definition of 'urgent' in the context of referral for further assessment (see section 3.2).

Delphi consensus

Background

Parents or carers often phone healthcare professionals for advice (e.g. NHS Direct, GP surgery) when their child has a fever.

The GDG has identified a number of symptoms which may indicate SBI (such as bacterial meningitis or pneumonia) and should prompt a 999 call. Other symptoms have been identified which warrant an urgent referral for a face-to-face assessment.

Delphi statement 2.1

An urgent face-to-face assessment means that a child should be seen within:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
43 (83%)	5 (10%)	1 (2%)	0	3 (6%)	52	2

In the first round consensus (83%) was reached that an urgent face-to-face assessment means that a child should be seen within 2 hours.

Health economics

The GDG recognised that the requirement for a face-to-face assessment within 2 hours for children with 'red' features may have health economic implications. In particular, the recommendation could be seen as producing an increase in the number of children referred from remote assessment to face-to- face assessment within this timescale. A detailed justification of this recommendation on clinical and health economic grounds was therefore developed. This is included in the guideline as Appendix E. In summary, the GDG concluded that the recommendation on urgent assessment would not represent an uplift in the provision of care for the following reasons:

- Children with 'red' features are at significant risk of serious illness and death.
- The traffic light system would encourage the referral of children with 'red' features for urgent assessment while discouraging the referral of the much larger number of children with 'green' features and most children with 'amber' features.
- 2 hours is an existing standard for referral for face-to-face assessment by out-of-hours providers and NHS Direct.*

- Fewer than 3% children undergoing remote assessment are likely to have 'red' features. At present a greater proportion of children with fever undergoing assessment by NHS Direct* are referred for urgent consultation.

GDG translation

The GDG recognised that remote assessment of symptoms and signs can be difficult as the quality of the information provided can vary.

However, some children will need an immediate assessment in view of the serious nature of the symptoms or combination of symptoms reported.

Other children will need an urgent face-to-face review by a healthcare professional who can examine the child.

The GDG felt it was not appropriate to identify individual symptoms as immediately life threatening because healthcare professionals will need to make a judgment in individual cases, based on the overall picture described.

As a result of stakeholder feedback and to ensure clarity of the recommendation, the GDG made the decision to combine the recommendation about which children should have an urgent face-to-face assessment and the recommendation about the time frame within which that assessment should take place into a single recommendation.

The GDG recognised that owing to the limitations of remote assessment, some children who are not seriously ill will be referred for urgent face-to-face assessment based on symptoms reported but not subsequently confirmed on examination. Nevertheless, the health economic analysis suggested that the recommendation of a 2 hour limit for urgent assessment could save lives and would not present an undue burden to the health service.

The GDG recognised that there have been no prognostic or validation studies on the predictive value of symptoms reported to remote assessors in children with feverish illness. It was therefore decided to call for research in this area.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Research recommendations

Number	Research recommendation
	Management by remote assessment
RR3	The GDG recommends that a UK study is undertaken to determine the validity of symptoms reported on remote assessment for children with fever. [2007] Why this is important Traditionally, symptomatic patients have been assessed in a face-to-face setting but increasingly, remote assessment (for example, assessment over the telephone) determines the urgency of the patient's need, the level of care required and from that the most appropriate next step for the patient. This might include referral to emergency services, referral to acute or non-acute services or closing the call with self-care advice/support. Clinical and cost effectiveness will only be achieved through remote assessment if perceived need equates to actual need. There is currently a lack of data available that demonstrate the validity of remote assessment.

7 Management by the non-paediatric practitioner

Introduction

Parents or carers of young children may seek a face-to-face assessment of their feverish child or be directed to do so following a remote assessment. There are an increasing number of professionals who may make this assessment. These include their GP, a nurse-practitioner in a walk-in centre, a pharmacist or an emergency department doctor. This guideline uses the term non-paediatric practitioner for this group. The setting of the assessment, although important, is less relevant than the experience and training of the healthcare professional undertaking the assessment. For this reason, the guideline development group (GDG) has separated recommendations pertaining to the non- paediatric practitioner assessment from those of the paediatric specialist. It has been assumed throughout that both the paediatric specialist and non-paediatric practitioner have the skills required to make a clinical assessment of a feverish child.

The initial face-to-face assessment of the feverish child is very important. The vast majority of children presenting to the non-paediatric practitioner with fever will have a condition that can be diagnosed, assessed and treated appropriately there and then or with simple follow-up arrangements.

In some cases, following assessment, the non-paediatric practitioner may refer the child to paediatric services for an opinion, for further necessary investigations that cannot be carried out in primary care, or for further treatment and care.

Fever without apparent source

A small number of children with fever will present with no obvious underlying source, and a small number of these will have a serious illness requiring further investigation and treatment by a paediatric specialist.

It is not always possible to distinguish serious illness from non-serious illness in the early stages of the condition. Safety netting is therefore vital to ensure that parents/carers and clinician agree when further care should be accessed and how. This may include, but not exclusively, a fixed appointment, formal liaison with other parts of the health system such as out-of-hours providers, or simple advice.

Safety netting

Following a consultation and the making of a provisional diagnosis and management plan, it is good practice for the healthcare professional to consider the following three questions:

- If I am right, what do I expect to happen?
- How will we know if I am wrong?
- What should happen then?

Safety netting is not a new concept.[151] It may take a number of forms, from dialogue with carer/parent about ‘amber’ and ‘red’ symptoms and signs they should watch for, review after a set period or liaising with other healthcare services. Good safety netting ensures continuity of care and a provision for possible deterioration of a child.

The GDG was unable to be prescriptive about safety netting since this will be determined by the actual practitioner carrying out the assessment and their professional competences and the range of services available locally. For example, a rural GP might use a different set of safety nets than a nurse working in an urban walk-in centre when dealing with the same child.

The GDG felt that safety netting was particularly important when a child presents with 'amber' features (see below), which were not felt to require automatic referral to secondary care at that time.

7.1 Clinical assessment

It is assumed that children with feverish illnesses presenting to a non-paediatric practitioner will undergo a face-to-face clinical assessment as described in Chapter 5. The first priority is to identify any immediately life-threatening features, including compromise of the airway, breathing, circulation and level of consciousness. Children with feverish illness should then be assessed for the presence or absence of symptoms and signs that predict the risk of serious illness using the traffic light system (see Table 5.2). Finally, the healthcare professional should look for a focus of infection or other symptoms and signs that might suggest a particular diagnosis.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

7.2 Management according to risk of serious illness

Evidence summary

The GDG was unable to find evidence to direct the management of children with fever in terms of referral to specialist care or care at home according to the risk of serious illness.

GDG statement

After an assessment of a febrile child has been made, the non-paediatric specialist has the following management options:

- If a diagnosis has been reached:
- reassurance to parents and carers that this is a self-limiting illness
- explanation, discussion and organising treatment options
- home care advice and safety netting
- refer for specialist paediatric treatment.

If no diagnosis has been reached:

- reassurance to parents and carers that this is probably a self-limiting illness given the absence of significant symptoms or signs
- perform some tests to help determine the diagnosis
- provide a safety net
- refer for specialist paediatric assessment.

A feverish child considered to have an immediately life-threatening illness should be transferred without delay to the care of a paediatric specialist by the most appropriate means of transport (usually 999 ambulance).

Health economics

The GDG recognised that in order to improve the NHS's ability to detect serious illness in children, it might be necessary to assess more, both in primary care and secondary care. The GDG also recognised that the number of children with 'amber' features with no focus on infection is a small proportion of face-to-face and remote access healthcare contacts by children with fever, and children with 'red' features make up an even smaller proportion of these children. Data on this is lacking, but the GDG consensus was that a normal GP practice will see an incidence of 1/100 children/year with 'red' symptoms, and a district general hospital may see three patients a week.

Attempts at modelling this were made but the number of possible variables and lack of evidence regarding outcomes impeded these attempts (see section11.2).

GDG translation

The GDG determined that children with fever receiving non-specialist care should be referred or allowed home according to their risk of serious illness, as defined in the traffic light table. Children with 'red' features are at risk of serious illness and should usually be referred to a paediatric specialist by the most appropriate route. Children with 'amber' features are at intermediate risk and should be provided with a safety net that may also involve referral to a specialist. The decision as to what form the safety net takes will depend on the experience, training and expertise of the non-specialist clinician. It will also depend on the local health service configuration and the family's social situation.

The GDG recognised that adherence to the recommendations in this section may cause changes in referral patterns between primary and secondary care. The health economists attempted to model these patterns but could not find sufficient evidence about current referral patterns and the associated risks. The GDG called for research to be undertaken so that the health economic model could be populated.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Research recommendations

Number	Research recommendation
RR 4	The GDG recommends that research is carried out on referral patterns between primary and secondary care for children with fever, so the health economic impact of this and future guidelines can be estimated. [2007]

7.3 Tests by the non-paediatric practitioner

In children with fever who are not referred to hospital, the use of investigations is determined by both pragmatic factors and clinical value. The delay in obtaining results of blood tests may preclude their use in non-specialist care.

Review question

In children presenting to primary care with fever and no obvious focus of infection, what is the predictive value of the following investigations in identifying children with a serious illness?

- urinalysis
- chest X-ray
- pulse oximetry
- capillary glucose.

The use of pulse oximetry and capillary glucose in the evaluation of children with fever was discussed but no evidence was found for or against their use. The GDG was unable to make a recommendation about these two investigations. Evidence was available regarding the use of chest X-rays and urine testing.

Chest X-rays

The GDG considered the question whether clinical acumen plus chest X-ray is better than clinical acumen alone in diagnosing chest infection in children aged 2 months to 59 months.

Narrative evidence

One EL 1+ systematic review (SR)[152] including one randomised controlled trial (RCT)[153] investigating the effects of chest radiography for children with acute lower respiratory infections was identified. They found that the odds of recovery by 7 days were 1.03 (95% confidence interval [CI] 0.64 to 1.64). The odds ratio (OR) for remaining ill at both 4 and 14 days were 0.74 (95% CI 0.45 to 1.23) and 0.82 (95% CI 0.45 to 1.48) for the study and control group, respectively. Thirty-three percent of radiography participants and 32% of control participants made a subsequent hospital visit within 4 weeks (OR 1.02, 95% CI 0.71 to 1.48); 3% of both radiography and control participants were

subsequently admitted to hospital within 4 weeks (OR 1.02, 95% CI 0.40 to 2.60).

Evidence summary

There was one systematic review of chest radiographs in children who met the criteria for clinical pneumonia, which included only one randomised controlled trial. This study of 522 children aged 2 months to 5 years demonstrated that children with clinical features of pneumonia based on the World Health Organization (WHO) criteria were less likely to be prescribed antibiotics, more likely to be diagnosed with bronchiolitis and had exactly the same rates of recovery, repeat attendance rates and subsequent admission rates when compared with those children who underwent a chest X-ray.

GDG translation

The GDG felt that in the presence of clinical signs of pneumonia or bronchiolitis, a chest X-ray is of no added diagnostic benefit in ambulatory care.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Urinalysis

In children with fever, urine should be tested for infection as described in Urinary Tract Infection in Children.*

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

7.4 Use of antibiotics by the non-paediatric practitioner

There are two situations in which a GP or prescribing professional may want to give antibiotics to a child with fever in the absence of a firm diagnosis of a bacterial infection. These are, firstly, in a child who is not particularly unwell and where the focus of infection cannot be found or initially established, and, secondly, in a very unwell child where the prescribing professional wants to prevent deterioration before transfer to hospital. This guideline relates to fever in children in both circumstances. Antibiotics have sometimes been prescribed empirically in this situation. The rationale behind this is sometimes put that these antibiotics might treat an unapparent bacterial infection or prevent development of severe bacterial infection (SBI). The temptation for a healthcare professional to recommend antibiotics may be increased by parental expectations and pressure. However, inappropriate prescribing of antibiotics is a major cause of antibiotic resistance. Antibiotics also have adverse effects, commonly rash and diarrhoea but also severe reactions such as allergy, anaphylaxis and Stevens–Johnson syndrome.

The use of antibiotics in children without a specific bacterial infection is thus not regarded as good clinical practice except when meningococcal disease is suspected, where immediate parenteral benzylpenicillin is currently recommended.[154]

Oral antibiotics

Review question

What are the benefits and risks of giving oral antibiotics to febrile children with no known focus of infection and no symptoms or signs of serious illness?

Narrative evidence

Three studies were found that evaluated antibiotics in children with no major focus of infection and who were well appearing. Two were EL 2+ SRs comprising eleven and four papers, respectively.[155,156] They examined the effect of oral and parenteral antibiotics in preventing SBI in well-appearing

* See Urinary tract infection in children, NICE clinical guideline 54 (2007)

children with *Streptococcus pneumoniae* occult bacteraemia. Fewer cases of SBIs but not bacterial meningitis were observed to develop in those children treated with antibiotics, compared with those who were not (*P* = 0.003). Furthermore, both oral and parenteral antibiotics were found to be equally effective in preventing SBI, which resulted in extremely low rates of complications observed in both groups (pooled OR = 1.48 in each group). Similarly, in another EL 1+ RCT[157] which looked at the effect of antibiotic treatment (amoxicillin) for acute otitis media in children between 6 months and 2 years, there was a reduced risk of 13% in the persistence of symptoms on day 4 in the amoxicillin group compared with the group which did not take amoxicillin (risk difference 13%, 95% CI 1% to 25%). In addition, median duration of fever was 2 days in the amoxicillin group versus 3 days in the placebo group (*P* = 0.004). Analgesic consumption was also higher in the group that went without antibiotics during the first 10 days (4.1 versus 2.3 doses, *P* = 0.004). However, no significant difference was observed in duration of pain or crying. No otoscopic differences were observed at days 4 and 11, and hearing tests findings were similar in both groups at 6 weeks The researchers concluded that, since seven to eight children aged 6–24 months with acute otitis media needed to be treated with antibiotics to improve symptomatic outcome on day 4 in one child, the modest effect does not justify the prescription of antibiotics at first visit.

Decreasing inappropriate antibiotic prescribing for children may also help decrease antibiotic resistance. In Finland, after nationwide reductions in the use of macrolide antibiotics for outpatient therapy, there was a significant decline in the frequency of erythromycin resistance among group A streptococci.[158]

Evidence summary

There is some evidence that oral antibiotics may decrease the risk of developing complications in children with *Streptococcus pneumoniae* occult bactaeremia, but insufficient evidence to conclude that it prevents bacterial meningitis.

There was no significant difference between children who were treated with oral or parenteral antibiotics.

However, over 1000 children at risk of occult pneumococcal bacteraemia would need to be treated to possibly reduce one case of meningitis.[159] There is evidence that campaigns to reduce the prescription of oral antibiotics are associated with a reduction in antimicrobial resistance.[158]

Health economics

There are very wide variations at both local and national levels in both rates and costs of antibiotic prescribing, with little evidence of associated variations in morbidity from infections. A decrease in inappropriate prescribing might also reduce antibiotic resistance. A decrease in inappropriate antibiotic prescribing would provide a saving in the overall NHS prescribing costs and delay antibiotic resistence. It is also possible that reduced antibiotic prescribing might increase the need or demand for reassessment and hospital admission of a febrile child either during surgery hours or by out-of- hours service providers, but while it would be possible to undertake research to assess the impact on healthcare demand (and costs and savings) of changes in antibiotic prescribing for children with suspected SBI, the GDG did not identify relevant data on this for the guideline.

GDG translation

The vast majority of well-appearing children (97%) with fever without cause do not have occult bacteraemia, and they will therefore not benefit from empirical oral antibiotics.

Occult pneumococccal bacteraemia is likely to be reduced markedly after conjugate pneumococcal vaccine was introduced in the routine UK immunisation schedule in September 2006.

Even for infections such as otitis media, the modest effect does not justify the prescription of antibiotics at first visit (number needed to treat [NNT] = 7–8).

The GDG also recognised the risks of the unnecessary prescribing of antibiotics such as adverse side effects and the development of antimicrobial resistance. The GDG also acknowledged the possibility of cost savings.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Empirical treatment with parenteral antibiotics

Review question

When should children in primary care be treated with empirical parenteral antibiotics in an attempt to decrease mortality or morbidity?

Narrative evidence

Two studies[159,160] that reported on the effect of empirical antibiotics on reducing mortality and morbidity were identified. An EL 2++ SR[159] comprising 14 studies evaluated the effectiveness of such antibiotics in reducing case fatality in meningococcal disease in patients of all ages. Twelve of the papers contained information on parenteral antibiotics given before admission and outcome, of which eight showed that there was a beneficial effect in giving parenteral antibiotics before admission and four reported an adverse effect. Risk ratios for mortality in these studies ranged from 0.16 (95% CI 0.01 to 2.63) to 2.36 (95% CI 0.25 to 22.54). Only one study reported a statistically significant result (risk ratio 0.35, 95% CI 0.16 to 0.80).[161] Since the proportion of cases treated differed among the reported studies (differences ranged from 15% to 59%, chi-squared for heterogeneity was 11.02 (P = 0.09), I_2 = 46% [95% uncertainty interval 0% to 77%]), studies were reported and examined on an individual basis. The reviewers could not conclude whether or not antibiotics given before admission had an effect on case fatality. However, they stated that the data are consistent with benefit when a substantial proportion of cases are treated.

A recent EL 2++[160] case–control study that was not included in the SR was also found. The study looked at the use of parenteral penicillin by GPs who had made the diagnosis of meningococcal disease in 26 children who died from the condition, and 132 survivors. Administration of parenteral penicillin was associated with increased risk of death (OR 7.4, 95% CI 1.5 to 37.7). Children who received penicillin had more severe disease on admission (median Glasgow meningococcal septicaemia prognostic score 6.5 versus 4.0, P = 0.002). The association between parenteral penicillin and poor outcome may be because children who were more severely ill were given penicillin before admission.

Evidence summary

In meningococcal disease, the evidence cannot conclude whether or not parenteral antibiotics given before admission have an effect on case fatality. However, the data are consistent with benefit when a substantial proportion of cases are treated.

Health economics

Since the evidence of effectiveness is equivocal, the cost-effectiveness of parenteral antibiotics cannot be established.

GDG translation

The GDG noted that all good-quality evidence referred to meningococcal disease and therefore looked at meningococcal disease in great detail compared with the other SBIs. Meningococcal disease is the leading infectious cause of mortality among children in the UK. No evidence on empirical treatment with parenteral antibiotics was found for other conditions, including meningitis, and therefore these conditions do not appear in the evidence tables. However, the GDG noted that current advice on immediate treatment in primary care refers to meningitis as well as meningococcal disease.

Children with meningococcal disease may benefit from pre-admission parenteral antibiotics, especially if most children with meningococcal disease are treated.

The GDG considers that there is insufficient evidence of effectiveness or cost-effectiveness to change the current UK practice (to give parenteral antibiotics at the earliest opportunity). As with oral antibiotics, the difference in costs (including consumables) should be taken into account when prescribing. Treatment should normally be initiated with the drug with the lowest cost (taking consumables into account).

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

8 Management by the paediatric specialist

Introduction

Young children with fever presenting to a paediatric specialist may be assessed initially by a non-paediatric practitioner or they may present directly to specialist care. Those children referred by a healthcare professional after an initial assessment are probably in a higher risk group for having a serious illness than those who are self-referred, although some may be referred simply for the opinion of a specialist because of uncertainty. Children who are reassessed because of parental concerns are probably also in a higher risk group for having a serious illness. For this reason, the recommendations have been separated into the assessment made by the non-paediatric practitioner and by the paediatric specialist. It has been assumed that both the paediatric specialist and non-paediatric practitioner have the skills required to make a clinical assessment of a feverish child. However, it has also been assumed that the paediatric specialist will have the training to perform, and access to, some investigations that may be necessary to complete the assessment of some febrile children. Almost all the tests and initial management considered in this chapter are part of the standard package of routine care for children with suspected severe bacterial infection (SBI) referred for specialist paediatric management. The guideline has reviewed the evidence of effectiveness for each intervention individually. In cases where the clinical benefit of a specific test or intervention has not been established, the recommendation is that these tests should not be performed, thus increasing the potential cost-effectiveness of care in this setting.

8.1 Clinical assessment

It is assumed that children with feverish illnesses presenting to paediatric specialist care will be assessed or reassessed using the 'traffic light' features described in Chapter 5. In addition to looking for these features, the clinician will look for a focus of infection or other symptoms and signs that might suggest a particular diagnosis.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

8.2 Children less than 3 months old

Although fever in the young infant is relatively uncommon, when it occurs there is a higher risk of SBI than in later life. Hospital Episode Statistics suggest that the incidence of the serious illnesses defined in this guideline are 19,316 per 100,000 for infants less than 3 months old in England, compared with 1400 per 100,000 for all children less than 5 years old. The neonate is at risk of rapidly developing infection because of a relatively poorly developed immune system and of permanent disability, especially from meningitis. Babies born preterm or with low birthweight are particularly vulnerable.

The infections may be those acquired from the mother at the time of delivery (e.g. group B streptococcus), or hospital- or community-acquired infections. Rarely, -devastating infections such as disseminated herpes simplex may present in the neonatal period. The host response to these infections and those presenting later in early infancy is fairly non-specific. For this reason, the GDG decided to provide separate recommendations for this group.

Narrative evidence

The studies suggested that SBI, particularly bacterial meningitis and urinary tract infection (UTI), are more common in the first 3 months than later in childhood. Among a series of infants in this age group with fever, the incidence of SBI lies in the range 6–10%.[108,162,163]

Three EL 2+ studies[108,162,164] and an EL 2+ meta-analysis[163] were found suggesting that neither clinical examination alone nor any single test is able to identify those with SBI. However, clinical assessment

and investigations combined can help to identify those infants more likely to have SBI. These babies appear ill to the clinician and/or have one or more abnormal test results from the following:

- white blood cell count (WBC) > 15 × 10^9/litre
- urine microscopy > 10 WBC per high power field (hpf)
- cerebrospinal fluid (CSF) with > 8 WBC per hpf or positive gram stain
- if diarrhoea is present more than 5 WBC per hpf in stool.

Another meta-analysis[152] of febrile infants less than 3 months old studied the usefulness of chest X-rays. This showed that chest radiographs were normal in 361 infants without respiratory signs. However, of 256 infants with one or more respiratory sign, 85 (33.2%) had positive chest radiographs for pneumonia. Signs included tachypnoea more than 50 breaths/minute, rales (crackles), rhonchi (wheeze), coryza, grunting, stridor, nasal flaring and cough.

GDG translation

Because young infants with fever are at relatively high risk of SBI (especially meningitis) which cannot be predicted by clinical features alone, the guideline development group (GDG) concluded that, on the basis of clinical effectiveness and cost-effectiveness, all febrile infants less than 3 months old require basic investigation as well as observation. This is not a change to usual clinical practice for this patient group. Those in the high-risk groups (neonates and those appearing unwell or with WBC < 5 × 10^9/litre or > 15 × 10^9/litre) should also be investigated for meningitis and receive empirical parenteral antibiotics, since they have the highest risk of infection. The GDG was unable to recommend a specific cut-off level for C-reactive protein (CRP), but expected paediatric specialists to use the CRP result as part of their overall assessment of a child with fever.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

8.3 Children aged 3 months or older

This section was partially updated in 2013.

Investigation by the paediatric specialist

Young children with fever will present to the paediatric specialist in three groups. The first group will appear well, with no symptoms or signs of serious illness, the vast majority of these children having viral or self-limiting illnesses (children with only ‘green’ symptoms/signs). A few of these children will have bacterial infections but they will not be identifiable by clinical assessment alone. This is particularly true of children less than 3 months of age and for this reason their management by the paediatric specialist is covered in a dedicated section of this chapter (section 8.2). Information is required regarding which serious illnesses occur in well-appearing children with fever, together with evidence of which investigations may help to identify these children.

A second group of children will arrive appearing very unwell with symptoms and signs of serious illness (mostly ‘red’ symptoms/signs) and will often be given immediate empirical antibiotic treatment.

The final group comprises those children with fever displaying symptoms and/or signs which may indicate the presence of a serious illness (one or more ‘amber’ or ‘red’ symptoms/signs). Few investigations will give results quickly enough to definitively identify serious illness in this group. For example, bacterial cultures will identify those with bacterial meningitis or bacteraemia but these results take 24–36 hours to become available. Treatment for these conditions should not be delayed until these results are available. It may be that identification of serious infection comes from a combination of signs and symptoms as well as simple tests such as WBC, etc. Markers of inflammation (e.g. WBC, CRP) may help to identify children with serious illness.

One controversial area is occult bacteraemia. Well-appearing children with fever can have bacteria in their blood, often pneumococcus. Most of these children will clear the bacteria without any antibiotic treatment, whereas a few will go on to develop significant sequelae, such as persistent bacteraemia and meningitis. Most information on this condition is from the USA and Australia, with little if any from the UK. In the USA, meningococcal disease occurs much less frequently than in the UK. A raised WBC has been used in the USA to identify those at increased risk of occult bacteraemia; however, in the UK this might not detect cases of meningococcaemia, as only one-third of cases have a raised WBC on

presentation. US data on the prevalence and causes of occult bacteraemia need to be viewed cautiously and UK data sought. The pattern of occult pneumococcal bacteraemia is also likely to change in the UK in 2006–07 following the introduction of conjugate pneumococcal vaccine to the childhood immunisation schedule.

Review question

In a febrile child what is the predictive value of the following in detecting serious illness?

- WBC
- absolute neutrophil count (ANC)
- CRP
- procalcitonin (PCT)
- erythrocyte sedimentation rate (ESR)
- urinalysis
- lumbar puncture
- chest X-ray
- combination of those above.

Narrative evidence

White blood cell count

Nine studies[166–174] evaluating WBC as a diagnostic marker for serious illness were found. The age ranges for these studies were birth to 16 years but in seven studies the upper limit was 36 months (age range mode: 3–36 months). Conditions studied were serious bacterial infection (SBI), meningococcal disease (MCD), bacterial meningitis, occult bacterial infection (OBI) and bacterial pneumonia. The cut-off value for WBC ranged from 15 to 17.1 × 10^9/litre. The ranges of performance of WBC as a marker of the presence of these serious illnesses were reported as sensitivity 20–76%, specificity 58–100% and relative risk (RR) 1.5–5.56.

Although one EL II study[168] did demonstrate a 'perfect' specificity of 100% with a WBC of > 15 × 10^9/litre identifying all children with SBI, the next highest result was 77%. Another EL II study[175] demonstrated an increased prevalence of occult bacteraemia with increasing height of fever and increasing WBC, but this was a US study conducted before the introduction of the conjugate pneumococcal vaccine, recently added to the UK childhood immunisation programme. These data are therefore likely to be less useful now.

One EL II prospective cohort study[176] looked at the combination of WBC > 20 × 10^9/litre combined with fever > 39°C in identifying 'occult pneumonia' (i.e. those with no clinical evidence of pneumonia) in children less than 5 years old. Between 26% and 30% of children with both these features had pneumonia on chest X-ray.

Absolute neutrophil count

Three EL II studies[169–171]evaluating absolute neutrophil count (ANC) were found. Two looked at children aged 1–36 months[169,171] and one at children aged 3–36 months.[170] The studies evaluated markers to identify SBI and OBI or to differentiate invasive bacterial infection from localised bacterial or viral infection.[170] The cut-off values for ANC were 10.2,[169] 10.6[170] and 9.6 × 10^9/litre.[170] The ranges of performance of ANC in identifying SBI were reported as sensitivity 50–71%, specificity 76–83% and RR 1.5–6.4.

Sepsis and meningitis

In children greater than 3 months old, PCT was found to have a significantly better diagnostic performance than CRP or WBC in identifying sepsis, septic shock and meningitis. PCT is also excellent in discriminating between viral and bacterial, and localised and invasive, bacterial infections. There was variation in the cut-off values used for PCT in the studies, with 2 ng/ml being most commonly reported as the best cut-off for distinguishing these groups. PCT was also found to perform better than CRP in identifying bacterial infection in children who had developed fever less than 12 hours prior to presentation. However, the authors added that since the negative predictive value of PCT is not always 100%, it cannot be considered a gold standard and a normal PCT level could conceivably falsely reassure clinicians.[165]

Lower respiratory tract infection

Six of the studies looked at PCT as a marker for bacterial lower respiratory tract infection (LRTI) in children. Of these, three found PCT to be more effective than either CRP or WBC in differentiating

bacterial from viral LRTI, whereas the other three studies found PCT to be of little value. This inconsistency may have been due to difficulty and differences in the confirmation of bacterial LRTI and also confounded by the use of antibiotics prior to measurement of PCT. PCT is known to fall rapidly once a bacterial infection is appropriately treated compared with CRP, which will fall more slowly and may even rise initially.[165]

Fever without localising signs

In another EL II study,[178] the authors reported the results of PCT assessed in children with fever without localising signs. Children treated with antibiotics during the preceding 2 days were excluded. PCT was more sensitive (93% versus 79%) but less specific (74% versus 79%) than CRP for predicting SBI (bacteraemia, pyelonephritis, lobar pneumonia and meningitis) in children with fever without apparent source.

In addition to this systematic review,[165] one prospective EL II cohort study[167] studied 72 children 1– 36 months old with fever without apparent source. Eight (11.1%) children had SBI (one pneumonia, two meningitis, four septicaemia/occult bacteraemia, two pyelonephritis), In identifying SBI in this group, PCT at a cut-off value of 2 ng/ml showed a sensitivity of 50% and a specificity of 85.9%. In comparison, at a cut-off of 50 mg/litre, CRP showed a sensitivity and specificity of 75% and 68.7% respectively, while the Yale Observation Score had a sensitivity of 87.5% and specificity of67.2%.

Chest X-ray

The diagnostic performance of chest X-ray in children with fever without apparent source (FWS) in relation to WBC is described above. In addition, one EL 1b SR[179] and one EL II prospective cohort study[180] were found that examined the diagnostic performance of chest radiography in differentiating bacterial and viral pneumonia in children.

The SR looked at five studies which used credible reference standards for identifying bacterial and viral infection. The authors considered identification of a bacterial pneumonia to be a positive test and of a viral pneumonia to be a negative test. As a result of heterogeneity in the studies, the authors could not report on comparable measures of diagnostic accuracy for each of the five studies. Rather, the researchers calculated likelihood ratios (LRs) for each study, as a measure of clinical usefulness of the chest X-ray. Commenting that LRs between 0.5 and 2.0 are rarely clinically useful, they reported no LRs outside these levels in the studies reviewed. The authors concluded that no clinically useful degree of accuracy had been demonstrated with regard to differentiating bacterial from viral pneumonia using chest radiography.

In an EL II study[180] of children admitted to hospital with community-acquired pneumonia, those with bacterial pneumonia had a significantly higher incidence of alveolar infiltrates compared with those with exclusively viral disease (72% versus 49%, $P = 0.001$). In children with exclusively interstitial infiltrates, half had bacterial infection and half viral.

Evidence summary

In children older than 3 months with fever without apparent source who appear well, 5% will have a bacterial infection, likely to be UTI or pneumonia. Occult bacteraemia is not often seen in the UK and is likely to decrease with the introduction of the universal pneumococcal vaccination. The currently available tests (CRP, PCT and WBC) do not improve the detection of SBI in this group, compared with features from the Yale Observation Score (YOS).

WBC and ANC perform less well than either CRP or PCT in helping to identify the presence of SBI. A combination of temperature > 39°C and a WBC.> 20 × 10^9/litre does, however, have a high specificity for occult pneumonia. Evidence is conflicting regarding the performance of chest radiography in differentiating bacterial and viral pneumonia in children but, at best, it has limited clinical usefulness.

Few studies were found looking at the usefulness of markers of bacterial infection in the management of children with fever without apparent source presenting to the paediatric specialist who were considered sufficiently unwell that intravenous anti-bacterial treatment should be initiated empirically.

GDG translation

‘Green’ group

Because tests such as CRP, PCT and WBC do not improve the detection of SBI in this group, the GDG concluded that routine blood tests on well-appearing children with fever are not justified. This would not change current practice since well-appearing children over 3 months old with fever rarely have blood tests in the UK at present. In contrast, there is a significant risk of UTI in this group and only by testing the urine will this be identified.

‘Amber’ and ‘Red’ groups

Although PCT is more sensitive than CRP in identifying sepsis and meningitis in young children with

fever, the GDG did not feel that this difference was sufficient to recommend PCT over CRP, potentially changing current UK practice. The GDG noted that there was only limited evidence on the use of PCT in children with fever without apparent source, and they decided to call for more research in this area. In children with no symptoms or signs of pneumonia, a combination of temperature > 39°C and a WBC > 20 × 10^9/litre has a high specificity for bacterial pneumonia and therefore the GDG concluded that a chest X-ray is indicated in this small group of children. In children considered sufficiently unwell to require empiric antibiotics, the GDG acknowledged that the result of a CRP or WBC would not influence immediate management. However, they should be measured as an aid to ongoing management of this group.

Procalcitonin and C-reactive protein

Introduction

A review question comparing procalitonin (PCT) and C-reactive protein (CRP) was outlined as new evidence had become available since the 2007 guideline was published.

PCTCRP are found in the bloodstream and the levels of both increase in response to the presence of bacterial infection, but not (or less so) to viral illness. This response starts approximately 6 hours after the start of infection with PCT and 12 hours afterwards for CRP. The tests are used to differentiate between viral and bacterial infections, and to determine the seriousness of bacterial infection.

Review question

The clinical question set out on the scope asks for: 'The predictive value of pro-calcitonin and/or C-reactive protein markers.' This translates into the following review question "What is the predictive value of procalcitonin compared to C-reactive protein for detecting serious illness in fever without apparent source in children under 5?"

Overview of review

In the 2007 guideline the use of CRP was recommended, but not the use of PCT. A research recommendation was outlined stating the need for studies comparing PCT and CRP. The focus of this question was to review the literature comparing PCT and CRP.

A literature search was undertaken from 2005 onwards. A total of 594 studies were identified. In addition, studies included in the 2007 guideline were reviewed for inclusion in the updated guideline.

Description of included studies

A total of 16 observational studies were included in this review (Galetto-Lacour et al., 2003; Guen et al., 2007; Lacour et al., 2001; Thayyil et al., 2005; Manzano et al., 2011; Olaciregui et al., 2009; Andreola et al., 2007; Maniaci et al., 2008; Hsaio et al, 2006; Berger et al, 1996; Isaacman et a, 2002; Pratt et al, 2007 ; Pulliam et al, 2001; Gomez et al, 2010; Luaces-Cubells et al, 2012; Woelker et al, 2012). Fifteen of these assessed CRP and ten assessed PCT. Eight of these studies directly compared CRP with PCT (Galetto-Lacour et al., 2003; Guen et al., 2007; Lacour et al., 2001; Thayyil et al., 2005; Manzano et al., 2011; Olaciregui et al., 2009; Andreola et al., 2007; Luaces-Cubells et al, 2012). Six studies included CRP but not PCT (Hsaio et al, 2006; Berger et al, 1996; Gomez et al, 2010; Isaacman et a, 2002; Pratt et al, 2007; Pulliam et al, 2001). Two studies examined PCT only (Maniaci et al., 2008; Woelker et al, 2012). Fourteen studies were prospective studies and two were retrospective (Olaciregui et al., 2009; Gomez et al., 2010). All the studies investigated populations with suspected bacterial illness.

Six studies (Guen et al., 2007; Olaciregui et al., 2009; Isaacman et al, 2002; Pratt et al, 2007; Pulliam et al, 2001; Gomez et al, 2010) investigated the same infection (bacteremia); the rest assessed a range of infections. Five studies carried out subgroup analyses. Andreola (2007) performed a subgroup analysis by duration of evolution of fever. Lacour (2001) split the results by age (younger than12 months and 12 months or older). Manzano (2011) reported separate results for children who had normal urine analysis. Pratt (2007) examined differences between children who presented less than 12 hours after becoming febrile and those who presented more than 12 hours after becoming febrile. Luaces-Cubells (2012) examined result for children who presented 8 hour or less after fever had started. Two studies reported results of combined tests of PCT and CRP (Guen et al., 2007; Lacour et al., 2001).

A range of gold standard tests were then undertaken to confirm diagnosis in each child. Prevalence of bacterial illness ranged from 0.6% to 29% across studies. Reported average age ranged from 4 days to 36 months.

Further information is shown in the evidence tables.

Evidence profile

The GRADE profiles presented show results of included studies for the review question.

- Table 8.1 – GRADE findings for comparison of different procalcitonin thresholds
- Table 8.2 – GRADE findings for comparison of different C-reactive protein thresholds
- Table 8.3 – GRADE findings for comparison of combined procalcitonin and C-reactive protein thresholds

Table 8.1 GRADE findings for comparison of different procalcitonin thresholds

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Bacteremia, pyelonephriitis, pneumonia, bacterial meningitis, sepsis, bone infections. Prevalence = 23%								
0.5 ng/ml								
1 (Andreola et al, 2007)	N = 408	73.4 (63 to 82)	76 (71 to 81)	48 (40 to 56)[a]	91 (87, 94)[a]	3.1 (2.5, 3.9)[b]	0.4 (0.2, 0.5)[b]	Very Low
1 ng/ml								
1 (Andreola et al, 2007)	N = 408	64 (53 to 74)	90 (86 to 93)	65 (55 to 75)[a]	89 (85, 93)[a]	6.2 (4.4, 9.0)[b]	0.4 (0.3, 0.5)[b]	Very Low
2 ng/ml								
1 (Andreola et al, 2007)	N = 408	48 (38 to 58)	97 (94 to 98)	80 (70 to 91)[a]	86 (82 to 90)[a]	13.6 (7.4 to 25.3)[b]	0.5 (0.4 to 0.7)[b]	Very Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Fever < 8 hours								
1 ng/ml								
1 (Andreola et al, 2007)	N = 45	86 (Not reported)	100 (Not reported)	Not reported	Not reported	Not reported	Not reported	Very Low
Bacteraemia, pyelonephritis, lobar pulmonary condensation. Prevalence = 22.6%								
0.9 ng/ml								
1 (Lacour et al, 2001)	N = 124	93 (77 to 99)	78 (69 to 86)	55 (41 to 70)[a]	97 (94 to 101)[a]	4.2 (2.9 to 6.3)[b]	0.1 (0.0 to 0.3)[b]	Low
< 12 months of age								
1 (Lacour et al, 2001)	N = 80	94 (Not reported)	87 (Not reported)	68 (Not reported)	98 (Not reported)	Not reported	Not reported	Very low
> 12 months of age								
1 (Lacour et al, 2001)	N = 44	90 (Not reported)	62 (Not reported)	41 (Not reported)	96 (Not reported)	Not reported	Not reported	Very low
Bacteremia, pyelonephritis, pneumonia, mastoiditis and retropharyngeal abscess. Prevalence = 29%								
0.5 ng/ml								
1 (Galetto-Lacour et al, 2003)	N = 99	93 (77 to 99)	74 (62 to 84)	60 (46, 74)[a]	96 (91, 101)[a]	3.6 (2.4, 5.5)[a]	0.1 (0.0 to 0.4)	Low
Bacteremia, bacterial meningitis, sepsis, UTI, pneumonia, gastroenteritis, cellulitis. Prevalence = 23.6%								
≥0.5ng/ml								
1 (Olaciregui et al, 2009)	N = 347	63 (52 to 74)	87 (83 to 91)	59 (48 to 70)	89 (85 to 93)	4.8 (3.5 to 7.0)[b]	0.4 (0.3 to 0.5)[b]	Low
Bacteremia/sepsis. Prevalence = 0.6%								
> 0.5 ng/ml								
1 (Olaciregui et al, 2009)	N = 347	86 (58 to100)	93 (90 to 96)	35 (19 to 51)	99 (98 to100)	12.3 (Not reported)	0.2 (Not reported)	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Bacteremia, UTI and bacteremia/UTI . Prevalence = 13%								
0.13 ng/ml								
1 (Maniaci et al, 2008)	N = 234	97 (81 to 100)	30 (24 to 38)	17 (11 to 23)	98 (90 to 100)	1.4 (1.2 to 1.6)	0.1 (0.0 to 0.8)	Low
Bacteremia, UTI, bacteremia/UTI, bacterial pneumonia. Prevalence = 18%								
0.12 ng/ml								
1 (Maniaci et al, 2008)	N = 234	95 (83 to 99)	26 (20 to 32)	22 (16 to 28)	96 (85 to 99)	1.3 (1.1 to 1.4)	0.2 (0.1 to 0.7)	Low
Bacteremia, UTI, pneumonia and bacterial meningitis. Prevalence = 16%								
> 0.2 ng/ml								
1 (Manzano et al, 2011)	N = 328	85 (74 to 92)	70 (68 to 71)	36 (31 to 39)	96 (93 to 98)	2.8 (2.3 to 3.2)	0.2 (0.1 to 0.4)	Low
Children with normal urine analysis only								
> 0.2 ng/ml								
1 (Manzano et al, 2011)	N = 262	88 (54 to 98)	71 (69 to 71)	9 (5 to 10)	99 (98 to 100)	3.0 (1.8 to 3.3)	0.2 (0.0 to 0.7)	Low
Bacterial pneumonia, bacterial meningitis, septicaemia and pyelonephritis. Prevalence = 1.1%								
> 500 ng/ml (> 0.5 ng/l)								
1 (Thayyil et al, 2005)	N = 72	88 (65 to 110)[b]	50 (38 to 62)[b]	18 (6 to 30)2	97 (91 to 103)[b]	1.8 (1.2 to 2.5)[b]	0.3 (0.0 to 1.6)[b]	Very Low
> 2000 ng/ml (> 2 ng/l)								
1 (Thayyil et al, 2005)	N = 72	50 (15 to 85)[b]	86 (77 to 94)[b]	31 (6 to 56)[b]	93 (87 to 100)[b]	3.6 (1.4 to 8.9)[b]	0.6 (0.3 to 1.2)[b]	Very Low
Bacteremia. Prevalence = 3.2%								
≥ 2 ng/ml (± IC 95%)								
1 (Guen et al, 2007)	N = 215	57.1 ±0.37	86.4±0.05	13.8 ±0.26	98.1 ±0.06	4.19	0.49	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Invasive bacterial infections: Bacterial bacterial meningitis, occult bacteremia and sepis. Prevalence = 1.7%								
≥ 0.5 ng/mL								
1 (Luaces-Cubells et al, 2012)	N = 868	0.87 (0.60 to 0.98)	0.83 (0.81 to 0.86)	0.09 (0.05 to 0.14) [a]	1.00 (0.99 to 1.00) [a]	5.15 (4.04 to 6.66) [a]	0.16 (0.04 to 0.58) [a]	Low
≥ 0.9 ng/mL								
1 (Luaces-Cubells et al, 2012)	N = 868	0.87 (0.60 to 0.98)	0.91 (0.88 to 0.92)	0.14 (0.08 to 0.23) [a]	1.00 (0.99 to 1.00) [a]	9.13 (6.84 to 12.18) [a]	0.15 (0.04 to 0.54) [a]	Low
≥ 1 ng/mL								
1 (Luaces-Cubells et al, 2012)	N = 868	0.73 (0.45 to 0.92)	0.92 (0.89 to 0.93)	0.14 (0.07 to 0.23) [a]	0.99 (0.99 to 1.00) [a]	8.72 (5.97 to 12.73) [a]	0.29 (0.13to 0.67) [a]	Low
≥ 2 ng/mL								
1 (Luaces-Cubells et al, 2012)	N = 868	0.60 (0.32 to 0.84)	0.95 (0.94 to 0.97)	0.19 (0.09 to 0.33) [a]	0.99 (0.98 to 1.00) [a]	12.80 (7.65 to 21.41) [a]	0.42 (0.23 to 0.78) [a]	Low
Serious bacterial infections: Bacterial bacterial meningitis, occult bacteremia and UTI. Prevalence = 8.3%								
≥ 0.2 ng/mL								
1 (Woelker et al, 2012)	N = 155	1.0 (0.72 to 1.0)	0.41 (0.33 to 0.49)	0.13 (0.08 to 0.22) [a]	1.0 (0.92 to 1.0) [a]	1.69 (1.47 to 1.94) [a]	NC	Very low
≥ 0.26 ng/mL								
1 (Woelker et al, 2012)	N = 155	0.92 (0.62 to 1.0)	0.64 (0.55 to 0.72)	0.19 (0.11 to 0.31) [a]	0.99 (0.93 to 1.0) [a]	2.57 (1.96 to 3.37) [a]	0.12 (0.02 to 0.80	Very low
≥ 0.3 ng/mL								
1 (Woelker et al, 2012)	N = 155	0.85 (0.54 to 0.97)	0.64 (0.55 to 0.72)	0.19 (0.10 0.32) [a]	0.98 (0.92 to 1.0) [a]	2.56 (1.84 to 3.55) [a]	0.23 (0.06 to 0.83) [a]	Very low

CI confidence interval, UTI urinary tract infection

[a] Estimates and confidence intervals were calculated by the NCC-WCH technical team.

[b] Confidence intervals were calculated by the NCC-WCH technical team.

Table 8.2 GRADE findings for comparison of different C-reactive protein thresholds

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Bacteremia, pyelonephritis, pneumonia, bacterial meningitis, bone infections, sepsis. Prevalence = 23%								
20 mg/l								
1 (Andreola et al, 2007)	N = 408	88 (80 to 94)	61 (55 to 66)	40 (34 to 47)[a]	95 (91 to 98)[a]	2.3 (1.9 to 2.6)[b]	0.2 (0.1 to 0.3)[b]	Very Low
40 mg/l								
1 (Andreola et al, 2007)	N = 408	71 (61 to 80)	81 (76 to 85)	53 (44 to 66)[a]	90 (87 to 94)[a]	3.8 (2.9 to 4.9)[b]	0.4 (0.3 to 0.5)[b]	Very Low
80 mg/l								
1 (Andreola et al, 2007)	N = 408	46 (36 to 57)	95 (92 to 97)	72 (60 to 83)[a]	85 (82 to 89)[a]	8.7 (5.1 to 14.1)[b]	0.6 (0.5 to 0.7)[b]	Very Low
Bacteraemia, pyelonephritis, lobar pulmonary condensation. Prevalence = 22.6%								
40 mg/l								
1 (Lacour et al, 2001)	N = 124	89 (72 to 98)	75 (65 to 83)	96 (92 to 100)[a]	51 (37 to 65)[a]	3.6 (2.5 to 5.2)[b]	0.1 (0.0 to 0.4)[b]	Low
< 12 months of age								
1 (Lacour et al, 2001)	N = 80	94 (Not reported)	84 (Not reported)	63 (Not reported)	98 (Not reported)	Not reported	Not reported	Very low
> 12 months of age								
1 (Lacour et al, 2001)	N = 80	80 (Not reported)	59 (Not reported)	91 (Not reported)	36 (Not reported)	Not reported	Not reported	Very low
Bacteremia, pyelonephritis, pneumonia, bacterial meningitis and deep abscess. Prevalence = 29%								
40 mg/l								
1 (Galetto-Lacour et al, 2003)	N = 99	79 (65 to 94)	79 (69 to 88)	61 (45 to 76)	90 (83 to 98)	3.7 (2.3 to 6.0)	0.3 (0.1 to 0.5)	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Bacteremia, bacterial meningitis, sepsis, UTI, pneumonia, gastroenteritis, cellulitis. Prevalence = 23.6%								
≥ 20 mg/l								
1 (Olaciregui et al, 2009)	N = 347	64 (54 to 74)	84 (80 to 88)	55 (45 to 65)	88 (84 to 92)	4.0 (2.9 to 5.5)[b]	0.4 (0.3 to 0.6)[b]	Low
≥ 30 mg/l								
1 (Olaciregui et al, 2009)	N = 347	59 (48 to 70)	89 (85 to 93)	63 (52 to 74)	83 (87 to 91)	5.4 (3.6 to 7.9)[b]	0.5 (0.4 to 0.6)[b]	Low
Bacteremia, UTI, pneumonia and bacterial meningitis. Prevalence = 16%								
> 17.7 mg/l								
1 (Manzano et al, 2011)	N = 328	94 (96 to 98)	69 (67 to 69)	37 (34 to 39)	98 (96 to 100)	3.0 (2.6 to 3.2)	0.1 (0.0 to 0.2)	Low
Children with normal urine analysis only								
> 17.7 mg/l								
1 (Manzano et al, 2011)	N = 262	88 (54 to 98)	70 (69 to 70)	8 (5 to 9)	99 (98 to 100)	2.9 (2.4 to 3.5)	0.2 (0.1 to 0.4)	Low
Bacterial pneumonia, bacterial meningitis, septicaemia and pyelonephritis. Prevalence = 11.1%								
> 50 mg/l								
1 (Thayyil et al, 2005)	N = 72	75 (45 to 105)[b]	69 (57 to 80)[b]	23 (7 to 39)[b]	96 (90 to 102)[b]	2.4 (1.4 4.1)[b]	0.4 (0.1 to 1.2)[b]	Very Low
Bacteremia/sepsis. Prevalence = 0.6%								
> 30 mg/l								
1 (Olaciregui et al, 2009)	N = 347	56 (32 to 80)*	74 (69 to 79)	10 (4 to 16)	95 (97 to 99)	2.2 (Not reported)	0.6 (Not reported)	Very Low
Bacteremia. Prevalence = **3.2%**								
≥ 40 mg/l (± IC 95%)								
1 (Guen et al, 2007)	N = 215	42.8±0.37	64.8±0.07	3.8±0.22	97.2±0.06	1.21	0.88	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Bacteraemia, urinary tract infection. Prevalence = 10.3%								
> 2 mg/l								
1 (Hsiao et al, 2006)	N = 387	100 (89 to 100)	29 (24 to 34)	74 (69 to 79)	26 (22 to 31)	1.4 (1.3 to 1.5)	-	Very Low
> 5.2 mg/l								
1 (Hsiao et al, 2006)	N = 387	84 (70 to 94)	54 (49 to 60)	50 (45 to 55)	50 (45 to 55)	1.9 (1.6 to 2.3)	0.27 (0.13 to 0.57)	Very Low
> 9.8 mg/l								
1 (Hsiao et al, 2006)	N = 387	51 (31 to 67)	80 (76 to 84)	23 (16 to 34)	77 (72 to 81)	2.6 (1.8 to 3.8)	0.6 (0.4 to 0.8)	Very Low
Pneumonia, urinary tract infection, bacteraemia, bacterial meningitis, cellulitis, septic arthritis, osteomyelitis, otitis media, bacterial gastroenteritis. Prevalence= 23.9%								
> 20 mg/l								
1 (Berger et al, 1996)	N = 138	83.3 (70.0-96.7	67.0 (57.7-76.4)	43.9 (31.0-56.7)	92.9 (86.8-98.9)	2.53 (1.82-3.50)	0.25 (0.11-0.56)	Moderate
Occult bacteremia, bacterial meningitis, UTI. Prevalence = 0.9%								
20 g/l								
1 (Gomez et al, 2010)	N = 1018	73.9 (53.5 to 87.5)	74.8 (72 to 77.5)	3 (1 to 5)	100 (99 to 100)	3.1 (2.1 to 4.5)	0.3 (0.1 to 1.0)	Very low
70 g/l								
1 (Gomez et al, 2010)	N = 1018	69.6 (49.1 to 89.4)	93.8 (92.1 to 95.1)	9 (2 to 15)	99.3 (98.5 to 99.6)	10.7 (6.3 to 18.0)	0.4 (0.1 to 0.9)	Very low
Occult bacteremia, UTI, Pneumonia. Prevalence = 14.3%								
3 mg/dl								
1 (Pratt et al, 2007)	N = 119	88 (62 to 98)	68 (58 to 76)	0.31 (19 to 46)	97 (89 to 100)	2.7 91.96 to 3.80)	0.17 (0.05 to 0.64)	Very Low
5 mg/dl								
1 (Pratt et al, 2007)	N = 119	71 (44 to 89)	84 (75 to 100)	43 (25 to 63)	94 (87 to 98)	4.5 (2.6 to 7.8)	0.35 (0.17 to 0.73)	Very Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
7 mg/dl								
1 (Pratt et al, 2007)	N = 119	59 (33 to 81)	87 (79 to 93)	43 (24 to 65)	93 (85 to 97)	4.6 (2.4 to 8.8)	0.47 (0.27 to 0.83)	Very Low
3 mg/dl < 12 hours								
1 (Pratt et al, 2007)	N = 45	67 (24 to 94)	74 (58 to 86)	28 (5 to 52)	94 (85 to 102)	2.6 (1 to 5.2)	0.4 (0.1 to 1.4)	Very Low
5 mg/dl								
1 (Pratt et al, 2007)	N = 45	50 (14 to 86)	92 (78 to 98)	50 (10 to 90)	92 (84 to 101)	6.5 (1.7 to 22.3)	0.5 (0.2 to 1.2)	Very Low
7 mg/dl								
1 (Pratt et al, 2007)	N = 45	33 (6 to 76)	97 (85 to 100)	67 (13 to 120)	90 (82 to 99)	13 (1.8 to 88.4)	0.7 (0.4 to 1.2)	Very Low
3 mg/dl > 12 hours								
1 (Pratt et al, 2007)	N = 74	100 (72 to 100)	63 (50 to 75)	32 (17 to 48)	100 (98 to 101)	2.7 (1.7 to 3.8)	0.0 (0.0 to 6.8)	Very Low
5mg/dl								
1 (Pratt et al, 2007)	N = 74	82 (48 to 97)	79 (67 to 88)	41 (20 to 61)	96 (91 to 101)	4 (2.1 to 6.9)	0.2 (0.1 to 0.8)	Very Low
7 mg/dl								
1 (Pratt et al, 2007)	N = 74	73 (40 to 93)	81 (69 to 89)	40 (19 to 61)	94 (88 to 101)	3.8 (1.9 to 7)	0.3 (0.1 to 0.9)	Very Low
Occult bacteremia, UTI, Pneumonia. Prevalence = 11.3%								
4.4 mg/dl								
1 (Isaacman et al, 2002)	N = 256	63 (43 to 82)	81 (76 to 87)	30 (18 to 43)	94 (91 to 98)	3.3 (2.2 to 4.8)	0.5 (0.3 to 0.7)	Low
Occult bacteremia, UTI, Pneumonia. Prevalence = 18%								
7 mg/dl								
1 (Pulliam et al, 2001)	N = 77	79 (49 to 94.2)	91 (79.8 to 96)	65 (38.3 to 85.8)	95 (86.1 to 99)	8.3 (3.8 to 27.3)	0.2 (0.1 to 0.6)	Low

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Invasive bacterial infections: Bacterial meningitis, occult bacteremia & sepis. Prevalence = 1.7%								
≥ 20 mg/L								
1 (Luaces-Cubells et al, 2012)	N = 868	0.80 (0.52 to 0.96)	0.66 (0.63 to 0.69)	0.04 (0.02 to 0.07) [a]	0.99 (0.98 to 1.00) [a]	2.36 (1.80 to 3.09) [a]	0.30 (0.11 to 0.83) [a]	Low
≥ 40 mg/L								
1 (Luaces-Cubells et al, 2012)	N = 868	0.47 (0.21 to 0.73)	0.83 (0.80 to 0.85)	0.05 (0.02 to 0.10) [a]	0.99 (0.98 to 0.99) [a]	2.72 (1.55 to 4.76) [a]	0.64 (0.40 to 1.03) [a]	Low
≥ 80 mg/L								
1 (Luaces-Cubells et al, 2012)	N = 868	0.33 (0.12 to 0.62)	0.95 (0.93 to 0.96)	0.10 (0.04 to 0.23) [a]	0.99 (0.98 to 0.99) [a]	6.45 (2.98 to 13.97) [a]	0.70 (0.49 to 1.01) [a]	Low
≥ 91 mg/L								
1 (Luaces-Cubells et al, 2012)	N = 868	0.33 (0.12 to 0.62)	0.96 (0.94 to 0.97)	0.13 (0.05 to 0.28) [a]	0.99 (0.98 to 0.99) [a]	8.16 (3.71 to 17.93) [a]	0.70 (0.49 to 0.99) [a]	Low

CI confidence interval, UTI urinary tract infection

[a] Estimates and confidence intervals were calculated by the NCC-WCH technical team.

[b] Confidence intervals were calculated by the NCC-WCH technical team.

Table 8.3 GRADE findings for combined procalcitonin (PCT) and C-reactive protein (CRP) tests

Number of studies	Number of children	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive predictive value (95% confidence interval)	Negative predictive value (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)	Quality
Bacteraemia, pyelonephritis, lobar pulmonary condensation. Prevalence = 22.6%								
PCT 0.9 ng/ml or CRP 40 mg/l								
1 (Lacour et al, 2001)	N = 124	96 (82 to 100)	67 (56 to 76)	46 (33 to 58)[a]	98 (95 to 101)[a]	2.9 (2.2 to 3.9)[b]	0.1 (0.0 to 0.4)[b]	Low
Bacteremia. Prevalence = 3.2%								
PCT ≥ 2ng/ml and/or CRP ≥ 40mg/l								
1 (Guen et al, 2007)	N = 215	71.4 ±0.33	61.4±0.07	6.5 ±0.37	98.2 ±0.06	1.85	0.46	Low

CRP C-reactive protein, PCT procalcitonin

[a] Estimates and confidence intervals were calculated by the NCC-WCH technical team.

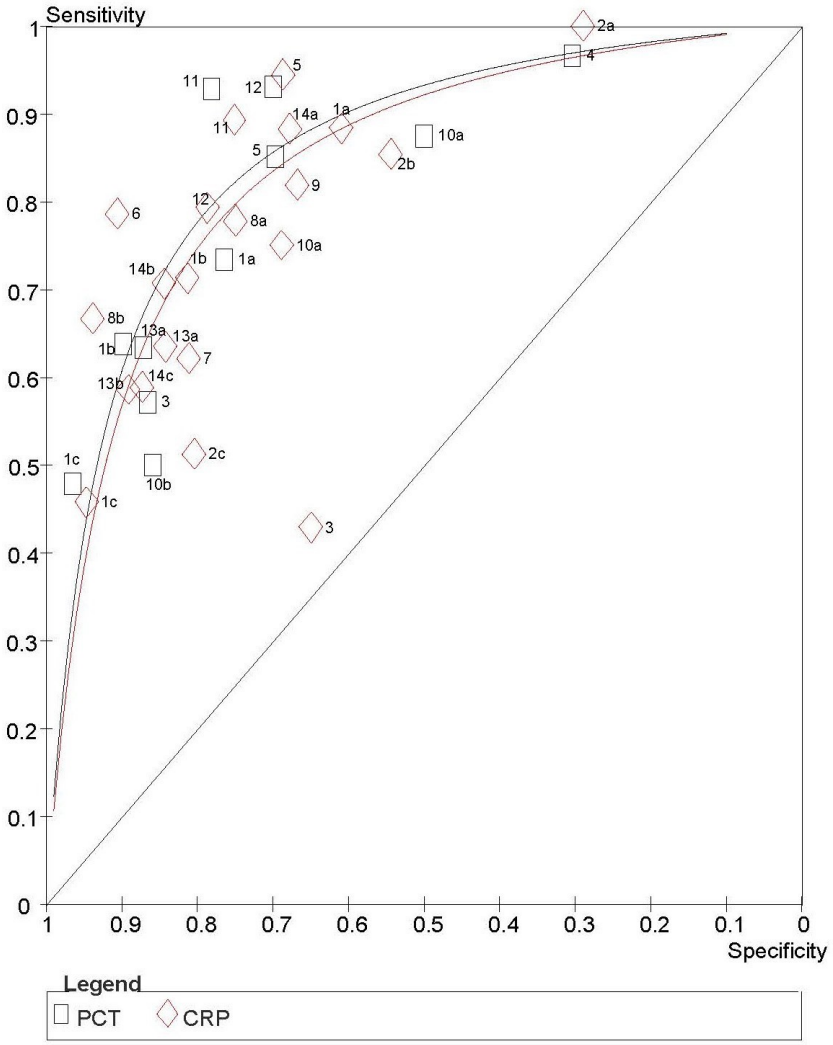
[b] Confidence intervals were calculated by the NCC-WCH technical team.

Comparison of PCT and CRP

The section below provides a comparison of the reported results for PCT and CRP.

A plot of all the sensitivity and specificity points reported in the included studies is shown in Figure 8.1. Based on these, the predicted receiver operating characteristic (ROC) curves have been generated for each test. However, this is only a descriptive plot and not a formal meta-analysis.

Figure 8.1 Plot of all reported sensitivities and specificities for PCT and CRP (1a-c Andreola 2007; 2a-c Hsiao 2006; 3 Guen; 4 Maniaci 2008; 5 Mazano; 6 Pulliam 2001; 7 Isaacman; 8 Gomez; 9 Berger 1996; 10a-b Thayyil 2005; 11 Lacour 2001; 12 Galetto-Lacour 2003; 13 Olaciregui 2009; 14 Pratt 2007; 15 - Luaces-Cubells, 2012 ; 16 - Woelker, 2012).



Next, a meta-analysis was undertaken for each test. In order to reduce bias only the eight studies that directly compared PCT and CRP were included and a single point estimate from each study. The results of these are shown in Table 8.4. These show that both tests predict a high area under curve (greater than 0.8) on the ROC curve, with a figure higher than 0.8 suggesting a useful test, and only small differences between the tests. The summary sensitivity results show both tests are moderately useful, although PCT performs slightly better. The summary specificity result shows both tests have moderate specificity, with PCT performing slightly better. However, the confidence intervals around these point estimates range from low to high predictive value, which suggests uncertainty in the findings. In addition, there was a high level of heterogeneity between studies in terms of how conditions were classified as serious or not and the setting where the study was undertaken, which limits the use of these results.

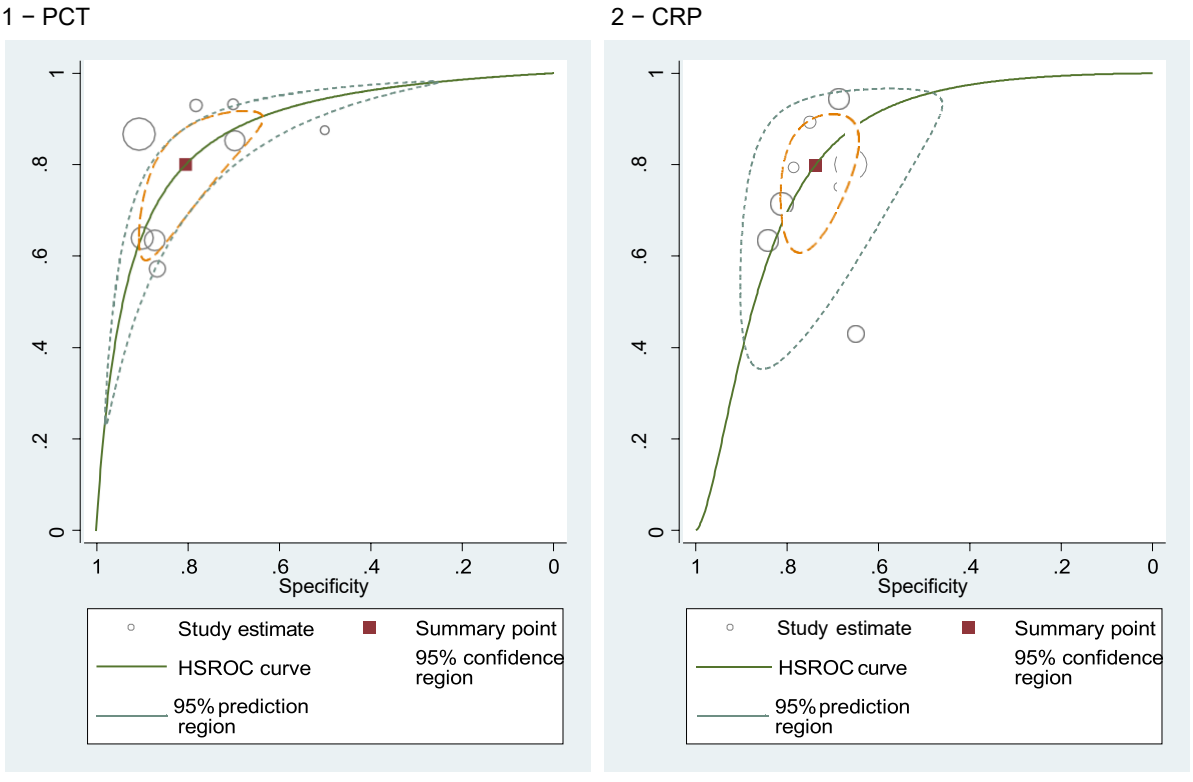
Table 8.4 Summary of results of meta-analysis of studies of PCT and CRP

	PCT result (95% CI)	CRP result (95% CI)
Number of studies	8	8
ROC area, AUROC	0.87 (0.84 to 0.90)	0.83 (0.79 to 0.86)
Sensitivity	0.80 (0.68 to 0.88)	0.80 (0.69 to 0.88)
Specificity	0.80 (0.71 to 0.87)	0.74 (0.68, 0.79)
Positive likelihood ratio	4.1 (2.9 to 5.7)	3.1 (2.5 to 3.7)
Negative likelihood ratio	0.25 (0.16 to 0.38)	0.27 (0.18 to 0.42)
Inconsistency (I-square)	96 (94 to 99)	92 (85 to 97)

AUROC area under the receiver operating characteristic curve, CI confidence interval, CRP C-reactive protein, PCT procalcitonin, ROC receiver operating characteristic

The results of the meta-analysis are further summarised in Figure 8.2 shown below. The hierarchical summary receiver operating characteristic (HSROC) is the summary ROC curve. Individual studies are shown as circles, which are proportion to the study sample size. The 95% prediction region is where the ROC curve could vary and the 95% confidence region is where the summary point estimate could vary.

Figure 8.2 Predicted ROC curve with confidence intervals for PCT and CRP



Evidence statements

The following definitions have been used when summarising the levels of sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV):

- High – 90% and above
- Moderate – 75% to 89%
- Low – 74% or below

C-reactive Protein

Fifteen studies of moderate to very low quality examining different combinations of bacterial illness were included in the review. The results showed that CRP has moderate sensitivity and specificity. A level below 20 mg/ml was needed to maximise sensitivity and a level above 80 mg/ml is needed to maximise specificity.

Procalcitonin

Ten studies of low to very low quality examining different combinations of bacterial illness were included in the review. The results showed that PCT has moderate sensitivity and specificity. A level below 0.5 ng/ml is needed to maximise sensitivity and a level above 2 ng/ml is needed to maximise specificity.

Comparison of procalcitonin with C-reactive protein

Eight studies of low to very low quality comparing PCT and CRP were included in the review.

One study of very low quality evidence had a population with 23% prevalence of bacteremia, pyeonephritis, pneumonia, bacterial meningitis, bone infection and sepsis. Procalcitonin showed low sensitivity and moderate specificity at 0.5 ng/ml; low sensitivity and high specificity at 1 ng/ml; and low sensitivity and high specificity at 2 ng/ml. C-reactive protein showed moderate sensitivity and low specificity at 20 mg/l; low sensitivity and moderate specificity at 40 mg/l; and low sensitivity and high specificity at 80 mg/l.

One study of low quality evidence had a population with 29% prevalence of bacteremia, pyelonephritis, pneumonia, mastoiditis and retropharyngeal abscess. Procalcitonin showed high sensitivity and low specificity at 0.5 ng/ml. C-reactive protein showed moderate sensitivity and moderate specificity at 40 mg/l.

One study of low quality evidence had a population with 23.6% prevalence of bacteremia, meningitis, sepsis, UTI, pneumonia, gastroenteritis and cellulitis. Procalcitonin showed low sensitivity and moderate specificity at a cut-off of 0.5 ng/ml or more. C-reactive protein showed low sensitivity and moderate specificity at a cut-off of 20 mg/l or more: it also showed low sensitivity and moderate specificity at a cut-off of 30 mg/l or more.

One study of low quality evidence had a population with 16% prevalence of bacteremia, UTI, pneumonia and bacterial meningitis. Procalcitonin showed moderate sensitivity and low specificity at a cut-off greater than 0.2 ng/ml. C-reactive protein showed high sensitivity and low specificity at a cut- off of 17.7 mg/l or more.

One study of very low quality evidence had a population with 11.1% prevalence of bacterial pneumonia, bacterial meningitis, septicaemia and pyelonephritis. Procalcitonin showed moderate sensitivity and low specificity with a cut-off greater than 500 ng/ml but it showed low sensitivity and moderate specificity at a cut-off greater than 2000 ng/ml. C-reactive protein showed moderate sensitivity and low specificity at a cut-off greater than 50 mg/l.

One study had a population with 0.6% prevalence of bacteremia and sepsis. Procalcitonin showed moderate sensitivity and high specificity at a cut-off greater than 0.5 ng/ml. Evidence for this finding was of low quality. C-reactive protein showed low sensitivity and low specificity at a cut-off greater than 30 mg/l. Evidence for this finding was of very low quality.

One study of low quality evidence had a population with 3.2% prevalence of bacteremia. Procalcitonin showed low sensitivity and moderate specificity at a cut-off of 2 ng/ml or more. C-reactive protein showed low sensitivity and low specificity at a cut-off of 40 mg/l or more.

One study of low quality had a population with 1.7% prevalence of bacterial meningitis, occult bacteremia and sepis. Procalcitonin showed moderate sensitivity and high specificity at a cut-off of 0.9 ng/ml or more. C-reactive protein showed moderate sensitivity and low specificity at a cut-off of 20 mg/l or more.

Combined procalcitonin with C-reactive protein

One study of low quality evidence had a population with 22.6% prevalence of bacteremia, pyelonephritis, lobar pulmonary consolidation. Combined PCT or CRP tests showed high sensitivity and low specificity at 0.9 ng/ml and 40 mg/l respectively.

One study of low quality evidence had a population with 3.2% prevalence of bacteremia. Combined PCT and CRP tests showed low sensitivity and low specificity at a cut-off of 2 ng/ml or more and 40 mg/l or more respectively.

Health economic evidence statements

No new health economic studies were identified and no significant changes to costs were identified. Therefore, no health new economic evaluation was undertaken for this question (see Evidence to Recommendations below for the GDG's view of why an additional analsyis was not required).

An economic evaluation was undertaken in the previous guideline to assess the cost effectiveness of using CRP versus using PCT to investigate the presence of SBI in children without apparent source (see Appendix D). Health economic evaluation was required since PCT is not routinely used. All other diagnostic tests are offered on the NHS and are part of the usual package of tests for children over 3 months where SBI is suspected. The results indicated that under certain assumptions CRP is both less costly and more effective than PCT in correctly diagnosing and ruling out SBI in children with fever without apparent source (FWS). However, the results were sensitive to the prevalence of SBI. CRP no longer dominated PCT when the prevalence of SBI was over 27%, keeping all the other baseline assumptions constant. Nevertheless, given the lack of robust evidence underpinning these baseline assumptions, the analysis cannot support the replacement of CRP with PCT at present. The GDG has recommended more research on the performance characteristics of CRP and PCT in children with feverish illness of uncertain cause.

Evidence to recommendations

Relative value placed on the outcomes considered

The GDG stated that the overarching aim of the guideline was the early and accurate detection of serious illness in children with fever. This allows for suitable treatment to begin, which will then reduce mortality and morbidity. Diagnostic tests are part of this process.

Consideration of clinical benefits and harms

The GDG members stated that, to their knowledge, the evidence presented was accurate and complete.

The GDG highlighted that the new data showed that both CRP and PCT were moderately useful diagnostic tests. The GDG members noted that the data comparing CRP and PCT showed a statistical difference in favour of PCT. However, they were also aware of the small absolute difference, low quality of the data and heterogeneity between the studies in terms of settings and populations. Furthermore, the GDG highlighted that while CRP was routinely available in secondary care within the NHS, no one was aware of PCT being used outside a research setting in the NHS for children. CRP and PCT are rarely available in primary care and any child who was unwell enough to require a CRP or PCT test should be immediately referred to a paediatric specialist.

The GDG discussed if there were any situations in which PCT would be more beneficial than CRP. The main focus of this was early detection of bacterial illness, as PCT levels increase earlier in response to infection than CRP. However, the evidence suggested that few children are taken to an emergency department within 6 hours of a fever starting.

Based on its assessment of the data, the GDG concluded that no clinically important difference between PCT and CRP could be identified, and therefore that CRP should still be recommended and PCT should not. Furthermore, the GDG decided that no change was needed concerning when a CRP test should be ordered.

The GDG examined the use of CRP and PCT test results together. The GDG concluded that the evidence showed that there was little additional benefit from using the tests together compared with usnig each on its own and there was no clinical reason for doing so.

The GDG debated if specific cut-offs for CRP could be recommended. The evidence suggested a level above 80 mg/ml would maximise specificity and a level lower than 20 mg/ml would maximise sensitivity. However, the GDG highlighted that there was known variation between laboratories. Furthermore, the GDG emphasised that the test results should not be used in isolation to decide clinical action, but should be used in conjunction with other results and a clinical assessment. Therefore, the GDG decided not to make recommendations on specific cut-offs for CRP.

Consideration of health benefits and resource uses

It was highlighted that CRP costs approximately £3 per test compared to £25 per test for PCT. In addition, CRP is currently available across England, whereas PCT is only used in a few research settings and more widespread use would require substantial training. The GDG concluded that given there was no clear clinical advantage to using PCT compared with CRP in children presenting with fever then it was not cost effective.

Quality of evidence

Evidence was of moderate to very low quality. There were a number of common issues which influenced the quality of evidence including: differing study populations; lack of blinding; not all subjects receiving reference tests; and imprecision of results caused by small sample sizes. Furthermore, there was heterogeneity between studies in terms of the settings where tests were undertaken and how conditions were classified as serious or non-serious.

The GDG noted that the study by Guen (2007) appeared to be an outlier. The possible reasons for this were discussed; these included the fact that occult bacteremia was being investigated and that tests were carried out within 3 hours of presentation on the children who were found to have SBI. However, the GDG concluded that even if this study was excluded it would not change its recommendations.

Other considerations

Equalities

No equality issues were raised in relation to this question.

Health economics

An economic evaluation was undertaken to assess the cost effectiveness of using CRP versus using PCT to investigate the presence of SBI in children without apparent source (see Appendix D). Health economic evaluation was required since PCT is not routinely used. All other diagnostic tests are offered on the NHS and are part of the usual package of tests for children over 3 months where SBI is suspected. The results indicated that under certain assumptions CRP is both less costly and more effective than PCT in correctly diagnosing and ruling out SBI in children with FWS. However, the results were sensitive to the prevalence of SBI. CRP no longer dominated PCT when the prevalence of SBI was over 27%, keeping all the other baseline assumptions constant. Nevertheless, given the lack of robust evidence underpinning these baseline assumptions, the analysis cannot support the replacement of CRP with PCT at present. The GDG has recommended more research on the performance characteristics of CRP and PCT in children with feverish illness of uncertain cause.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Research recommendations

Number	Research recommendations
	Diagnosis
RR 5	The GDG recommends that a UK study of the performance characteristics and cost-effectiveness of procalcitonin versus C-reactive protein in identifying serious bacterial infection in children with fever without apparent source be carried out. [2007].

Why this is important

Many young children with fever appear well with no symptoms or signs of serious illness. The vast majority of these children will have self-limiting illnesses. However, a few will have serious bacterial infections which may not be identifiable by clinical assessment alone. Investigations that help to identify these children with serious bacterial infections could lead to prompt antibiotic treatment, which may improve their outcome. These investigations need to be both sensitive and specific so that most serious bacterial infections are identified and so that antibiotics are not given to children who don't need them. The inflammatory markers C-reactive protein and procalcitonin have shown varying performance characteristics for identifying bacterial infection in a variety of populations. If either or both were found to be sensitive and specific for identifying serious bacterial infection in children with fever without apparent source, there would be evidence for their more widespread use. The cost effectiveness of this approach would need to be calculated.

Viral co-infection

Only a minority of young children with fever have bacterial infections. The rest are presumed to have viral infections, although these are rarely confirmed and mostly do not need treatment. If it were possible to identify those children with definite viral infections, this might help identify those at low risk of serious illness. However, if bacterial infection co-existed with viral infection then differentiating between serious and non-serious illness would not be helped by identifying those with viral infection.

Review question

What is the incidence of co-existing bacterial infection in a child presenting with fever in which a virus (e.g. influenza or RSV) is detected (with a rapid test)?

Narrative evidence

Three EL 3 retrospective studies[181–183] which investigated co-existing bacterial infection in children with respiratory syncytial virus (RSV) infection were found. One retrospective cohort[181] investigated the prevalence of co-existing SBI in 178 children less than 8 weeks old with proven RSV infection and fever. Those children with RSV were over five times more likely to have an increased work of breathing compared with those who were RSV negative (RR 5.1, 95% confidence interval [CI] 2.9 to 8.9). The other two retrospective cross-sectional studies investigated children with influenza virus[182] and RSV respiratory tract infection.[183] The odds of any SBI were 72% less in children who tested positive for influenza than in those who did not (odds ratio [OR] 0.28, 95% CI 0.16 to 0.48).[182] Febrile RSV-positive infants had a lower rate of bacteraemia compared with febrile RSV-negative infants (1.1% versus 2.3%). Similarly, none of the febrile children with RSV respiratory tract infection tested had positive cerebrospinal cultures, but urinary tract infection was found in 14% of those less than 3 months old and 8.4% of those over 3 months old.[183]

Evidence summary

The incidence of SBI is lower in feverish children with proven RSV or influenza infections compared with those in whom viral investigations are negative. However, SBI, especially UTI and influenza/RSV, infections can co-exist.

GDG translation

Since children with proven viral infection still have a risk of SBI (although this was reduced compared with children without proven viral infection), the GDG felt that they should be assessed for serious illness in the same way as other children. Those with no features of serious illness should have urine tested, while those with features of serious illness should be assessed by a paediatric specialist. Given that rapid detection of viral illness (such as influenza or RSV infection) does not exclude a co- existing SBI, the GDG recognised that the use of these tests is not an efficient use of scarce healthcare resources.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Observation in hospital

Children with fever are often observed in hospital for a period of time to help differentiate those with serious illness from those with non-serious illness. This observation usually involves the repeated measurement of 'vital signs' such as heart rate, respiratory rate and temperature, as well as repeated assessments of the child to look for the development of any clinical features that would give cause for concern. Investigations, if indicated, can also be done and their results sometimes obtained during a period of observation.

Review question

In a child with fever what are the benefits, if any, of a period of observation on an assessment facility?

GDG statement

The GDG found limited research to show the overall benefits of a period of observation in the paediatric assessment unit of the child with fever, in terms of cases of serious illness identified, hospital admission, morbidity, mortality and recovery. Delphi consensus (see section 3.2) was sought in an attempt to answer the question as to whether or not observation itself can help to differentiate feverish children with non-serious and serious illness. In addition, the Delphi panel were asked to decide how long such a period of observation should be.

Delphi statement 5.1

A period of observation in hospital (with or without investigations) as part of an assessment can help differentiate minor from serious bacterial illness (such as bacterial meningitis or pneumonia) in a young child who has a fever without obvious cause.

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
0	6 (12%)	44 (85%)	2 (4%)		52	8

Delphi statement 5.2

The period of observation in a hospital to help differentiate minor from serious illness in a young child over 3 months of age with fever without obvious cause should be approximately:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	3 (6%)	26 (50%)	10 (19%)	12 (23%)	52	6

There was 85% agreement (consensus achieved) for Statement 5.1 but no consensus reached for Statement 5.2.

GDG translation

The GDG accepted that Delphi consensus agreeing that a period of observation of young children with fever in hospital was useful in differentiating those with minor illness from those with serious illness. The GDG believes that this period of observation is likely to be cost-effective for the NHS since the cost of observation is outweighed by savings from preventing unnecessary diagnostic tests from being undertaken in children with minor illness. The GDG acknowledged that no evidence was found nor consensus reached to determine the ideal duration of such a period of observation. Since febrile infants less than 3 months of age have an increased risk of SBI which can be missed by observation alone, the guideline does not suggest observation alone in this age group.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Response to antipyretic medication

This section was partially updated in 2013.

It has been suggested that response to antipyretic medication may help differentiate serious from non-serious illness in febrile children. This could occur in two ways:

- a difference in the response to antipyretics being reflective of the seriousness of the underlying condition

- improved clinical appearance.

Decrease in fever after antipyretics

Some healthcare professionals think that a decrease in fever with antipyretic therapy indicates a lower likelihood of serious bacterial infection (SBI), and that a lack of response makes an SBI more likely. In contrast to this, other healthcare professionals fear that giving antipyretics to reduce fever in febrile children may make the detection of serious illness more difficult as the high fever and other symptoms of SBI is 'masked' by antipyretics. Evidence about fever response to antipyretics in children with both serious and non-serious illness would be useful to help in the assessment of these children.

Improved clinical appearance after antipyretics

Antipyretics may also improve the child's general condition. Many healthcare professionals feel that clinical review of a febrile child 1–2 hours after they have been given antipyretics improves the ability to differentiate between serious and non-serious illness. The antipyretic and analgesic effect of antipyretics may lead to the improvement of features which may suggest serious illness (e.g. irritability, tachycardia, etc.). If this improvement in features occurred only in those with non-serious illness, this would help to identify these children. However, if this improvement also occurred in children with serious illness, then these children may not have their illness identified correctly.

Evidence about improved clinical appearance after antipyretics would be useful to help in the assessment of children and would also be relevant to the use of observation in febrile children.

Updated review for 2013

Review question

The clinical question outlined in the scope asks for 'The predictive value of the clinical response to paracetamol or NSAIDs' (non-steroidal anti-inflammatory drugs). This translated to a review question of: "What is the predictive value of the clinical response to paracetamol or NSAIDs?"

Description of included studies

Eight studies were identified (Torrey et al, 1984; Baker et al, 1989; Yamamoto et al, 1987; Mazur, 1989 et al; Weisse et al, 1987; Baker et al, 1987; Mazur et al, 1994; Bonadio et al, 1993), including four used in the 2007 guideline.

Five of the studies were prospective (Torrey et al., 1984; Baker et al., 1989; Yamamoto et al., 1987; Weisse et al., 1987; Baker et al., 1987) and three were retrospective (Mazur et al., 1989; Mazur et al., 1994; Bonadio et al., 1993). Two used the same dataset (Mazur et al., 1989; Mazur et al., 1994). All of the studies were undertaken in hospital settings in the USA, and the most recent was undertaken in 1994. All studies focused on differentiating bacterial illness from other illnesses. All studies reported on change in temperature, four reported on change in febrile state and one reported on change in symptomology. Studies used different definitions of fever, disease and timing of follow-up. The majority of studies used paracetamol (acetaminophen), with aspirin being used when paracetamol had already been administered (this reflects the age and setting of the studies as this is not acceptable practice in the UK). The dosage of paracetamol used was either 10 or 15 mg/kg. The age of children ranged up to 17 years in one study, but the majority were 2 years or younger.

Evidence profile

The GRADE profiles presented show results of included studies for the review question. .

Table 8.5 GRADE findings of response to antipyretics by children with bacterial or non-bacterial illnesses.

Number of studies	Number of children		Effect		Quality
	Serious disease (Δ ℃ [SD], n)	Not serious disease (Δ ℃ [SD], n)	Relative (95% confidence interval) (MD and Standardised MD) (95% confidence interval)	Absolute mean difference	
Final symptoms score – Yale Observation Score					
Baker et al, 1989	7.5, (+/- 1.4), n=15	7.7, (+/- 2.2), n=135	0.2 (NS)	-	Very low

Change in symptoms – Yale Observation Score					
Baker et al, 1989	-3.8 (+/- 3.2), n=15	-1.6 (+/- 2.5), n=135	2.2 (*P* < 0.001)	-	Very low
Change in temperature °C between serious and non-serious disease					
Torrey et al, 1984	-1.32, - , n=16	-1.05, - , n=239	0.27 (*P* = 0.14)	-	Very low
Baker et al, 1989	-1.7, (+/- 0.8), n=15	-1.6, (+/- 0.6), n=135	SMD -0.16 (-0.69 to +0.37)	-	Very low
Yamamoto et al, 1987	-1.606 (+/- 0.722), n=17	-1.639 (+/- 0.705), N=216	SMD 0.05 (-0.45 to +0.54)	-	Very low
Mazur et al, 1989	-1.0 (+/- 0.6), N=34	-1.5 (+/- 0.5), N=68	SMD 0.92 (0.49 to 1.36)	-	Very low
Weisse et al, 1987	1.48°F, -, n=17	1.16°F, -, 1n=6	0.32°F (*P* = 0.37)	-	Very low
Baker et al, 1987	1.3 (+/- 0.8), n= 62	1.0 (+/- 0.6), n= 234	SMD -0.46 (-0.75 to -0.18) *P* < 0.01 against all groups	-	Very low
Mazur et al, 1994	-1.0, (+/- 0.6, n=34	-1.2, (+/- 0.6), n=450	SMD -0.33 (-0.68 to +0.02)	-	Very low
Bonadio et al, 1993	-1.40, -, n=59	-1.44, -, n=59	0.04 (NS)	-	Very low
Final temperature C between serious and non-serious disease					

Number of studies	**Number of children**		**Effect**		**Quality**
	Serious disease (Δ ℃ [SD], n)	**Not serious disease (Δ ℃ [SD], n)**	**Relative (95% confidence interval) (MD and Standardised MD) (95% confidence interval)**	**Absolute mean difference**	
Torrey et al, 1984	38.8, - , n=16	38.8, - , n=239	(*P* = 0.46)	-	Very low
Baker et al, 1989	38.5 (SD +/- 0.6), n=15	38.4 (SD +/- 0.6), n=135	(NS) SMD 0.17 (-0.37 to +0.70)	-	Very low
Change in febrile state					
Yamamoto et al, 1987	15 of 17	180 of 216	RR 1.06 (0.88 to 1.27)	-	Very low
Mazur et al, 1989	18 of 34	62 of 68	Univariate OR = 9.2 (95% CI 2.7 to 32.0) Multivariate OR = 9.4 (95% CI 2.6 to 34.2)	-	Very low
Weisse et al, 1987	4 of 35	10 of 65	RR 0.74 (0.25 to 2.20)	-	Very low

Mazur et al, 1994	18 of 34	335 of 450	RR 0.71 (0.52, 0.98) Univariate OR = 2.6 (95% CI 1.3 to 5.2) Multivariate OR = 3.4 (95% CI 1.6 to 7.3)	-	Very low

CI confidence interval, MD mean difference, NS non-significant, OR odds ratio, RR relative risk, SD standard deviation, SMD standard mean difference

Evidence statements

One study found no difference in the Yale Observation Score between children with bacterial illness and those without bacterial illness following treatment with antipyretics. However, the same study did find a significant difference in the change in Yale Observation Score before and after treatment between the groups. The evidence for this finding was of very low quality.

Three studies found that temperature was reduced more in children with bacterial illness compared with children without bacterial illness after antipyretics were administered. Three other studies found no difference in temperature reduction in children with bacterial illness compared with children without bacterial illness after antipyretics (paracetamol or ibuprofen) were administered. A further two studies that analysed the same dataset found temperature was reduced less in children with bacterial illness compared with children without bacterial illness after antipyretics were administered. The quality of evidence was very low.

Two separate studies that analysed the same dataset found that the proportion of children who responded to antipyretics (paracetamol or ibuprofen) was lower in children with bacterial illness compared with children without bacterial illness. Two other studies found no difference in the proportion of children who responded to antipyretics. This evidence was very low quality.

Health economics profile

No health economic studies were identified and no health economic analysis was undertaken for this question as it did not consider the effectiveness of alternative interventions.

Evidence to recommendations

Relative value placed on the outcomes considered

The GDG stated that the overarching aim of the guideline was the early and accurate detection of serious illness in children with fever. This allows for suitable treatment to begin, which will then reduce morbidity and mortality.

Consideration of clinical benefits and harms

The GDG members stated that, to their knowledge, all the relevant available evidence had been reviewed.

The GDG believed that some healthcare professionals think that a faster or greater decrease in temperature after antipyretics would suggest that a serious illness is less likely. The GDG concluded that this is not supported by evidence.

The GDG found evidence from one study showing that if a child's Yale Observation Score was measured before and after the use of antipyretics, the clinical features may have resolved in those without serious illness. As the traffic light system contains many of the same features as the YOS, the GDG believed that reassessment after antipyretics may help differentiate those with and without serious illness. However, the GDG concluded that more research should be undertaken on this before any recommendation could be made.

The GDG considered the possibility that antipyretics, by reducing symptoms, might 'mask' the severity of a serious illness. The GDG concluded that there was insufficient evidence to make a conclusion on this matter. The GDG stated that this question should be a priority for future research.

However, the GDG was concerned that the recommendation not to use response to antipyretics as a diagnostic test could lead to children not being regularly reassessed, as often this was done to see if a child had responded to antipyretics. The GDG was worried that this could lead to unnecessary delays in treatment for children with serious illnesses. Therefore, the GDG added a caveat to the recommendation stating that a child admitted to hospital with any amber or red features on the traffic light table needed be regularly assessed to ensure their condition had not worsened.

Consideration of health benefits and resource uses

The GDG emphasised that antipyretics were considerably cheaper than any formal diagnostic test. However, as the GDG concluded that antipyretics were of no diagnostic value, switching to antipyretics from diagnostic tests would not be a cost-effective option for the NHS.

Quality of evidence

The available evidence was of very low quality due to poor study design. In addition, heterogeneity between studies in terms of definitions of fever, disease, dosage of antipyretics, age of children and timing of follow-up made comparison of outcomes difficult.

The studies were also relatively old and in many cases did not reflect what would be considered safe practice in the UK, especially in relation to giving Aspirin to children as this would not be used in the UK.

Other considerations

No equalities issue were identified in relation to this question.

Recommendation

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Research recommendation

Number	Research recommendations
	Antipyretics
RR 6	The GDG recommends that studies are conducted in primary care and secondary care to determine whether examination or re-examination after a dose of antipyretic medication is of benefit in differentiating children with serious illness from those with other conditions. [2007] Why this is important Antipyretic medications are widely used in primary and secondary settings by parents and healthcare professionals. Children may therefore present to healthcare facilities having had a dose of antipyretics. Furthermore, the child's response to antipyretic drugs may be used as an indication of severity of illness, the rationale being that those with milder illness will either show greater improvement in condition or a greater reduction in their fever than children with more serious illnesses. However, it is not clear if such changes in condition are a valid and reliable method of differentiating children with serious illness from those with less serious conditions.

8.4 Immediate treatment by the paediatric specialist

Some children with fever have life-threatening serious illness which requires immediate treatment to improve their chances of survival. These treatments will be:

- directed against the causative organism (antibiotics, aciclovir)
- directed against the consequences of the infection, such as shock or respiratory failure (intravenous fluids, oxygen)
- directed against the inflammation caused by the infection (corticosteroids).

Many of these immediate treatments are endorsed in paediatric advanced life support courses and are therefore commonly used in the UK. Specific guidance for the immediate treatment of suspected meningococcal disease was also considered.

Review question

For children with symptoms and signs of a serious illness what immediate treatments improve their outcome?

Evidence of the effect of the following interventions in the treatment of serious illness was looked for:

- intravenous fluids
- steroids
- antibiotics
- aciclovir
- oxygen.

Intravenous fluids

Narrative evidence

Two systematic reviews (SRs) and three randomised controlled trials (RCTs) which looked at the use of intravenous fluids as immediate treatments were identified.

The first EL 1++ SR[190] evaluated three RCTs investigating the effect of maintenance fluid volumes in meningitis. Maintenance fluid was calculated as 100 ml/kg per day given for the first 10 kg body weight of the child, 50 ml/kg for the second 10 kg, and 20 ml/kg for over 20 kg. This was given intravenously for the first 48 hours for all three studies. The maintenance fluid volumes were compared with restricted fluid volumes 60% of the initial maintenance fluids. All three studies investigated both children and adults with acute bacterial meningitis. Pooling of the results of all three trials showed no significant difference between deaths in the maintenance and restricted fluid groups (RR 0.82, 95% CI 0.53 to 1.27). However, the risk of long-term neurological sequelae (spasticity, hemiparesis/hemiplegia, visual impairment and response to sound) was found to be significantly lower in the maintenance fluid group compared with the restricted fluid group (RR 0.42, 95% CI 0.20 to 0.89).

The second EL 1+ SR[191] involving 30 RCTs quantified the effect on mortality of administering either human albumin or plasma protein fraction during the management of 1419 critically ill patients. All patients were reported to have been critically ill as a result of hypovolaemia (state of decrease in the volume of blood plasma, which is characteristic of shock) due to trauma, surgery, burns or hypoalbuminaemia. The risk of death was 1.68 times more in the albumin group compared with the plasma protein group when the results of all the trials were summarised and pooled together (RR 1.68, 95% CI 1.26 to 2.23).

Three studies of which one was an EL 1++[192] study and two EL 1+ studies[50,193] were also found. The first RCT[192] EL 1++ compared the effect of fluid resuscitation with albumin or saline on mortality in both children and adults in the intensive care unit (n = 6997). There was no significant difference in the risk of death in the albumin group compared with the saline group (P = 0.87). At 28 days, there was still no difference in either group in the number of participants that remained in the ICU or hospital (P = 0.09 and 0.10, respectively). These researchers concluded that there was no appreciable difference in the survival times of either group.

The second RCT[50] evaluated the efficacy of normal saline and colloid (polymer from degraded gelatine in saline [Haemaccel]) intravenous fluid in restoration of circulating volume in children aged 0–12 years with septic shock. The median volume of fluid needed for initial resuscitation was significantly higher in the saline group compared with the gelatine group: 50 ml/kg (range 20–108) versus 30 ml (range 20–70) (P = 0.018). However, there was no difference in the time taken for resuscitation between the groups (P = 0.41).

The third RCT[193] determined whether moderate oral fluid restriction (nasogastric tube at 60% of normal maintenance volumes) or intravenous fluid (half-normal saline + 5% dextrose at 100% of normal maintenance volumes at full maintenance volumes) would result in a better outcome, for 346 children with bacterial meningitis, for the first 48 hours of treatment. There was no appreciable reduction in the risk of death or neurological sequelae in either group (P = 0.11).[193]

A fourth EL 2+ case–control study[11] investigated 143 children under 17 years who died from meningococcal diseases matched by age with 355 survivors from the same region of the country. The aim of the study was to determine whether suboptimal management in hospital contributed to poor outcome in children admitted with meningococcal disease. Inadequacies in fluid therapy in terms of too little versus adequate fluid therapy (OR 2.5, 95% CI 1.4 to 4.7, P < 0.004) and inadequate inotropes (OR 5.8, 95% CI 2.3 to14, P < 0.001) were significantly associated with death.

A further retrospective cohort study of children who presented to local hospitals with septic shock reviewed shock reversal (defined by return of normal systolic blood pressure and capillary refill time) and outcome. Shock reversal was successfully achieved in 24 (26%) children, which was associated with 96% survival and a nine-fold increased odds of survival (OR 9.49, 95% CI 1.07 to 3.89). Shock reversal was achieved by both fluid boluses and the early use of inotropes.[194]

Evidence summary

Many of the papers in the evidence table referred to maintenance intravenous therapy for bacterial meningitis, a subject that is outside the scope of this guideline. The GDG decided to address only studies that dealt with intravenous fluids for immediate resuscitation. Resuscitation with intravenous fluids in children with fever and signs of circulatory insufficiency is associated with lower mortality. Failure to administer sufficient intravenous fluids in children with meningococcal disease and septic shock is associated with higher risk of mortality. There is insufficient evidence to recommend colloid over crystalloid fluid and vice versa.

Health economics

The GDG recognises that there is a substantial cost difference, with crystalloids being considerably cheaper than colloids.

GDG translation

The GDG concluded that children with fever and signs of circulatory insufficiency have reduced mortality when given intravenous fluid resuscitation. Current practice would be to give a bolus of 20 ml/kg. The GDG recognises that there is unresolved debate about the relative merits of crystalloid and colloid fluids for this purpose. There remain concerns about the risks of infection from blood products, such as albumin. From a health economics perspective the GDG would favour the use of crystalloids. The GDG was aware that there is particular debate about the relative merits of albumin and crystalloid in the initial treatment of meningococcal disease, but making a recommendation on this issue was considered beyond the scope of this guideline.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Steroids

Narrative evidence

One EL 1+ SR[195] which looked at 18 RCTs investigating the effect of adjuvant corticosteroids on mortality, severe hearing loss and neurological sequelae, in the treatment of children and adults with acute bacterial meningitis was found. Overall, the number of participants who died was significantly smaller in the corticosteroid group compared with the placebo group: 8.5% versus 11.6% (RR 0.76, 95% CI 0.59 to 0.97). However, this effect on mortality was not seen in the subgroup of children (RR 0.95, 95% CI 0.65 to 1.37).

The administration of corticosteroids before or with the first dose of antibiotics was associated with a decreased risk of hearing loss. This was also evident for children with *Haemophilusinfluenzae* type b meningitis (RR 0.31, 95% CI 0.15 to 0.62) and for those with pathogens other than *Haemophilusinfluenzae* (RR 0.42, 95% CI 0.20 to 0.89).

Evidence summary

For children with bacterial meningitis the early use of steroids may decrease hearing loss. However, this was most evident for children with *Haemophilusinfluenzae* type b and possibly pneumococcal meningitis.

GDG translation

The GDG found no evidence to support the use of steroids other than in the early treatment of bacterial meningitis, which falls outside the scope of this guideline. The GDG noted the effect of steroids reported in the systematic review, but was unsure about the applicability in the UK, especially in the era of *Haemophilusinfluenzae* type b and pneumococcal vaccines. The GDG was unable to make a recommendation.

Antibiotics

Narrative evidence

One EL 2- cohort study[196] which evaluated the effect of empirical antibiotics on the outcome of SBI was found.

The prospective cohort study of critically ill adults[196] studied the relationship between inadequate antimicrobial treatment of infections (community-acquired and hospital-acquired) and hospital mortality for patients requiring ICU admission. The mortality rate of infected patients receiving inadequate antimicrobial treatment (52%) was significantly greater than the hospital mortality rate of patients without this risk factor (12%) (RR 4.26, 95% CI 3.52 to 5.15, $P < 0.001$).

Evidence summary

Critically ill children with SBI who are given no or ineffective antibiotics have an increased risk of mortality.

GDG translation

A diagnosis of SBI (especially bacteraemia) may not be confirmed until 12–36 hours from time of culture, since it takes this period of time to grow bacteria. Antibiotic treatment should not be delayed in a critically ill child until bacterial illness is confirmed, since the child may die during this period. Empirical antibiotic treatment should be given to critically ill children, at the earliest opportunity once SBI is suspected.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Aciclovir

Narrative evidence

Three EL 1- RCTs[197–199] looking at the treatment of serious illness with aciclovir were identified. Two of the RCTs[197,198] compared vidarabine and aciclovir as treatment in adults and children with herpes simplex encephalitis. The study which examined 208 adults reported more deaths (54% versus 28%, *P* = 0.008) and increased mortality (38% versus 14%, *P* = 0.021) in the vidarabine recipients than in the aciclovir recipients.[197] The study which looked at 210 infants less than 1 month old found no difference between vidarabine and aciclovir in either morbidity (*P* = 0.83) or mortality (*P* = 0.27).[198]

The third open-label RCT[199] estimated the treatment efficiency of high-dose aciclovir (HD, 60 mg/kg per day), intermediate dose (ID, 45 mg/kg per day) and standard dose (SD, 30 mg/kg per day) with regard to mortality and morbidity in 88 infants less than 28 days old. The survival rate for neonatal herpex simplex virus infection was found to be 3.3 times higher in those children treated with HD (OR 3.3, 95% CI 1.5 to 7.3). In addition, the children treated with HD aciclovir were 6.6 times more likely to be developmentally normal at 12 months of age, compared with children treated with standard dose therapy.

A large EL 3 retrospective multicentre study[200] studied prognostic factors for herpes simplex encephalitis in adult patients. A delay of greater than 2 days between admission to the hospital and initiation of aciclovir therapy was strongly associated with a poor outcome (OR 3.1, 95% CI 1.1 to 9.1, *P* = 0.037). However, there was still a favourable outcome for 55 of the patients (65%).

Evidence summary

Treatment with aciclovir decreases morbidity and mortality in adults and children with herpes simplex encephalitis. Treatment with aciclovir within 48 hours of admission improves the outcome in herpes simplex encephalitis.

GDG translation

The GDG recognised the difficulty in the early identification and treatment of children with herpes simplex encephalitis as the early features may be non-specific. The diagnosis of herpes simplex encephalitis may not be confirmed for a number of days after admission as initial investigations can be normal. Early treatment with aciclovir improves outcome in herpes simplex encephalitis.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Oxygen

Evidence summary

There was a lack of evidence meeting the inclusion criteria examining the effect upon outcome of administering oxygen to the child with symptoms and signs of serious illness.

GDG translation

Recommendations regarding treatment with oxygen were made based on GDG consensus.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

8.5 Causes and incidence of serious bacterial infection

Antimicrobial therapy has significantly improved the outcome for children with SBI. The appropriate antibiotic treatment for SBI will often not be determined for 24–36 hours, since it takes this period of time to grow bacteria and determine their antibiotic sensitivities. However, antibiotic treatment should not be withheld until the causative organism and its antibiotic sensitivities are confirmed, since the child may die or suffer harm in the meantime. Empirical antibiotic treatment is therefore given to children likely to have serious illness. Knowledge of the common organisms causing SBI in children will help decide which antibiotics should be used as empirical treatment for children likely to have SBI.

Review questions

What are the most common organisms causing serious illness in young children with fever?

What is the incidence of serious illness in young children with fever?

Narrative evidence

A search for UK-based cohort studies after 1992 found four EL 2+ retrospective studies.[121,201–203] The studies varied in baseline characteristics. For example, one study[121] recruited children aged 8 days to 16 years and another had children of 2 weeks to 4.8 years.[202] Moreover, some studies[201] recruited based on the presenting features of infectious disease or meningococcal disease[121] while others recruited children with a diagnosis of pneumonia[202] or bacterial meningitis.[203]

Hospital Episode Statistics (HES) was also reviewed as a proxy of incidence of serious illness in England and Wales. The data suggested that UTI (217.2/100,000), pneumonia (111.9/100,000), bacteraemia (105.3/100,000) and meningitis (23.8/100,000) were the most likely infections in children aged 7 days to 5 years admitted to hospital in England and Wales.[204]

Moreover, the likely organisms to cause these infections are *Neisseria meningitidis*, *Streptococcus pneumoniae*, *Escherichia coli*, *Staphylococcus aureus* and *Haemophilus influenzae* type b. In children less than 3 months of age, group B streptococcus and listeria may also cause SBI.[203]

Evidence summary

Serious bacterial infection in a child presenting to hospital with fever but without an identified focus is likely to be bacteraemia, meningitis, UTI or pneumonia. The likely organisms to cause these infections are *Neisseria meningitidis*, *Streptococcus pneumoniae*, *Escherichia coli*, *Staphylococcus aureus* and *Haemophilus influenzae* type b (rare in immunised children). In children less than 3 months of age, group B streptococcus and listeria may also cause SBI.

GDG translation

The GDG noted the causes of SBI and the likely organisms at various ages. The GDG believes that this information could be used to decide which antibiotics could be used when it is decided to treat a suspected SBI without apparent source and in the absence of the results of microbiological cultures. A third-generation cephalosporin (e.g. cefotaxime or ceftriaxone) might not be the treatment of choice for all these organisms but was felt to be adequate initial treatment. This empirical antibiotic treatment could be altered once culture results became available or the focus of infection became apparent.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

8.6 Admission to and discharge from hospital

Admission to hospital is frightening for many young children and disruptive for their families. A child with fever should only be admitted to hospital when absolutely necessary. Some conditions require frequent monitoring and treatment adjustments, which can only be done in hospital. Other conditions may be managed at home, sometimes with community healthcare support, such as ‘Hospital at Home’ schemes. The ability to manage a child at home will vary according to local facilities. The conditions that need admission to hospital will therefore vary.

Factors other than the child’s clinical condition can also influence the decision to admit a child with fever to hospital. These will include particular risk factors, such as travel to an area where malaria occurs, the family’s previous experience of illness and the ability of the family to return if their child’s condition worsens.

Review question

What factors other than the child's clinical condition should be considered when deciding to admit a child with fever to hospital?

Evidence summary

No evidence was found about when to admit children with fever to hospital.

GDG statement

The GDG agreed that the decision to admit or discharge a child with feverish illness should be made on the basis of clinical acumen after the child has been assessed (or reassessed) for the features of serious illness (i.e. 'red' or 'amber') and taking into account the results of investigations. The GDG also recognised that personal and social factors should also be taken into account when deciding whether or not to admit a child with fever to hospital. In the absence of evidence as to what these factors should be, the GDG decided it was appropriate to use the Delphi technique to inform the recommendation on admission to hospital.

When a child has a fever and no features of serious illness it is not usually necessary or appropriate for them to be cared for in hospital. However, there are circumstances where healthcare professionals should consider things apart from the child's clinical condition when deciding whether or not a child needs to be admitted to hospital, especially if alternative support systems, such as children's community nurses, are not available. No evidence was available for this topic. The GDG therefore used the Delphi panel to help produce broadly applicable recommendations in this area (see section 3.2).

Delphi statement 6

Healthcare professionals should consider the following factors, as well as the child's clinical condition, when deciding whether to admit a child with fever to hospital.

6.a Social and family circumstances

First round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
7 (13%)	20 (38%)	25 (47%)	1 (2%)		53	6

Second round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
2 (4%)	17 (33%)	33 (64%)			52	7

6.b Other illnesses suffered by the child or other family members

First round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
2 (4%)	17 (33%)	32 (60%)			53	7

Second round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	10 (19%)	41 (79%)			52	7.5

6.c Parental anxiety and instinct (based on their knowledge of their child)

First round

1 to 3	to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	14 (26%)	37 (70%)	1 (2%)		53	8

Second round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
2 (4%)	7 (13%)	43 (83%)			52	8

6.g Contacts with other people who have serious infectious diseases

First round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
4 (8%)	17 (32%)	28 (53%)	4 (8%)		53	7

Second round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	8 (15%)	42 (81%)	1 (2%)		52	8

6.h Recent travel abroad to tropical/subtropical areas, or areas with a high risk of endemic infectious disease

First round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
7 (13%)	12 (23%)	32 (60%)	2 (4%)		53	7

Second round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	2 (4%)	48 (92%)			51	8

6.i When the parent or carer's concern for their child's current illness has caused them to seek support or advice repeatedly

First round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
7 (13%)	15 (28%)	30 (57%)	1 (2%)		53	7

Second round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
2 (11%)	11 (22%)	38 (75%)			51	8

6.j Where the family has experienced a previous illness or death due to feverish illness which has increased their anxiety levels

First round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
2 (4%)	13 (25%)	37 (70%)	1 (2%)		53	8

Second round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	9 (17%)	42 (81%)	1 (2%)		52	8

6.k When a feverish illness has no obvious cause, but the child remains ill longer than expected for a self-limiting illness

First round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median

2 (4%)	13 (25%)	36 (70%)	1 (2%)	1	52	7

Second round

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
2 (4%)	9 (17%)	41 (79%)			52	8

GDG translation

Seven statements achieved agreement by the Delphi panel and were therefore used as recommendations.

An eighth factor (6.a Social and family circumstances) did not achieve the required level of agreement (64% scored 7–9; Median score 7). However, the GDG was aware of the associations between social deprivation and infection, hospital admission and death. The GDG decided this was an important factor to consider and unanimously agreed to include this as a recommendation.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

8.7 Referral to paediatric intensive care

Children with life-threatening infections may require paediatric intensive care. This is most likely to be beneficial if intensivists are involved in the child's management at an early stage.

GDG translation

The GDG agreed that children with the features of life-threatening illness that require immediate antibiotic treatment are also those likely to require paediatric intensive care. These children should be assessed and discussed with an intensivist at an early stage of their management.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Suspected meningococcal disease

The management of individual serious illnesses is strictly beyond the scope of this guideline. However, the GDG did come across evidence from the literature searches that they felt should be included in the guidance. The use of fluids for resuscitation in meningococcal disease is discussed in section 8.5 above.

Narrative evidence

Evidence for the use of immediate parenteral antibiotics is presented in Sections 9.3. An EL 2+[11] case–control study on the provision of health care for survivors and those who subsequently died from meningococcal disease was discussed earlier. In this study,[11] the failure to recognise disease complications, particularly in the absence of specific paediatric care, was associated with an 8.7-fold increase in the risk of death (P = 0.002). Not being under the care of a paediatrician was associated with a 66-fold increase (P = 0.005), failure of supervision a 19.5-fold increase (P = 0.015) and failure to administer inotropes a 23.7-fold increase (P = 0.005) in the risk of death. Not being under paediatric care was also highly correlated with a failure to recognise complications (P = 0.002; Fisher's exact test).

Evidence summary

In meningococcal disease, the evidence cannot conclude whether or not parenteral antibiotics given before admission have an effect on case fatality. However, the data are consistent with benefit when a substantial proportion of cases are treated. Failure to recognise complications of the disease increases the risk of death, as does not being under the care of a paediatric specialist.

GDG translation

The GDG noted that meningococcal disease is the leading cause of mortality among infectious diseases in childhood. Children with meningococcal disease may benefit from immediate parenteral antibiotics, especially if most children with meningococcal disease are treated. The GDG considers that there is

insufficient evidence of effectiveness or cost-effectiveness to change the current UK practice, which is to give parenteral antibiotics at the earliest opportunity. The GDG also recognises the importance of children with meningococcal disease being under the care of an experienced paediatric specialist. The GDG noted the need to anticipate complications.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

9 Antipyretic interventions

This section was partially updated in 2013.

Introduction

Fever is an increase in temperature that occurs as the result of the action of substances known as pyrogens upon the hypothalamus, the part of the brain that controls body temperature. These pyrogens have the effect of increasing the temperature set-point of the hypothalamus, which causes it to increase the temperature of the body.[205] The hypothalamus is sometimes likened to a thermostat, instigating heat promotion or loss procedures to achieve the desired set-point temperature. It is important to differentiate fever, which is regulated by the body, from hyperthermia, which is caused by external factors and is not regulated by the hypothalamus.

Fever is a normal physiological response to infection and a number of other conditions. Although it is a normal response, some people, including many doctors, nurses and parents, believe that fever should be treated to reduce temperature. This is usually either because of concerns about the damaging effect of fever or because it is thought to be a distressing symptom.[205,206] However, opinions differ about this, with others believing that fever should be allowed to run its course.[207]

If it is thought necessary to reduce fever, there are a number of interventions that are or have been used, either alone or in combination. Pharmacological treatments differ fundamentally from physical treatments, as they aim to lower the hypothalamic set-point rather than simply cool the body. If it is thought necessary to reduce fever, the safest, most clinically and cost-effective treatments and those most acceptable to the child should be used. The first question that the guideline development group (GDG) considered was what, if any, antipyretic interventions should be used. A variety of interventions were considered, specifically drugs, such as paracetamol and ibuprofen, and physical methods such as tepid sponging.

9.1 Effects of body temperature reduction

Antipyretics in the prevention of febrile convulsions

In addition to the underlying illness, fever may be accompanied by a number of unpleasant symptoms including pain, reduced eating and drinking, and reduced activity. In some cases, for example pain, this is likely to be the result of the illness causing the fever or the immune response to it. However, in other cases it is not always clear whether these are the direct result of the fever, or of the underlying illness, or a combination of the two. The GDG therefore considered the use of antipyretic interventions in the treatment of these symptoms. However, it is difficult to know what symptoms to measure and how to do so reliably.

A particular concern of many parents about fever in children is that it may cause fits, or febrile convulsions.[206] These are common in young children, and are very rarely associated with epilepsy or other problems in later life.[230] Because antipyretics reduce temperature, there is a theoretical rationale for their use in the prevention of febrile convulsions.

Review question

Does the use of antipyretic interventions in children with fever serve a benefit or harm in terms of any of the following:

- time to recovery
- wellbeing
- activity
- eating and drinking
- prevention of febrile convulsions?

We did not find any evidence against other interventions.

Narrative evidence

Research regarding the use of antipyretics in the prevention and treatment of febrile convulsions is limited. One EL 1+ review[231] that was judged to be adequate for inclusion owing to its clinical relevance, after obtaining methodological details from the author, and one EL 1+ systematic review (SR)[232] examining the use of antipyretic drugs as prophylaxis against febrile convulsions were found.

The first SR[231] investigated the hypothesis that paracetamol and ibuprofen, used prophylactically, will reduce the incidence of febrile convulsions across a wide variety of conditions. It found no evidence that the prophylactic use of antipyretics has any effect in reducing the incidence of febrile convulsions. The second review[232] assessed 12 studies of the effects of paracetamol for treating children in relation to fever clearance time, febrile convulsions and resolution of associated symptoms. It also found no evidence that the use of prophylactic paracetamol influenced the risk of febrile convulsions.

An EL 1+ double-blind randomised controlled trial (RCT)[228] analysing [225] datasets was also identified, which found that there was no significant difference in mean duration of fever (34.7 hours versus 36.1 hours, *P* not given) or of other symptoms (72.9 hours versus 71.7 hours). Children treated with paracetamol were more likely to be rated as having at least a 1-category improvement in activity ($P = 0.005$) and alertness ($P = 0.036$).

Evidence summary

Limited evidence was found regarding the use of antipyretic medications in the promotion of well- being, activity, eating and drinking, and no evidence of cost-effectiveness. One study suggested that parents could identify some improvement in activity and alertness after the administration of paracetamol, but not in mood, comfort, appetite or fluid intake. There is no evidence that the use of antipyretic agents reduces the incidence of febrile convulsions. (EL 1)

GDG translation

The GDG noted that, from the evidence, antipyretic agents do not appear to be effective in the prevention of febrile convulsions. There is very limited evidence regarding the effect of paracetamol on activity or other areas contained within the clinical question, which showed inconsistent effects.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Effect of antipyretics on the outcome of disease

This section was updated in 2013.

For the 2013 update, a review was undertaken on the effect of antipyretics on the outcome of disease, specifically to examine the hypothesis that the use of antipyretics could worsen severity of illness. The reason for addressing this question was that the GDG was aware of studies showing that the use of antipyretics to reduce fever could have an adverse effect on overall outcome; specifically, studies on adult patients in Intensive Care Units have shown higher mortality rates associated with use of antipyretics (Schulman et al., 2005; Lee et al., 2012) and a study of vaccination in children has shown that antibody production is inhibited when antipyretics were used to prevent post-vaccination fever (Prymula et al., 2009). The GDG wanted to see if the same pattern was found in children with feverish illnesses.

Review question

The clinical question outlined in the scope and examined in the review was “Whether reducing fever with paracetamol or non-steroidal anti-inflammatory drugs (NSAIDs) affects the course of the disease?”

Description of included studies

Seven studies were included in the review (Sugimura et al., 1994; Byington et al., 2002; Francois et al., 2010; Mikaeloff et al., 2007; Dubos et al., 2008; Lesko et al., 2001; Doran et al., 1988). One study was an RCT (Doran et al., 1988), four were prospective observational studies and two were retrospective observational studies.

Four studies examined the effect of antipyretics on outcomes in children with chickenpox (Mikaeloff et al., 2007; Dubos et al., 2008; Lesko et al., 2001; Doran et al., 1988) and three examined the effect on pneumonia (Sugimura et al., 1994; Byington et al., 2002; Francois et al., 2010). Sample sizes range from 156,034 in the retrospective study to 72 in the RCT.

For this review, post-vaccination fever was not counted as an illness.

Evidence statements

The GRADE profiles presented show results of included studies for the review question.

Table 9.1 GRADE findings for outcome of disease in children after antipyretics

Number of studies	Number of children		Effect		Quality
	Antipyretic	**No treatment**	**Relative (95% confidence interval)**	**Absolute (95% confidence interval)**	
Cases of complicated pneumonia vs. uncomplicated pneumonia using ibuprofen					
1 (Byington et al., 2002)	Ibuprofen	No treatment	Adjusted OR 4.0 (2.5 to 6.5), $P < 0.001$[a]	-	Very low
1 (Francois et al., 2010)	Ibuprofen	No treatment	Adjusted OR 2.57 (1.51 to 4.35), $P < 0.001$ [a]	-	Very low
Primary varicella with skin or soft tissue complications using paracetamol					
1 (Mikaeloff et al., 2007)	Paracetamol	No treatment	Adjusted RR 4.9 (2.1 to 11.4) [a]	-	Very low
1 (Mikaeloff et al., 2007)	Paracetamol	No treatment	Adjusted RR 1.5 (1.0 to 2.2) [a]	-	Very low
1 (Dubos et al., 2008)	Paracetamol	No treatment	Adjusted OR 4.8 (1.6 to 14.4), $P = 0.005$[a]	-	Very low
Risk of any invasive group A streptococcal infection using ibuprofen or paracetamol					
1 (Lesko et al., 2001)	Any ibuprofen during illness	No ibuprofen	OR 3.9 (1.3 to 12) [a]	-	Very low
1 (Lesko et al., 2001)	Any acetaminophen during illness	No acetaminophen	OR 1.2 (0.50 to 3.0) [a]	-	Very low

Number of studies	Number of children		Effect		Quality
	Antipyretic	No treatment	Relative (95% confidence interval)	Absolute (95% confidence interval)	
1 (Lesko et al., 2001)	Ibuprofen only	No medication	Matched OR 1.5 (0.58 to 11) [a]	-	Very low
1 (Lesko et al., 2001)	Acetaminophen only	No medication	Matched OR 0.98 (0.43 to 2.2)[a], Adjusted OR 0.94 (0.34 to 2.6) [a]	-	Very low
1 (Lesko et al., 2001)	Acetaminophen and ibuprofen	Neither	Matched OR 5.0 (1.6 to 16) [a] Adjusted OR 5.6 (1.2 to 25) [a]	-	Very low
Time to total scabbing using paracetamol					
1 (Doran et al., 1988)	6.7 days (SD 2.3)	5.6 days (SD 2.5)	*P* < 0.05 [a]	-	Very low
Time to last new vesicle using paracetamol					
1 (Doran et al., 1988)	3.9 days (SD 1.4)	4.1 days (SD 1.2)	*P* = 0.64 [a]	-	Very low
Time to total healing using paracetamol					
1 (Doran et al., 1988)	16.1 (SD 5.6)	16.2 (SD 5.8)	*P* = 0.45 [a]	-	Very low
Number of paracetamol doses used by parents					
1 (Sugimura et al., 1994)	Complicated pneumonia 2.52 (SD 0.80)	Pneumonia 1.37 (SD 0.72)	*P* < 0.001 [a]	-	Very low

OR odds ratio, P probability, RR relative risk

Note: Observational studies are set at low quality unless they have design aspects that increase this.

[a] As reported by authors.

Evidence statements

Two observational studies found higher rates of pneumonia with complications were associated with use of ibuprofen. This finding was statistically significant. The evidence for this finding was of very low quality, specifically because this type of study cannot confirm a causal link between use of ibuprofen and complicated pneumonia. Patients with severe pneumonia might be more likely to have a high fever and hence receive antipyretics.

One observational study found higher rates of group A streptococcal infection with the use of ibuprofen or ibuprofen and paracetamol combined, but not with use of paracetamol alone. The evidence for this finding was of very low quality.

Three observational studies found higher rates of varicella with skin complications associated with the use of paracetamol. This finding was statistically significant. The evidence for this finding was of very low quality. Again, this type of study cannot confirm a causal pathway between use of paracetamol and varicella with complications.

One RCT study found that time to 'scabbing' was shorter in children with chickenpox who received placebo compared with children who received paracetamol. This finding was statistically significant. However, for two further outcomes (last new vesicle and total healing time) there was no statistical association between paracetamol use and outcome. The evidence for these findings was of very low quality.

One observational study found that use of paracetamol was more frequent in patients with pneumonia with complications compared with those with pneumonia without complications. This finding was statistically significant. The evidence for this finding was of very low quality, and this study could not confirm a causal pathway between use of paracetamol and outcome.

Health economics profile

No health economic studies were identified for this question and no formal health economic analysis was undertaken.

Evidence to recommendations

Relative value placed on the outcomes considered

The GDG stated that the overarching aim of the guideline was the early and accurate detection of serious illness in children with fever. In addition, the GDG stressed the importance of avoiding unnecessary investigation or treatments.

Consideration of clinical benefits and harms

There is evidence to show that antipyretic use may be associated with more severe symptoms in underlying conditions. However, the GDG recognised that it was not possible to determine if the relationship between treatment and symptoms was causative. There are a number of possible pathways:

- Antipyretics reduce the body's ability to react to a disease and result in worsening symptoms, for example by altering the immune response.
- Antipyretics reduce symptoms, so they delay or stop treatment of underlying disease.
- Children with greater severity of illness are given antipyretics to relieve symptoms but these have no effect on the underlying severity of the condition.
- A combination of the above.

Whilst the GDG took account of the possibility of harm with antipyretic treatment, it recognised that there was no convincing evidence that this was likely to occur, and hence it recognised a role for antipyretics in treating some children with fever.

The GDG did not to make a recommendation on the effect of antipyretics on the outcome of disease as this was implicitly included in the later recommendations on the use of antipyretics.

Consideration of health benefits and resource uses

The GDG stated that as no definite link between antipyretic use and increased severity of illness could be established, the cost effectiveness of antipyretics could not be determined.

Quality of evidence

Seven studies were identified. The available evidence was of very low quality either due to using observational study designs or small sample sizes or high levels of missing values. In six of the seven studies it was not possible to establish the treatment regimen. Four of the studies examined chickenpox, where the use of antipyretics to relieve non-febrile symptoms is unclear. All the studies included children aged over 5 years.

In addition, the causal link between antipyretic use and increased severity of disease was not clearly established. Therefore, antipyretic use could be a consequence of disease severity rather than a cause of it, or could be a confounding factor on the causal pathway.

Other considerations

No inequalities issues were raised in relation to this question.

Recommendations

No recommendations were made based on this review.

9.2 Physical and drug interventions to reduce body temperature

Review question

What, if any, antipyretic interventions are effective in reducing body temperature in children with fever?

There are a number of interventions that can be undertaken to reduce temperature, both pharmacological and physical; however, it is not clear whether these treatments are either beneficial or necessary, or what the indications for the treatment of fever should be. Consequently, there is wide variation in practice, both with the use of interventions, and the outcomes that are aimed for. Some healthcare professionals aim to reduce temperature to what they consider to be normal, while others aim simply to reduce temperature. Although the circumstances under which interventions are used will vary, it is important that the possible benefits and harms of treating fever are understood. This includes any adverse effects from the interventions.

Elevations in body temperature result from rising levels of substances such as prostaglandins in the hypothalamus. This has the effect of resetting the hypothalamic temperature set-point and increasing temperature. Paracetamol and no steroidal anti-inflammatory agents such as ibuprofen inhibit the action of the cyclooxygenase enzymes involved in the production of prostaglandins, and this is the basis of their antipyretic activity, although inflammatory mediators other than prostaglandins may also be potential drug targets. Peripherally, the production of pyrogenic cytokines is also suppressed and the production of endogenous anti-inflammatory compounds is promoted.

Physical treatments such as tepid sponging cool the part of the body being sponged but do not reduce the levels of prostaglandins and so the temperature of the whole body is not reduced.

Furthermore, because the hypothalamus is still set at a higher temperature level, physical treatments may cause shivering and other adverse effects as the body aims to meet the hypothalamic set-point temperature, which continues to be raised. Shivering with a high temperature is sometimes referred to as a rigor.

Physical interventions

There are a number of physical interventions that can be used to reduce body temperature, including undressing, fanning and sponging with cool or cold water. These take advantage of heat loss through convection and evaporation but do not treat the underlying causes of the fever; either the disease or the alteration in hypothalamic set-point.

Narrative evidence

Two reviews[208,209] with EL 1+ and EL 2+ ratings, respectively, due to the nature of the included studies, were found. These compared tepid sponging with antipyretic drugs. One systematic review (SR)[210] which evaluated the benefits and harms of sponging techniques was also found. One further study compared undressing with paracetamol and tepid sponging.[211] There is a lack of evidence regarding opening windows or fanning as methods of reducing temperature. Tepid sponging offers no significant benefit over antipyretic agents alone.[209] In studies looking at combinations of sponging techniques and drugs, sponging seemed to have no or only short-lived additive effects on the reduction in temperature. Adverse effects in some children included crying and shivering in those treated with sponging. Undressing alone had little effect on temperature. A small study in adult volunteers with artificially induced fever showed that, during active external cooling, shivering was common, and both heat production and blood pressure were raised.[212] Discomfort was also significant, a finding that is supported by some studies of tepid sponging in children.[213]

GDG translation

Physical methods of temperature reduction do not treat the cause of fever, which is the action of circulating pyrogens occurring as the result of the underlying condition. Tepid sponging is time consuming, may cause distress, and has minimal medium- to long-term effects on temperature. Undressing appears to have little, if any, effect on temperature. There was no evidence regarding other physical methods of temperature control, for example fanning, although this shares the above

limitation. Physical methods may also cause shivering if the cooling is too much or too quick.[213] This may cause vasoconstriction and an increase in temperature and metabolism.

Because there is limited evidence regarding clothing of the feverish child, the GDG agreed by consensus that children with fever should be clothed appropriately for their surroundings, with the aim of preventing overheating or shivering. The major consideration should be the comfort of the child, and the prevention of over-rapid cooling that may cause shivering which may be distressing for child and parents. Care also needs to be taken not to overdress febrile children. It is not possible to be prescriptive about this because of varying environmental and other conditions, and the provision of information about appropriate clothing is an important role for healthcare professionals. In view of the lack of evidence from clinical studies for or against the use of physical cooling methods, the GDG concluded that research in this area may be beneficial.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Drug interventions

This section was updated in 2013.

The primary method of temperature control is the use of antipyretic drugs such as paracetamol and ibuprofen. Unlike the physical methods previously discussed, these do treat the proximal cause of fever, the increased hypothalamic set-point, although neither physical nor pharmacological methods treat the ultimate cause, for example an underlying infection. The GDG sought to identify the most appropriate pharmacological treatment for fever (as distinct from the cause of the fever), considering not only antipyretic efficacy but also safety and cost.

Review question

The clinical question outlined in the scope for the update is to establish the 'Effect on fever and associated symptoms of treatment with:

- paracetamol alone or non-steriodal anti-inflammatory drugs (NSAIDs) alone, compared with placebo and with one another
- alternating paracetamol and NSAIDs, compared with placebo, either drug alone, and taking both at the same time
- paracetamol and NSAIDs taken at the same time, compared with placebo, and either drug alone and either drug alone.'

Although the question states that any NSAID will be included, after assessment of available evidence, and based on the experience of the GDG, it was apparent that ibuprofen is the only NSAID in common use in the UK and it is the only NSAID licenced for this indication in children.

This question includes ten specific comparisons:

- paracetamol compared with placebo
- ibuprofen compared with placebo
- paracetamol and ibuprofen combined compared with placebo
- paracetamol and ibuprofen alternating compared with placebo
- paracetamol compared with ibuprofen
- paracetamol compared with paracetamol and ibuprofen combined
- paracetamol compared with paracetamol and ibuprofen alternating
- ibuprofen compared with paracetamol and ibuprofen combined
- ibuprofen compared with paracetamol and ibuprofen alternating
- paracetamol and ibuprofen combined compared with paracetamol and ibuprofen alternating.

Description of included studies

Twenty-five studies were included in this review (Gupta et al., 2007; Walson et al., 1989; Kauffman et

al., 1992; Wilson et al., 1991; Brewer et al., 1968; Autret et al., 1997; Nadal et al., 2002; Sarrell et al., 2006; Wong et al., 2001; Erlewyn-Lajeunesse et al., 2006; Sidler et al., 1990; Ulukol et al., 1999; Autret et al., 1994; McIntyre et al., 1996; Vauzelle-Kervrodan et al., 1997; Van Esch et al., 1995; Autret-Leca et al., 2007; Southey et al., 2009; Beasley et al., 2008; Pierce et al., 2010;Hay et al., 2009; Pashapour et al., 2009; Kramer et al., 2008; Paul et al., 2010; Nabulsi et al., 2006). No studies were identified comparing alternating or combining antipyretics with placebo.

All the identified studies were RCTs, meta-analyses of RCTs or large case series. Studies were mainly undertaken in Europe and North America. Study dates ranged from the 1980s to 2010.

Assessment of effectiveness was hampered by the heterogeneity of study protocols, with different entry criteria, outcomes, measurement times and dosages of drugs being used. As a result of this heterogeneity, a meta-analysis was not undertaken.

A number of calculations have been used in this review. For a relative risk, an effect size of 0.25 with the 95% confidence interval (CI) not crossing 1 (no effect) was considered a large effect. For standardised mean differences, an effect size of 0.1 was considered small, 0.3 medium and 0.5 large.

Evidence profile

Evidence was found for 8 of the 10 comparisons and results of these studies are summarised in the following GRADE tables (for full evidence profiles see Appendix I):

- Table 9.2 − paracetamol compared with placebo
- Table 9.3 − ibuprofen compared with placebo
- Table 9.4 − paracetamol compared with ibuprofen
- Table 9.5 −paracetamol compared with paracetamol and ibuprofen combined
- Table 9.6 − paracetamol compared with paracetamol and ibuprofen alternating
- Table 9.7 − ibuprofen compared with paracetamol and ibuprofen combined
- Table 9.8 − ibuprofen vparacetamol and ibuprofen alternating
- Table 9.9 − paracetamol and ibuprofen combined compared with paracetamol and ibuprofen alternating.

Table 9.2 GRADE findings for paracetamol vs. placebo

Number of studies	Number of children		Effect*		Quality
	Paracetamol	Placebo	Relative risk (95% confidence interval)	Absolute (95% confidence interval)	
Quality of life at 1 to 2 hours					
No data					
Quality of life at > 2 to 5 hours					
At 4 hours					
Comfort 1 (Gupta et al., 2007)	19 of 103 [a]	9 of 107[b]	RR 2.19 (1.04, 4.62)	-	Low
Activity 1 (Gupta et al., 2007)	29 of 103 [a]	4 of 107[b]	RR 7.53 (2.74, 20.67)	-	Low
Alertness 1 (Gupta et al., 2007)	22 of 103 [a]	4 of 107[b]	RR 5.71 (2.04, 16.01)	-	Low
Mood 1 (Gupta et al., 2007)	1 of 103 [a]	3 of 107[b]	RR 3.81 (1.09, 13.26)	-	Low
Appetite 1 (Gupta et al., 2007)	7 of 103 [a]	1 of 107[b]	RR 7.27 (0.91, 58.08)	-	Low
Quality of life at > 5 to 24 hours					
At 6 hours					
Comfort 1 (Gupta et al., 2007)	38 of 103 [a]	8 of 107[b]	RR 4.93 (2.42 to 10.06)	-	Low
Activity 1 (Gupta et al., 2007)	62 of 103 [a]	17 of 107[b]	RR 3.79 (2.38 to 6.02)	-	Low
Alertness 1 (Gupta et al., 2007)	60 of 103 [a]	22 of 107[b]	RR 2.83 (1.89, 4.26)	-	Low
Mood 1 (Gupta et al., 2007)	37 of 103 [a]	13 of 107[b]	RR 2.96 (1.67 to 5.23)	-	Low
Appetite 1 (Gupta et al., 2007)	21 of 103 [a]	1 of 103[b]	RR 21.00 (2.88 to 153.23)	-	Low

Number of studies	Number of children		Effect*		Quality
	Paracetamol	Placebo	Relative risk (95% confidence interval)	Absolute (95% confidence interval)	
Discomfort at > 24 hours					
No data					
Mean temperature at 1 to 2 hours					
1 hour					
1 (Walson et al., 1989a)	101.2ºF (SD 0.9), n = 31[c]	102.1ºF (SD 0.9), n = 33[b]	SMD -0.99 (-1.51 to -0.47)	-	Low
1 (Gupta et al., 2007)	38.4ºC (SD 1.0), n = 101 [a]	38.7ºC (SD 0.9), n = 102[b]	SMD -0.31 (-0.59 to -0.04)	-	Low
1 (Kauffman et al., 1992)	38.2ºC (SD 0.5657), n = 8[c]	38.8ºC (SD 0.6), n = 9[b]	SMD -0.97 (-2.00 to +0.05)	-	Very Low
2 hours					
1 (Walson et al., 1989a)	100.3ºF (SD 0.9) n = 31[c]	101.8ºF (SD 1.3), n = 33[b]	SMD -1.32 (-1.86 to 0.77)	-	Low
1 (Gupta et al., 2007)	38.0ºC (SD 0.8), N = 101 [a]	38.6ºC (SD 0.9), n = 102[b]	SMD -0.70 (-0.99 to -0.42)	-	Low
1 (Kauffman et al., 1992)	37.7ºC (SD 0.6), n = 8[c]	39.0ºC (SD 0.56), n = 9[b]	SMD -2.13 (-3.39 to -0.88)	-	Very Low
Mean in temperature at > 2 to 5 hours					
3 hours					
1 (Walson et al., 1989a)	100.1ºF (SD 1.0), n = 31[c]	101.7ºF (SD 1.4), n = 33[b]	SMD -1.29 (-1.83, -0.75)	-	Very Low
1 (Gupta et al., 2007)	37.8ºC (SD 0.8), n = 101 [a]	38.55ºC (SD 1.0), n = 102[b]	SMD -0.82 (-1.11 to -0.54)	-	Low
1 (Kauffman et al., 1992)	37.7ºC (SD 0.8485), n = 8[c]	39.2ºC (SD 0.9), n = 9[b]	SMD -1.62 (-2.76 to -0.49)	-	Very Low
4 hours					
1 (Walson et al., 1989a)	100.3ºF (SD 1.3), n = 31[c]	101.6ºF (SD 1.5), n = 33[b]	SMD -0.91 (-1.43 to -0.40)	-	Low
1 (Gupta et al., 2007)	37.6ºC (SD 0.8), n = 101 [a]	38.5ºC (SD 1.0), n = 102[b]	SMD -0.99 (-1.28 to -0.70)	-	Low
1 (Kauffman et al., 1992)	37.8ºC (SD 0.8485), n = 8[c]	39.4ºC (SD 0.6), n = 9[b]	SMD -2.09 (-3.33 to -0.85)	-	Very Low
5 hours					
1 (Walson et al., 1989a)	100.5ºF (SD 1.3), n = 31[c]	101.3ºF (SD 1.6), n = 33[b]	SMD -0.54 (-1.04 to -0.04)	-	Low
1 (Gupta et al., 2007)	37.6ºC (SD 0.7), n = 101 [a]	38.4ºC (SD 0.9), n = 102[b]	SMD -0.99 (-1.28 to -0.70)	-	Low

Number of studies	Number of children		Effect*		Quality
	Paracetamol	Placebo	Relative risk (95% confidence interval)	Absolute (95% confidence interval)	
1 (Kauffman et al., 1992)	38.1ºC (SD 0.5657), n = 8[c]	39.4ºC (SD 0.9), n = 9[b]	SMD -1.62 (-2.76 to -0.48)	-	Very Low
Mean in temperature at > 5 to 24 hours					
6 hours					
1 (Walson et al., 1989a)	100.8ºF (SD 1.9), n = 31[c]	101.2ºF (SD 1.5), n = 33[b]	SMD -0.23 (-0.72 to 0.26)	-	Low
1 (Gupta et al., 2007)	37.7ºC (SD 0.7), n = 101 [a]	38.3ºC (SD 1.0), n = 102[b]	SMD -0.69 (-0.98 to -0.41)	-	Low
1 (Kauffman., et al., 1992)	38.5ºC (SD 1.1314), n = 8[c]	39.3ºC (SD 0.6), n = 9[b]	SMD -0.85 (-1.86 to +0.15)	-	Very Low
8 hours					
1 (Walson et al., 1989a)	101.6ºF (SD 1.8), n =31[c]	101.2ºF (SD 1.7), n = 33[b]	SMD 0.23 (-0.27 to +0.72)	-	Very Low
1 (Kauffman et al., 1992)	38.8ºC (SD 0.8485), n = 8[c]	39.2ºC (SD 0.6), n = 9[b]	SMD -0.52 (-1.50 to +0.45)	-	Very Low
Mean in temperature at > 24 hours					
No data					
Mean change temperature at 1 to 2 hours					
2 hours					
1 (Gupta et al., 2007)	70.3% (SD 24.8), n = 101 [a]	30.7% (SD 26.1), n = 102[b]	SMD 1.55 (+1.24 to +1.86)	-	High
Mean change temperature at > 2 to 5 hours					
4 hours					
1 (Gupta et al., 2007)	85.4% (SD 22.4), n = 101 [a]	45.5% (SD 34.1), n = 102[b]	SMD 1.38 (+1.68 to +1.07)	-	High
Mean change temperature at > 5 to 24 hours					
6 hours					
1 (Gupta et al., 2007)	87.6% (SD18.6), n = 101[a]	51.0% (SD 33.3), n = 102[b]	SMD 1.35 (+1.66 to +1.04)	-	High
Mean change temperature at > 24 hours					
No data					
Afebrile at 1 to 2 hours					
No data					

Number of studies	Number of children		Effect*		Quality
	Paracetamol	Placebo	Relative risk (95% confidence interval)	Absolute (95% confidence interval)	
Afebrile at > 2 to 5 hours					
No data					
Afebrile at > 5 to 24 hours					
No data					
Afebrile at > 24 hours					
No data					
Temperature area under the curve					
0 to 8 hours					
1 (Kauffman et al., 1992)	328 (-356 to 686)[c]	- 67 (-629 to 120)[b]	*P* < 0.01	-	Low
1 (Walson et al., 1989)	365.0[c]	166.5[b]	*P* < 0.05	-	Low
0 to 6 hours					
1 (Wilson et al., 1991)	6.72 (+/- 0.58)[e]	11.70 (0.83)[b]	-	-	Low
Adverse events					
5 (Southey et al., 2009; Brewer et al., 1968; Gupta et al., 2007; Kauffman et al., 1992; and Walson et al., 1989)	23 of 385	12 of 371	RR 1.81 (0.94 to 3.50)	-	Very low
Long-term effects of paracetamol – Asthma					
1 (Beasley et al., 2008)	NR	NR	RR 1.46 (1.36 to 1.56)	-	Very low
Long-term effects of paracetamol – Rhinoconjunctivitis					
1 (Beasley et al., 2008)	NR	NR	RR 1.48 (1.38 to 1.60)	-	Very low
Long-term effects of paracetamol – Eczema					
1 (Beasley et al., 2008)	NR	NR	RR 1.35 (1.26 to 1.45)	-	Very low

NR not reported, P probability, RR relative risk, SD standard deviation, SMD standard mean difference

* Relative and absolute differences are calculated by the NCC technical team based on the data presented in the papers. When this data is unavailable the authors reported figures may be used.

[a] 15 mg/kg paracetamol repeated at 6 hours

[b] Placebo

[c] 10 mg/kg paracetamol single dose

[d] AUC of percentage decrease of temperature (from baseline to 98.6 ºF) vs. time

[e] 120 mgm/5 ml

Table 9.3 GRADE findings for ibuprofen vs. placebo

Number of studies	Number of children		Effect[a]		Quality
	Intervention	Control	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Discomfort at 1 to 2 hours					
No data					
Discomfort at > 2 to 5 hours					
No data					
Discomfort at > 5 to 24 hours					
No data					
Discomfort > 24 hours					
No data					
Mean change temperature at 1 to 2 hours					
No data					
Mean change in temperature at > 2 to 5 hours					
No data					
Mean change in temperature at > 5 to 24 hours					
No data					
Mean change in temperature at > 24 hours					
No data					
Mean temperature at 1 to 2 hours					
1 hours					
1 (Walson et al., 1989)	100.9ºF (SD 1.0), n = 29[b]	102.1ºF (SD 0.9), n = 33[c]	SMD -1.25 (-1.80 to -0.70)	-	Moderate
1 (Walson et al., 1989)	100.8ºF (SD 0.9), n = 25[d]	102.1ºF (SD 0.9), n = 33 [c]	SMD -1.43 (-2.01 to -0.84)	-	Moderate
2 hours					
1 (Walson et al., 1989)	99.8ºF (SD 1.1), n = 29[b]	101.8ºF (SD 1.3), n = 33 [c]	SMD -1.63 (-2.21 to -1.05)	-	Moderate
1 (Walson et al., 1989)	99.5ºF (SD 0.7), n = 25[e]	101.8ºF (SD 1.3), n = 33 [c]	SMD -2.09 (-2.75 to -1.44)	-	Moderate
Mean temperature at > 2 to 5 hours					
3 hours					
1 (Walson et al., 1989)	99.5ºF (SD 1.3), n = 29[b]	101.7ºF (SD 1.4), n = 33 [c]	SMD -1.60 (-2.18 to -1.03)	-	Moderate
1 (Walson et al., 1989)	99.3ºF (SD 0.7), n = 25[d]	101.7ºF (SD 1.4), n = 33 [c]	SMD -2.05 (-2.70 to -1.41)	-	Moderate

Number of studies	Number of children		Effect[a]		Quality
	Intervention	Control	Relative (95% confidence interval)	Absolute (95% confidence interval)	
4 hours					
1 (Walson et al., 1989)	99.5ºF (SD 1.6) n = 29[b]	101.6ºF (SD 1.5) n = 33 [c]	SMD -1.34 (-1.90 to -0.78)	-	Moderate
1 (Walson et al., 1989)	99.2ºF (SD 1.2) n = 25[e]	101.6ºF (SD 1.5) n = 33 [c]	SMD -1.72 (-2.33 to -1.10)	-	Moderate
5 hours					
1 (Walson et al., 1989)	99.8ºF (SD 1.9) n = 29[b]	101.3ºF (SD 1.6) n = 33 [c]	SMD -0.85 (-1.37 to -0.33)	-	Moderate
1 (Walson et al., 1989)	99.3ºF (SD 1.7 n = 25[e]	101.3ºF (SD 1.6 n = 33 [c]	SMD -1.20 (-1.77 to -0.63)	-	Moderate
Mean temperature at > 5 to 24 hours					
6 hours					
1 (Walson et al., 1989)	100.2ºF (SD 2.2) n = 29[b]	101.2ºF (SD 1.5) n = 33[c]	SMD -0.53 (-1.04 to -0.02)	-	Moderate
1 (Walson et al., 1989)	99.7ºF (SD 1.9) n = 25[d]	101.2ºF (SD 1.5) n = 33[c]	SMD -0.88 (-1.42 to -0.33)	-	Moderate
7 hours					
1 (Walson et al., 1989)	101.2ºF (SD 2.0) n = 29[b]	101.2ºF (SD 1.7) n = 33 [c]	SMD 0.00 (-0.50 to +0.50)	-	Low
1 (Walson et al., 1989)	100.6ºF (SD 2.2) n = 25[d]	101.2ºF (SD 1.7) n = 33 [c]	SMD -0.31 (-0.83 to +0.22)	-	Low
Mean temperature at > 24 hours					
No data					
Afebrile at 1 to 2 hours					
No data					
Afebrile at > 2 to 5 hours					
No data					
Afebrile at > 5 to 24 hours					
No data					
Afebrile at > 24 hours					
No data					
Temperature area under the curve					
0 to 8 hours					
1 (Kauffman et al., 1992)	730 (576 to 839)[f]	-67 (-629 to 120) [c]	*P* < 0.01	-	Low
1 (Kauffman et al., 1992)	590 (160 to 875)[d]	-67 (-629 to 120 [c]	*P* < 0.01	-	Low

Number of studies	Number of children		Effect[a]		Quality
	Intervention	Control	Relative (95% confidence interval)	Absolute (95% confidence interval)	
1 (Walson et al., 1989)	460.9 [b]	139.0 [c]	*P* < 0.05	-	Low
1 (Walson et al., 1989)	510.8 [d]	139.0 [c]	*P* < 0.05	-	Low
0 to 6 hours – Change in temperature					
1 (Wilson et al., 1991)	7.09 (0.58)[b]	11.70 (0.83) [c]		-	Low
1 (Wilson et al., 1991)	4.91 (0.47)[d]	11.70 (0.83) [c]		-	Low
Adverse events					
4 (Southey et al., 2009; Kauffman et al.,1992; Walson et al., 1989a; and Wilson et al., 1991)	55 of 357	27 of 294	RR 1.67 (1.12, 2.48)	-	Low

NR [n]ot reported, P probability, RR relative risk, SD standard deviation, SMD standard mean difference

[a] Relative and absolute differences are calculated by the NCC technical team based on the data presented in the papers. When this data is unavailable the authors reported figures may be used.

[b] Dose of 5 mg/kg

[c] Placebo

[d] Dose of 10 mg/kg ibuprofen

[e] Children aged more than 5 years included in study (Gupta up to 6; Walson up to 11; Kauffman up to 12; Wilson up to 12)

[f] 7.5 mg/kg ibuprofen

Table 9.4 GRADE findings for paracetamol vs. ibuprofen

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Discomfort at 1 to 2 hours					
No data					
Discomfort at 2 to 5 hours					
4 hours					
GBC-score 1 (Autret et al., 1997)	0.8 (SD 1), n = 116[a]	0.6 (SD 0.9), n = 113[b]	NS [c]	-	Low

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
GBC-VAS 1 (Autret et al., 1997)	27.8 (SD 29.5), n = 114[a]	18.3 (SD 26.5), n = 108[b]	SMD 0.34 (0.07, 0.60)	-	Low
CHEOPs discomfort 1 (Autret et al., 1997)	2.2 (SD 0.9), n = 114[a]	2.5 (1.0), n = 108[b]	NS [c]	-	Low
Overal efficacy 1 (Figueras Nadal et al., 2002)	64 of 94[d]	61 of 93[f]	RR 1.04 (0.85 to 1.27)	-	Very low
Discomfort at > 5 to 24 hours					
6 hours					
GBC-score 1 (Autret et al., 1997)	0.8 (SD 1.0), n = 114[a]	0.5 (SD 1.0), n = 112[b]	NS [c]	-	Low
GBC-VAS 1 (Autret et al., 1997)	26.7 (SD 30.6), n = 112[a]	15.9 (SD 31.1), n = 107[b]	SMD 0.35 (+0.08 to +0.62)	-	Low
CHEOPs discomfort 1 (Autret et al., 1997)	2.3 (SD 0.9), n = 112[a]	2.5 (SD 1), n = 107[b]	NS [c]	-	Low
Discomfort > 24 hours					
Day 1					
NCCPC stress test 1 (Sarrell et al., 2006)	11.48 (SD 2.58), n = 15511[g]	11.77 (SD 2.64), n=154[h]	SMD -0.11 (-0.33 to +0.11)	-	High
Day 2					
NCCPC stress test 1 (Sarrell et al., 2006)	8.83 (SD 2.67) n = 155 [g]	8.87 (SD 2.54) n=154[h]	SMD 0.02 (-0.24 to +0.21)	-	High
Day 3					
NCCPC stress test 1 (Sarrell et al., 2006)	7.96 (SD 2.71), n = 155[g]	7.66 (SD 2.96) n=154[g]	SMD 0.11 (-0.12 to +0.33)	-	High
Mean change temperature at 1 to 2 hours					
1 hour					
1 (Autret et al., 1997)	-0.97°C (SD 0.58) n=114 [a]	-0.90°C (SD 0.56) n=114[b]	SMD -0.12 (-0.38 to + 0.14)	-	Low

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
1 (Wong et al., 2001)	-1.00°C (SD 0.65), n=185[i]	-1.05°C (SD 0.70), n=191[j]	SMD 0.07 (-0.13 to +0.28)	-	Moderate
1 (Erlewyn-Lajeunesse et al., 2006)	-0.92°C (95% CI 0.70 to 1.14), n=35[g]	-0.95°C (95% CI 0.72 to 1.17), n=37[k]	SMD 0.04 (-0.42 to +0.51)	-	Low
1 (Wilson et al., 1991)	-0.8 (SD 0.3279), n = 43[g]	-0.8 (SD 0.3606), n = 52[h]	SMD 0.04 (-0.42 to +0.51)	-	Very low
1 (Wilson et al., 1991)	-0.8 (SD 0.3428), n = 47 [l]	-0.8 (SD 0.3606), n = 52[h]	SMD 0.00 (-0.40 to +0.40)	-	Very low
1.5 hours					
1 (Wong et al., 2001)	-1.33°C (SD 0.66), n=185[i]	-1.33°C (SD 0.68) n=191[j]	SMD 0.00 (-0.20 to +0.20)	-	Moderate
2 hours					
1 (Wong et al., 2001)	-1.56°C (SD 0.72), n = 185 [i]	-1.55°C (SD 0.68) n = 191[j]	SMD: -0.01 (-0.22 to +0.19)	-	Moderate
1 (Wilson et al., 1991)	-1.2 (SD 0.6557), n = 43[g]	-1.2 (SD 0.7211), n = 52[h]	SMD 0.00 (-0.40 to +0.40)	-	Very low
1 (Wilson et al., 1991)	-1.2 (SD 0.6856), n = 47 [l]	-1.2 (SD 0.7211), n = 52[h]	SMD 0.00 (-0.39 to +0.39)	-	Very low
Mean change in temperature at > 2 to 5 hours					
3 hours					
1 (Wong et al., 2001)	-1.58°C (SD 0.81), n=185 [i]	-1.52°C (SD 0.79) n=191 [j]	SMD -0.07 (-0.28 to +0.13)	-	Moderate
1 (Wilson et al., 1991)	-1.5 (SD 0.6856), n = 47[g]	-1.4 (SD 0.7211), n = 52[h]	-0.14 (-0.54 to +0.25)	-	Very low
1 (Wilson et al., 1991)	-1.4 (SD 0.6557), n = 43 [l]	-1.4 (SD 0.7211), n = 52[h]	0.00 (-0.40 to +0.40)	-	Very low
4 hours					
1 (Autret et al., 1997)	-1.42°C (SD 0.85) n=112[a]	-1.04°C (SD 0.85) n= 110[b]	SMD -0.45 (-0.71 to -0.18)	-	Very low
1 (Wilson et al., 1991)	-1.6 (SD 0.6856), n = 47[g]	-1.3 (SD 1.4422), n = 52[h]	SMD -0.26 (-0.66 to +0.14)		Very low
1 (Wilson et al., 1991)	-1.2 (SD 0.6557), n = 43 [l]	-1.3 (SD 1.4422), n = 52[h]	SMD 0.09 (-0.32 to +0.49)	-	Very low
1 (Ulukol et al., 1999)	-1.86°C (SD 0.74) n = 30 [l]	-1.29°C (SD 0.71) n= 30[b]	SMD 0.78 (+0.25 to +1.30)	-	Low

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
1 (Autret et al., 1994)	-1.32°C (SD 1.00), n = 77[a]	-1.02°C (SD 1.05), n = 74[b]	SMD -0.29 (-0.61 to +0.03)	-	Low
1 (McIntyre et al., 1996)	-1.80°C (SD -), n = 76[m]	-1.6°C (SD -), n = 74[n]	*P* = 0.39	-	Moderate
1 (Nadal et al., 2002)	-1.30°C (SD 1.1), n = 94[o]	-1.20°C (SD 0.96), n = 93[p]	SMD -0.10 (-0.38 to +0.19)	-	Very low
1 (Wong et al., 2001)	-1.44°C (SD 0.98), n = 185 [i]	-1.47°C (SD 0.91), n = 191[j]	SMD 0.03 (-0.17 to +0.23)	-	Moderate
5 hours					
1 (Wong et al., 2001)	-1.35°C (SD 1.06), n = 185 [i]	-1.34°C (SD 1.05)[j] n = 191[j]	SMD -0.01 (-0.21 to +0.19)	-	Moderate
1 (Wilson et al., 1991)	-1.4 (SD 0.6856), n = 47 [g]	-1.0 (SD 1.4422), n = 52[h]	SMD -0.35 (-0.74 to +0.05)	-	Very low
1 (Wilson et al., 1991)	-1.1 (SD0.6557), n = 43 [l]	-1.0 (SD 1.4422), n = 52[h]	SMD -0.09 (-0.49 to +0.32)	-	Very low
Mean change in temperature at > 5 to 24 hours					
6 hours					
1 (Autret et al., 1997)	-1.19°C (SD 0.94), n = 108[a]	-0.88°C (SD 0.85), n=108[b]	SMD -0.34 (-0.61 to -0.08)	-	Low
1 (Wong et al., 2001	-1.24°C (SD 1.08), n=185 [i]	-1.20°C (SD 1.09) n=191[j]	SMD -0.04 (-0.24 to +0.17)	-	Moderate
1 (Wilson et al., 1991)	-1.1 (SD 0.6557), n = 43[g]	-0.9 (SD 1.4422), n = 52[h]	SMD -0.17 (-0.58 to +0.23)	-	Very low
1 (Wilson et al., 1991)	-1.2 (SD 0.6856), n = 47 [l]	-0.9 (SD 1.4422), n = 52[h]	SMD -0.26 (-0.66 to +0.14)	-	Very low
Mean change in temperature at > 24 hours					
No studies found					
Mean temperature at 1 to 2 hours					
1 hour					
1 (Kauffman et al., 1992)	38.0ºC (SD 0.6928), n = 12[a]	38.2°C (SD 0.5657), n = 8[b]	SMD -0.30 (-1.20 to +0.60)	-	Very low
1 (Kauffman et al., 1992)	37.9 (SD 0.4243), n = 8 [l]	38.2 (SD 0.5657), n = 8[b]	SMD -0.57 (-1.57 to +0.44)	-	Very low
1 (Vauzelle-Kervrodan et al., 1997)	38.4°C (SD 0.6) n = 60[a]	38.3°C (SD 0.6), n = 56[b]	SMD: 0.17 (-0.20 to +0.53)	-	Low

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
1 (Erlewyn-Lajeunesse et al., 2006)	37.81°C (SD 0.69), n = 35[g]	37.98°C (SD 0.47), n = 37[k]	SMD: -0.29 (-0.75 to +0.18)	-	Low
1 (Walson et al., 1989)	100.9°F (SD 1), n = 29 [g]	102.1°F (SD 0.9), n = 31[b]	SMD -0.31 (-0.82 to +0.20)	-	Low
1 (Walson et al., 1989)	100.8°F (SD 0.9) n = 25 [l]	102.1°F (SD 0.9), n = 31[b]	SMD -0.44 (-0.97 to +0.10)	-	Low
1 (Nadal et al., 2002)	37.93°C (SD 0.72) , n = 100 [o]	38.06°C (SD 0.72), n = 99 [p]	SMD -0.18 (-0.46 to +0.10)	-	Very low
1.5 hours					
1 (Nadal et al., 2002)	37.61°C (SD 0.73), n = 100[o]	37.78°C (SD 0.70), n = 99[p]	SMD -0.24 (-0.52 to +0.04)	-	Very low
2 hours					
1 (Nadal et al., 2002)	37.50°C (SD 0.74) n = 100[o]	37.67°C (SD 0.78), n = 99[p]	SMD -0.22 (-0.50 to +0.06)	-	Very low
1 (Van Esch et al., 1995)	37.60°C (SD 0.6025), n = 30[g]	37.96°C (SD 0.9155), n = 29[b]	SMD -0.46 (-0.98 to +0.06)	-	Very low
1 (Vauzelle-Kervrodan et al., 1997)	37.9°C (SD 0.7), n = 58[l]	37.9°C (SD 0.7), n = 55[b]	SMD 0.00 (-0.37 to +0.37)	-	Low
1 (Walson et al., 1989)	99.8°F (SD 1.1), n = 29[g]	101.8°F (SD 0.9), n = 31[b]	SMD -0.49 (-1.01 to +0.02)	-	Low
1 (Walson et al., 1989)	99.5°F (SD 0.7) n = 25[l]	101.8°F (SD 0.9), n = 31[b]	SMD -0.97 (-1.52 to -0.41)	-	Moderate
1 (Kauffman et al., 1992)	37.3°C (SD 0.5196), n = 12[a]	37.7°C (SD 0.6), n = 8[b]	SMD -0.69 (-1.62 to +0.23)	-	Very low
1 (Kauffman et al., 1992)	37.2°C (SD 0.2828), n = 8[q]	37.7°C (SD 0.6), n = 8[b]	SMD -1.01 (-2.07 to +0.05)	-	Low
1 (Autret-Leca et al., 2007)	37.4 (SD 0.75), n = 151 [l]	37.4 (SD 0.8), n = 150[r]	SMD 0.00 (-0.23 to +0.23)	-	Very low
Mean temperature at > 2 to 5 hours					
3 hours					
1 (Walson et al., 1989)	99.5°F (SD 1.0), n = 29[g]	101.7°F (SD 1.0), n = 31[b]	SMD -0.51 (-1.03 to 0.00)	-	Low
1 (Walson et al., 1989)	99.3 °F (0.7), n = 25 [l]	101.7 °F (SD 1.0),n = 31[b]	SMD -0.90 (-1.45 to -0.34)	-	Moderate
1 (Vauzelle-Kervrodan et al., 1997)	37.6°C(SD 0.7), n = 58 [l]	37.8°C (SD 0.7), n = 56[b]	SMD -0.28 (-0.65 to +0.09)	-	Very Low
1 (Nadal et al., 2002)	37.57°C (SD 0.92), n = 100[o]	37.78°C (SD 0.92), n = 99[p]	SMD -0.23 (-0.51 to +0.05)	-	Very low

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
1 (Kauffman et al., 1992)	36.9°C (SD 0.6928), n = 12[a]	37.7°C (SD 0.8485), n = 8[b]	SMD -1.01 (-1.97 to -0.05)	-	Very low
1 (Kauffman et al., 1992)	36.7°C (SD 0.2828), n = 8 [l]	37.7°C (SD 0.8485), n = 8[b]	SMD -1.49 (-2.64 to -0.35)	-	Very low
1 (Autret-Leca et al., 2007)	37.3°C (SD 0.75), n = 151 [l]	37.3°C (SD 0.75), n = 150[ah]	SMD 0.00 (-0.23 to +0.23)	-	Very low
4 hours					
1 (Van Esch et al., 1995)	37.38°C (SD 1.0022), n = 31[g]	37.95°C (SD 1.2806), n = 31[b]	SMD -0.49 (-1.00 to +0.02)	-	Very low
1 (Vauzelle-Kervrodan et al., 1997)	37.6°C (SD 0.8), n = 58 [l]	37.8°C (SD 0.8), n = 55[b]	SMD -0.25 (-0.62 to +0.12)	-	Very Low
1 (Nadal et al., 2002)	37.82°C (SD 1.05), n = 100[o]	37.97°C (SD 1.02), n = 99 [p]	SMD -0.14 (-0.42 to +0.13)	-	Very low
1 (Walson et al., 1989)	99.5°F (SD 1.6), n = 29[g]	101.6°F (SD 1.3), n = 31[b]	SMD -0.54 (-1.06 to -0.03)	-	Moderate
1 (Walson et al., 1989)	99.2°F (SD 1.2), n = 25 [l]	101.6°F (SD 1.3), n = 31[b]	SMD -0.86 (-1.42 to -0.31)	-	Moderate
1 (Kauffman et al., 1992)	36.9°C (SD 0.6928), n = 12[a]	37.8°C (SD 0.8485), n = 8[b]	SMD -1.14 (-2.12 to -0.16)	-	Very low
1 (Kauffman et al., 1992)	36.7°C (SD 0.2828), n = 8[a]	37.8°C (SD 0.8485), n = 8[w]	SMD -1.64 (-2.82 to -0.47)	-	Very low
1 (Autret-Leca et al., 2007)	37.4(SD 0.9), n = 151 [l]	37.4(SD 1.0), n = 150[s]	SMD 0.00 (-0.23 to +0.23)		Very low
5 hours					
1 (Walson et al., 1989)	99.8°F (SD 1.9) ,n = 29[g]	101.3°F (SD 1.3), n = 31[b]	SMD -0.43 (-0.94 to +0.09)	-	Moderate
1 (Walson et al., 1989)	99.3°F (SD 1.7) n = 25 [w]	101.3°F (SD 1.3) n = 31[b]	SMD -0.79 (-1.34 to -0.25)	-	Moderate
1 (Nadal et al., 2002)	37.88°C (SD 1.07), n = 100[o]	37.85°C (SD 0.87), n = 99[p]	SMD 0.03 (-0.25 to +0.31)	-	Very low
1 (Kauffman et al., 1992)	37.0°C (SD 0.6928), n = 12[a]	38.1°C (SD 0.5657), n = 8[b]	SMD -1.63 (-2.69 to -0.57)	-	Very low
1 (Kauffman et al., 1992)	36.9°C (SD 0.5657), n = 8 [l]	38.1°C (SD 0.5657), n = 8[b]	SMD -2.01 (-3.27 to -0.74)	-	Very low
1 (Autret-Leca et al., 2007)	37.4 (SD 0.9), n = 151 [l]	37.6 (SD 1.0), n = 150[s]	SMD -0.21 (-0.44 to +0.02)	-	Very low

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
6 hours					
1 (Vauzelle-Kervrodan et al., 1997)	38°C (SD 0.8), n = 56 [l]	38°C (SD 0.8), n = 55[b]	SMD 0.00 (-0.37 to +0.37)	-	Low
1 (Van Esch et al., 1995)	37.82°C (SD 1.2828), n = 34[g]	38.23°C (SD 1.3015) ,n = 35 [b]	SMD -0.31 (-0.79 to +0.16)	-	Very low
1 (Nadal et al., 2002)	37.87°C (SD 0.96), n = 100[o]	38.10°C (SD 0.97), n = 99[p]	SMD -0.24 (-0.52 to +0.04)	-	Very low
1 (Walson et al., 1989)	100.2°F (SD 2.2),n = 29[g]	101.2°F (SD 1.9), n = 31[b]	SMD -0.29 (-0.80 to +0.22)	-	Low
1 (Walson et al., 1989)	99.7°F (SD 1.9) , n = 25 [l]	101.2°F (SD 1.9), n = 31[b]	SMD -0.57 (-1.11 to -0.03)	-	Low
1 (Kauffman et al., 1992)	37.3°C (SD 0.6928), n = 12[a]	38.5°C (SD 1.1314), n = 8[b]	SMD -1.29 (-2.29 to -0.29)	-	Very low
1 (Kauffman et al., 1992)	37.2°C (SD 0.5657), n = 8 [l]	38.5°C (SD 1.1314), n = 8[b]	SMD -1.37 (-2.50 to -0.25)	-	Very low
1 (Autret-Leca et al., 2007)	37.5°C (SD 0.9), n = 151 [l]	37.7°C (SD 1.0), n = 150[r]	SMD -0.21 (-0.44 to +0.02)	-	Very low
8 hours					
1 (Nadal et al., 2002)	38.0°C (SD 1.33), n = 100[o]	38.2°C (SD 0.84), n = 99[p]	SMD -0.18 (-0.46 to +0.10)	-	Very low
1 (Walson et al., 1989)	101.2°F (SD 2.0), n = 29[g]	101.2°F (SD 1.8), n = 31[b]	SMD -0.21 (-0.72 to +0.30)	-	Low
1 (Walson et al., 1989)	100.6°F (SD 2.2) n = 25 [l]	101.2°F (SD 1.8), n = 31[b]	SMD -0.50 (-1.03 to +0.04)	-	Low
1 (Kauffman et al., 1992)	37.7°C (SD 0.8485), n = 8[a]	38.8°C (SD 0.8485), n = 8[b]	SMD -1.23 (-2.32 to -0.13)	-	Very low
1 (Kauffman et al., 1992)	37.9°C (SD 1.3856), n = 12 [l]	38.8°C (SD 0.8485), n = 8[b]	SMD -0.72 (-1.64 to +0.21)	-	Very low
1 (Autret-Leca et al., 2007)	37.6°C (SD 0.9), n = 151 [l]	37.6°C (SD 0.95), n = 150[r]	SMD 0.00 (-0.23 to +0.23)	-	Very low
12 hours					
1 (Van Esch et al., 1995)	37.87°C (SD 1.3576), n = 32[e]	37.88°C (SD 1.1241), n = 35[b]	SMD -0.01 (-0.49 to +0.47)[b]	-	Low
24 hours					
1 (Van Esch et al., 1995)	37.92°C (SD 1.1432), n = 27[g]	38.18°C (SD 1.2638), n = 33[b]	SMD -0.21 (-0.72 to +0.30)	-	Very low
1 (Sarrell et al., 2006)	40.60°C (SD 1.46)n = 155[g]	40.55°C (SD 1.31), n = 154 [h]	SMD 0.04 (-0.19 to +0.26)	-	High

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Mean temperature at > 24 hours					
Day 2					
1 (Sarrell et al., 2006)	39.66°C (SD 1.48), n = 155[g]	39.74°C (SD 1.37), n = 154[h]	SMD -0.06 (-0.28 to +0.17)	-	High
Day 3					
1 (Sarrell et al., 2006)	39.64°C (SD 1.46), n = 155[g]	39.34°C (SD 1.19) n = 154 [h]	SMD 0.22 (0.00 to +0.45)	-	High
Afebrile at 1 to 2 hours					
1 hour					
1 (Autret et al., 1997)	33 of 116[a]	25 of 113[b]	RR 1.29 (0.82, 2.02)	-	Low
2 hours					
1 (Wong et al., 2001)	145 of 185[i]	130 of 191[j]	RR 1.15 (1.02, 1.30)	-	Moderate
1 (Van Esch et al., 1995)	27 of 30[g]	22 of 29 [b]	RR 1.19 (0.94, 1.50)[b]	-	Very low
Afebrile at > 2 to 5 hours					
4 hours					
1 (Autret et al., 1997)	69 of 116[a]	45 of 113[b]	RR 1.49 (1.14, 1.96)	-	Low
1 (Van Esch et al., 1995)	26 of 30[g]	22 of 29[b]	RR 1.18 (0.90, 1.55)	-	Very low
1 (Vauzelle-Kervroedan et al., 1997)	56 of 58 [l]	53 of 55[b]	RR 1.00 (0.93, 1.08)	-	Moderate
Afebrile at > 5 to 24 hours					
6 hours					
1 (Autret et al., 1997)	43 of 116[a]	40 of 113[b]	RR 1.05 (0.74, 1.48)	-	Low
1 (Van Esch et al., 1995)	20 of 34[g]	18 of 35[b]	RR 1.14 (0.75, 1.75)	-	Very low
12 hours					
1 (Van Esch et al., 1995)	21 of 34[g]	24 of 35[b]	RR 0.96 (0.68, 1.34)	-	Very low
24 hours					
1 (Van Esch et al., 1995)	20 of 34[g]	20 of 35[b]	RR 1.22 (0.86, 1.74)	-	Very low
Afebrile at > 24 hours					
No data					

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Temperature AUC					
0 to 8 hours					
1 (Kauffman et al., 1992)	730 (576 to 839)[a]	328 (-356 to 686)[b]	p = 0.05	-	Very low
1 (Kauffman et al., 1992)	590 (160 to 875) [l]	328 (-356 to 686)[b]	p = 0.05	-	Very low
0 to 6 hours					
1 (Wilson et al., 1991)	7.09 (SEM 0.58), n = 43[g]	6.72 (SEM 0.58), n = 51[h]	NS	-	Very Low
1 (Wilson et al., 1991)	4.91 (SEM 0.47), n = 47 [l]	6.72 (SEM 0.58), n = 51[h]	NS	-	Very low
0 to 8 hours					
1 (Walson et al., 1989)	460.9 [g]	365.0 [b]	NS	-	Low
1 (Walson et al., 1989)	510.9 [l]	365.0 [b]	p < 0.05	-	Low
0 to 6 hours – total temperature change per hour					
1 (Walson et al., 1992)	297[g]	377[b]	NS	-	Moderate
1 (Walson et al., 1992)	385 [l]	377[b]	NS	-	Moderate
0 to 12 hours – total temperature change per hour					
1 (Walson et al., 1992)	689 [g]	938[n]	p<0.05	-	Moderate
1 (Walson et al., 1992)	929 [l]	938[b]	NS	-	Moderate
0 to 24 hours – total temperature change per hour					
1 (Walson et al., 1992)	1572 [g]	2100[b]	p<0.05	-	Moderate
1 (Walson et al., 1992)	1995 [l]	2100[b]	NS	-	Moderate
0 to 48 hours – total temperature change per hour					
1 (Walson et al., 1992)	3286 [g]	4400[b]	NS	-	Moderate
1 (Walson et al., 1992)	3933 [l]	4400[b]	NS	-	Moderate

Number of studies	Number of children		Effect*		Quality
	Ibuprofen	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Adverse events					
5 (Southey et al., 2009; Pierce et al, 2010; Kauffman et al., 1992; Sarrell et al., 2006; and Walson et al., 1989)	2962 of 21843	1469 of 11678	RR 1.04 (0.98 to 1.10)	-	Very low
Discontinuation of treatment					
1 (Southey et al., 2009)	5 of 257	5 of 226	RR 0.54 (0.17 to 1.71)	-	Very low

AUC area under the curve, NC non-calculable, NR not reported, NS Not significant at $P < 0.05$, P probability, RR relative risk, SD standard deviation, SMD standard mean difference

* Relative and absolute differences are calculated by the NCC technical team based on the data presented in the papers. When this data is unavailable the authors reported figures may be used.

[a] Ibuprofen at 7.5 mg/kg

[b] Paracetamol at 10 mg/kg

[c] Not presented in correct format for analysis of categorical data

[d] 6.67 mg/kg of Ibuprofen

[e] Study used a non-validated scoring system

[f] 10.65 mg/kg of paracetamol

[g] 5 mg/kg of Ibuprofen

[h] 12.5 mg/kg of paracetamol

[i] At 5 mg/kg for initial temp <39.2°C and 10 mg/kg for initial temp ≥39.2°C

[j] 12 mg/kg of paracetamol. The dose of paracetamol was adjusted according to each patient's age following package insert instructions and averaged 12mg/kg

[k] 15.3 mg/kg paracetamol

[l] 10 mg/kg Ibuprofen

[m] At 20 mg/kg in 24 hours

[n] At 50 mg/kg in 24 hours

[o] 6.67 mg/kg of Ibuprofen

[p] 10.65 mg/kg of paracetamol

[q] Included children aged more than 5 (Nadal = 12 ; Wong Included children up to 6 years; Ulukol up to 14 years; McIntyre up to aged 12; Kaufmann up to 12; Vauzelle up to 12; Erlewyn; Autret-Leca 12)

[r] 1.96h paracetamol; 2.16h ibuprofen

[s] A crossover analysis comparing the study drugs was performed on 22 children with a second episode of fever.

Table 9.5 GRADE findings for paracetamol vs. paracetamol and ibuprofen combined

Number of studies	Number of children		Effect		Quality
	Combined	**Mono**	**Relative (95% confidence interval)**	**Absolute (95% confidence interval)**	
Discomfort at 1 to 2 hours					
No data					
Discomfort at > 2 to 5 hours					
No data					
Discomfort at > 5 to 24 hours					
24 hours					
Discomfort 1 (Hay et al., 2009)	29 of 50 [a]	22 of 52[b]	RR 1.37 (0.92 to 2.04)	-	Low
Activity 1 (Hay et al., 2009)	23 of 48 [a]	20 of 50[b]	RR 1.20 (0.76 to 1.88)	-	Very low
Appetite 1 (Hay et al., 2009)	14 of 48[a]	10 of 48[b]	RR 1.40 (0.69 to 2.84)	-	Very low
Sleep 1 (Hay et al., 2009)	20 of 52[a]	17 of 46[b]	RR 1.04 (0.62 to 1.73)	-	Very low
Discomfort at > 24 hours					
48 hours					
Discomfort 1 (Hay et al., 2009)	36 of 52[a]	34 of 52[b]	RR 1.06 (0.81 to 1.38)	-	Low
Activity 1 (Hay et al., 2009)	28 of 52[a]	31 of 52[b]	RR 0.90 (0.65 to 1.26)	-	Very low
Appetite 1 (Hay et al., 2009)	21 of 51[a]	21 of 51[b]	RR 1.00 (0.63 to 1.59)	-	Very low
Sleep 1 (Hay et al., 2009)	25 of 52[a]	27 of 52[b]	RR 0.93 (0.63 to 1.36)	-	Very low
Day 5					
Discomfort 1 (Hay et al., 2009)	38 of 50[a]	43 of 49[b]	RR 0.87 (0.72 to 1.04)	-	Moderate
Activity 1 (Hay et al., 2009)	37 of 51[a]	44 of 49[b]	RR 0.81 (0.67 to 0.98)	-	Moderate
Appetite 1 (Hay et al., 2009)	32 of 52[a]	29 of 50[b]	RR 1.06 (0.77 to 1.46)	-	Low

Number of studies	Number of children		Effect		Quality
	Combined	Mono	Relative (95% confidence interval)	Absolute (95% confidence interval)	
S*leep* 1 (Hay et al., 2009)	27 of 51[a]	31 of 50[b]	RR 0.85 (0.61 to 1.20)	-	Low
Mean change temperature at 1 to 2 hours					
1 hour					
1 (Erlewyn-Lajeunesse et al., 2006)	-1.22 (0.95 to 1.50), n = 36[c]	-0.95 (0.72 to 1.17), n = 37 [d]	RR 0.36 (-0.10 to 0.82)	-	Moderate
Mean change in temperature at > 2 to 5 hours					
No data					
Mean change in temperature at > 5 to 24 hours					
No data					
Mean change in temperature at > 24 hours					
No data					
Mean temperature at 1 to 2 hours					
1 hour					
1 (Erlewyn-Lajeunesse et al., 2006)	37.59°C (SD 0.61), n = 36[c]	37.98°C (SD 0.47), n = 37[d]	RR -0.71 (-1.18 to -0.24) Adjusted 0.35C (0.10 to 0.6), *P* = 0.028.	-	Moderate
Mean temperature at > 2 to 5 hours					
No data					
Mean temperature at > 5 to 24 hours					
1 (Hay et al., 2009)	36.6°C (SD 1.01), n = 52[a]	36.4°C (SD 0.89), n = 52[b]	SMD 0.21 (-0.18 to +0.59)	-	Low
Mean temperature at > 24 hours					
1 (Hay et al., 2009)	36.0°C (SD 0.66), n = 52[a]	36.2°C (SD 0.93), n = 52[b]	SMD -0.25 (-0.63 to +0.14)	-	Low
Afebrile at 1 to 2 hours					
1 (Hay et al., 2009)	47 of 52[a]	33 of 52[b]	RR 1.42 (1.14 to 1.78)	-	Moderate
Afebrile at > 2 to 5 hours					
1 (Hay et al., 2009)	51 of 52[a]	37 of 52[b]	RR 1.38 (1.15 to 1.65)	-	Moderate

Number of studies	Number of children		Effect		Quality
	Combined	Mono	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Afebrile at > 5 to 24 hours					
6 hours					
1 (Hay et al., 2009)	47 of 52[a]	39 of 52[b]	RR 1.21 (1.01 to 1.44)	-	Moderate
8 hours					
1 (Hay et al., 2009)	45 of 52[a]	42 of 52[b]	RR 1.07 (0.90 to 1.27)	-	Moderate
12 hours					
1 (Hay et al., 2009)	49 of 52[a]	39 of 52[b]	RR 1.26 (1.06 to 1.49)	-	Moderate
Afebrile at > 24 hours					
1 (Hay et al., 2009)	47 of 52[a]	46 of 52[b]	RR 1.02 (0.90 to 1.17)	-	Moderate
Time without fever					
0 to 4 hours					
1 (Hay et al., 2009)	116.2 (SD 65.0)[a]	171.1 (SD 40.8) [b]	Adjusted mean difference 55.3 (33.1 to 77.5), *P* < 0.001	-	Moderate
24 hours					
1 (Hay et al., 2009)	1217.4 (SD 237.6) [a]	940.3 (SD 362.9) [b]	Adjusted mean difference 4.4 (2.4 to 6.3), *P* < 0.001	-	Moderate
Adverse events					
Diarrhoea					
1 (Hay et al., 2009)	12 of 52[a]	10 of 52[b]	RR 1.20 (0.57 to 2.53)	-	Low
Vomiting					
1 (Hay et al., 2009)	2 of 52[a]	6 of 52	RR 0.33 (0.07 to 1.58)	-	Low

NC non-calculable, NR not reported, P probability, RR relative risk, SD standard deviation

[a] Hay – 15 mg/kg paracetamol + 10 mg/kg ibuprofen

[b] 15 mg/kg paracetamol

[c] 15 mg/kg + 5 mg/kg

[d] 15 mg/kg paracetamol

Table 9.6 GRADE findings for paracetamol vs. paracetamol and ibuprofen alternating

Number of studies	Number of children		Effect		Quality
	Alternating	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Discomfort at 1 to 2 hours					
No data					
Discomfort at > 2 to 5 hours					
No data					
Discomfort at > 5 to 24 hours					
Day 1					
1 (Sarrell et al., 2006)	9.26 (SD 2.49), n = 155[a]	11.77 (SD 2.64), n = 154[b]	SMD -0.98 (-1.21 to -0.74)	-	High
Discomfort > 24 hours					
Day 2					
1 (Sarrell et al., 2006)	5.09 (SD 2.78), n = 155[a]	8.87 (SD 2.54), n = 154[b]	SMD -1.42 (-1.67 to -1.17)	-	High
Day 3					
1 (Sarrell et al., 2006)	4.18 (SD 2.74), n = 155[a]	7.66 (SD 2.96), n = 154[b]	SMD -1.22 (-1.46 to -0.97)	-	High
Mean change temperature at 1 to 2 hours					
No data					
Mean change in temperature at > 2 to 5 hours					
No data					
Mean change in temperature at > 5 to 24 hours					
No data					
Mean change in temperature at > 24 hours					
No data					
Mean temperature at 1 to 2 hours					
1 (Pashapour et al., 2009)	38.8°C (SD 0.59), n = 35[c]	38.8°C (SD 0.47), n = 35[d]	SMD 0.00 (-0.47 to +0.47)	-	Low
Mean temperature at > 2 to 5 hours					
3 hours					
1 (Kramer et al., 2008)	37.7°C (SD 0.6224), n = 19[e]	37.7°C (SD 0.415), n = 19[d]	SMD 0.00 (-0.64 to +0.64)	-	Very low
4 hours					
1 (Pashapour et al., 2009)	38.4°C (SD 0.34), n = 35[c]	38.5°C (SD 0.3), n = 35[d]	SMD -0.31 (-0.78 to +0.16)	-	Very low
1 (Kramer et al., 2008)	37.4°C (SD 0.8299), n = 19[e]	38.0°C (SD 1.0374), n = 19[d]	SMD -0.63 (-1.28 to +0.03)	-	Very low

Number of studies	Number of children		Effect		Quality
	Alternating	Paracetamol	Relative (95% confidence interval)	Absolute (95% confidence interval)	
5 hours					
1 (Pashapour et al., 2009)	38.0°C (SD 0.47), n = 35[c]	38.2°C (SD 0.38), n = 35[d]	SMD -0.46 (-0.94 to +0.01)	-	Very low
1 (Kramer et al., 2008)	37.1°C (SD 0.6224), n = 19[e]	37.9°C (SD 0.8299), n = 19[d]	SMD -1.07 (-1.75 to -0.38)	-	Very low
Mean temperature at > 5 to 24 hours					
6 hours					
1 (Kramer et al., 2008)	37.4°C (SD 0.8299), n = 19[e]	37.5°C (SD 0.8299), n = 19[d]	SMD -0.12 (-0.75 to +0.52)	-	
7 hours					
1 (Pashapour et al., 2009)	38.0°C (SD 0.48), n = 35[c]	38.2°C (SD 0.57), n = 35[d]	SMD -0.38 (-0.85 to +0.10)	-	Moderate
8 hours					
1 (Pashapour et al., 2009)	37.7°C (SD 0.46), n = 35[c]	38.0°C (SD 0.52), n = 35[d]	SMD -0.60 (-1.08 to -0.12)	-	Moderate
Mean temperature at > 24 hours					
Day 1					
1 (Sarrell et al., 2006)	39.64°C (SD 1.17), n = 155[a]	40.55°C (SD 1.31), n = 155[b]	SMD -0.73 (-0.96 to -0.50)	-	High
Day 2					
1 (Sarrell et al., 2006)	38.78°C (SD 0.87), n = 155[a]	39.74°C (SD 1.37), n = 155[b]	SMD -0.83 (-1.07 to -0.60)	-	High
Day 3					
1 (Sarrell et al., 2006)	38.54°C (SD 0.74), n = 155[a]	39.34°C (SD 1.19), n = 155[b]	SMD -0.81 (-1.04 to -0.57)	-	High
Afebrile at 1 to 2 hours					
No data					
Afebrile at > 2 to 5 hours					
No data					
Afebrile at > 5 to 24 hours					
No data					
Afebrile at > 24 hours					
No data					
Temperature AUC					
No data					
Adverse events					
No reported					

AUC area under the curve, NC non-calculable, NR not reported, NS Not significant at $P < 0.05$, P probability, SD standard deviation, SMD standard mean difference

[a] Alternating acetaminophen (12.5 mg/kg) with ibuprofen (5 mg/kg) every 4 hours

[b] Acetaminophen (12.5 mg/kg) every 6 hours

[c] Alternating ibuprofen (10 mg/kg) with acetaminophen (15 mg/kg) every 4 hours

[d] Acetaminophen (15 mg/kg) every 4 hours

[e] Alternating acetaminophen (15 mg/kg) with ibuprofen (10 mg/kg) with) every 3 hours

Table 9.7 GRADE findings for ibuprofen vs. paracetamol and ibuprofen combined

Number of studies	**Number of children**		**Effect**		**Quality**
	Combined	**Mono**	**Relative (95% confidence interval)**	**Absolute (95% confidence interval)**	
Discomfort at 1 to 2 hours					
No data					
Discomfort at > 2 to 5 hours					
No data					
Discomfort at > 5 to 24 hours					
24 hours					
Discomfort 1 (Hay et al., 2009)	29 of 50[a]	36 of 52[b]	RR 0.84 (0.62 to 1.13)	-	Low
Activity 1 (Hay et al., 2009)	23 of 48[a]	20 of 34[b]	RR 0.81 (0.54 to 1.22)	-	Very low
Appetite 1 (Hay et al., 2009)	14 of 48[a]	14 of 52[b]	RR 1.08 (0.58 to 2.03)	-	Very low
Sleep 1 (Hay et al., 2009)	20 of 52[a]	13 of 26[b]	RR 0.77 (0.46 to 1.29)	-	Very low
Discomfort > 24 hours					
48 hours					
Comfort 1 (Hay et al., 2009)	36 of 52[a]	37 of 52[b]	RR 0.97 (0.76 to 1.25) Adjusted OR 0.89 (0.32 to 2.43)	-	Moderate
Activity 1 (Hay et al., 2009)	28 of 52[a]	37 of 51[b]	RR 0.74 (0.55 to 1.00)	-	Moderate
Appetite 1 (Hay et al., 2009)	21 of 51[a]	22 of 50[b]	RR 0.94 (0.59 to 1.47)	-	Very low
Sleep 1 (Hay et al., 2009)	25 of 52[a]	31 of 51[b]	RR 0.79 (0.55 to 1.13)	-	Low
Day 5					
Comfort 1 (Hay et al., 2009)	38 of 50[a]	38 of 47[b]	RR 0.94 (0.76 to 1.16)	-	Moderate

Number of studies	Number of children		Effect		Quality
	Combined	Mono	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Activity 1 (Hay et al., 2009)	37 of 51[a]	39 of 46[b]	RR 0.86 (0.69 to 1.05)	-	Low
Appetite 1 (Hay et al., 2009)	32 of 52[a]	29 of 49[b]	RR 1.04 (0.76 to 1.43)	-	Low
Sleep 1 (Hay et al., 2009)	27 of 51[a]	25 of 50[b]	RR 1.06 (0.72 to 1.55)	-	Low
Mean change temperature at 1 to 2 hours					
1 hour					
1 (Erlewyn-Lajeunesse et al., 2006)	-1.22 (0.95 to 1.50) n = 36[c]	-0.92 (0.70 to 1.14), n = 35[d]	SMD -0.33 (-0.80 to +0.13)	-	Moderate
Mean change in temperature at > 2 to 5 hours					
No data					
Mean change in temperature at > 5 to 24 hours					
No data					
Mean change in temperature at > 24 hours					
No data					
Mean temperature at 1 to 2 hours					
1 hour					
1 (Erlewyn-Lajeunesse et al., 2006)	37.59°C (SD 0.61) [c]	37.81°C (SD 0.69) [d]	SMD -0.33 (-0.80 to +0.13) Adjusted MD = 0.25C (-0.01 to 0.50), *P* = 0.166	-	Moderate
1 (Paul et al., 2010)	37.4°C (SD 0.5), n = 20[e]	37.6°C (SD 0.5), n = 20[f]	SMD -0.39 (-1.02 to +0.23)	-	
2 hours					
1 (Paul et al., 2010)	37.0°C (SD 0.5), n = 20[e]	37.1°C (SD 0.4), n = 20[f]	SMD -0.22 (-0.84 to +0.41)	-	
Mean temperature at > 2 to 5 hours					
3 hours					
1 (Paul et al., 2010)	36.9°C (SD 0.4), n = 20[e]	37.2°C (SD 0.6), n = 20[f]	SMD -0.58 (-1.21 to +0.06)	-	

Number of studies	Number of children		Effect		Quality
	Combined	Mono	Relative (95% confidence interval)	Absolute (95% confidence interval)	
4 hours					
1 (Paul et al., 2010)	36.9°C (SD 0.3), n = 20[e]	37.5°C (SD 1.1), n = 20[f]	SMD -0.73 (-1.37 to -0.09)	-	
5 hours					
1 (Paul et al., 2010)	36.9°C (SD 0.5), n = 20[e]	38.0°C (SD 1.1), n = 20[f]	SMD -1.26 (-1.95 to -0.58)	-	
Mean temperature at > 5 to 24 hours					
6 hours					
1 (Paul et al., 2010)	37.2°C (SD 0.6), n = 20[e]	38.5°C (SD 1.5), n = 20[f]	SMD -1.12 (-1.79 to -0.44)	-	
24 hours					
1 (Hay et al., 2009)	36.6°C (SD 1.01), n = 52[a]	36.4°C (SD 0.85), n = 52[b]	SMD 0.21 (-0.17 to +0.60)	-	Moderate
Mean temperature at > 24 hours					
1 (Hay et al., 2009)	36.0°C (SD 0.66), n = 52[a]	36.1°C (SD 0.78), n = 52[b]	SMD -0.14 (-0.52 to +0.25)	-	Moderate
Afebrile at 1 to 2 hours					
1 hour					
1 (Paul et al., 2010)	18 of 20[e]	16 of 20[f]	RR 1.13 (0.86 to 1.46)	-	Very low
2 hours					
1 (Hay et al., 2009)	47 of 52[a]	44 of 52[b]	SMD 1.07 (+0.92 to +1.24)	-	Moderate
1 (Paul et al., 2010)	20 of 20[fl]	19 of 20[f]	RR 1.05 (0.92 to 1.20)	-	Low
Afebrile at > 2 to 5 hours					
3 hours					
1 (Paul et al., 2010)	20 of 20[e]	18 of 20[f]	RR 1.11 (0.93 to1.31)	-	Very low
4 hours					
1 (Hay et al., 2009)	51 of 52[el]	44 of 52[f]	RR 1.16 (1.03 to 1.31)	-	Moderate
1 (Paul et al., 2010)	20 of 20[e]	14 of 20[f]	RR 1.41 (1.05 to 1.90)	-	Low
5 hours					
1 (Paul et al., 2010)	20 of 20[e]	12 of 20[f]	RR 1.64 (1.15 to 2.35)	-	Low

Number of studies	Number of children		Effect		Quality
	Combined	Mono	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Afebrile at > 5 to 24 hours					
6 hours					
1 (Hay et al., 2009)	47 of 52[a]	37 of 52[b]	RR 1.27 (1.05 to 1.54)	-	Moderate
1 (Paul et al., 2010)	19 of 20[e]	10 of 20[f]	RR 1.90 (1.21 to 2.98)	-	Low
8 hours					
1 (Hay et al., 2009)	45 of 52[a]	46 of 52[b]	RR 0.98 (0.85 to 1.13)	-	Moderate
12 hours					
1 (Hay et al., 2009)	49 of 52	47 of 52	RR 1.04 (0.93 to 1.17)	-	Moderate
Afebrile at > 24 hours					
24 hours					
1 (Hay et al., 2009)	47 of 52[a]	45 of 52[b]	RR 1.04 (0.91 to 1.20)	-	Moderate
Time without fever					
4 hours					
1 (Hay et al., 2009)	171.1 (40.8)	156.0 (57.6)	adjusted mean difference 16.2 (-7.0 to 39.4), *P* = 0.2	-	Moderate
24 hours					
1 (Hay et al., 2009)	1217.4 (237.6)	1055.2 (329.7)	adjusted mean difference 2.5 (0.6 to 4.4), *P* = 0.008	-	Moderate
Adverse events					
Diarrhoea					
1 (Hay et al., 2009)	12 of 52	9 of 52	RR 0.75 (0.35 to 1.63)	-	Very low
Vomiting					
1 (Hay et al., 2009)	2 of 52	3 of 52	RR 1.50 (0.26 to 8.61)	-	Very low

NC non-calculable, NR not reported, OP odds ratio, RR relative risk, SD standard deviation, SMD standard mean difference

[a] Hay – 15 mg/kg paracetamol + 10 mg/kg ibuprofen

[b] 10 mg/kg ibuprofen

[c] 15 mg/kg + 5 mg/kg

[d] 5 mg/kg ibuprofen

[e] 10 mg/kg Ibuprofen and 15 mg/kg acetaminophen single dose

[f] 10 mg/kg Ibuprofen single dose

Table 9.8 GRADE findings for ibuprofen vs. paracetamol and ibuprofen alternating

Number of studies	Number of children		Effect		Quality
	Intervention	**Control**	**Relative (95% confidence interval)**	**Absolute (95% confidence interval)**	
Discomfort at 1 to 2 hours					
No data					
Discomfort at > 2 to 5 hours					
No data					
Discomfort at > 5 to 24 hours					
Day 1					
NCCPC score 1 (Sarrell et al., 2006)	9.26 (SD 2.49), n = 155[a]	11.48 (SD 2.58), n = 155[b]	SMD -0.87 (-1.11 to -0.64)	-	
Discomfort at > 24 hours					
Day 2					
NCCPC score 1 (Sarrell et al., 2006)	5.09 (SD 2.78), n = 155[a]	8.83 (SD 2.67), n = 155[b]	SMD -1.37 (-1.62 to -1.12)	-	High
Day 3					
NCCPC score 1 (Sarrell et al., 2006)	4.18 (SD 2.74), n = 155[a]	7.96 (SD 2.71), n = 155[b]	SMD -1.38 (-1.63 to -1.14)	-	High
Mean change temperature at 1 to 2 hours					
No data					
Mean change in temperature at > 2 to 5 hours					
No data					
Mean change in temperature at > 5 to 24 hours					
No data					
Mean change in temperature at > 24 hours					
No data					
Mean temperature at 1 to 2 hours					
1 hour					
1 (Paul et al., 2010)	37.6°C (SD 0.4), n = 20[c]	37.6°C (SD 0.5), n = 20[c]	SMD 0.00 (-0.62 to +0.62)	-	Very low
2 hours					
1 (Paul et al., 2010)	37.2°C (SD 0.3), n = 20[c]	37.1°C (SD 0.4), n = 20[d]	SMD 0.28 (-0.35 to +0.90)	-	Very low

Number of studies	Number of children		Effect		Quality
	Intervention	Control	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Mean temperature at > 2 to 5 hours					
3 hours					
1 (Paul et al., 2010)	36.9°C (SD 0.4), n = 20[c]	37.2°C (SD 0.6), n = 20[d]	SMD -0.58 (-1.21 to +0.06)	-	Low
4 hours					
1 (Nabulsi et al., 2006)	37.5°C (SD 0.7), n = 37[e]	37.7°C (SD 0.9), n = 33[f]	SMD -0.25 (-0.72 to +0.22)	-	Very low
1 (Paul et al., 2010)	36.9°C (SD 0.3), n = 20[c]	37.5°C (SD 1.1), n = 20[d]	SMD -0.73 (-1.37 to -0.09)	-	Moderate
5 hours					
1 (Paul et al., 2010)	36.8°C (SD 0.3), n = 20[c]	38.0°C (SD 1.1), n = 20[d]	SMD -1.46 (-2.16 to -0.75)	-	Moderate
Mean temperature at > 5 to 24 hours					
6 hours					
1 (Paul et al., 2010)	36.9°C (SD 0.3), n = 20[c]	38.5°C (SD 1.5), n = 20[d]	SMD -1.45 (-2.15 to -0.75)	-	Low
Mean temperature at > 24 hours					
Day 1					
1 (Sarrell et al., 2006)	39.64°C (SD 1.17), n = 155[a]	40.6°C (SD 1.46), n = 155[b]	SMD -0.72 (-0.95 to -0.49)	-	High
Day 2					
1 (Sarrell et al., 2006)	38.78°C (SD 0.87), n = 155[a]	39.66°C (SD 1.48), n = 155[b]	SMD -0.72 (-0.95 to -0.49)	-	High
Day 3					
1 (Sarrell et al., 2006)	38.54°C (SD 0.74), n =155[a]	39.64°C (SD 1.46), n = 155[b]	SMD -0.95 (-1.18 to -0.71)	-	High
Afebrile at 1 to 2 hours					
1 hour					
1 (Paul et al., 2010)	16 of 20[c]	16 of 20[d]	RR 1.00 (0.73, to1.36)	-	Very low
2 hours					
1 (Paul et al., 2010)	20 of 20[c]	19 of 20[d]	RR 1.05 (0.92 to 1.20)	-	Moderate

Number of studies	Number of children		Effect		Quality
	Intervention	Control	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Afebrile at > 2 to 5 hours					
3 hours					
1 (Paul et al., 2010)	20 of 20[c]	18 of 20[d]	RR 1.11 (0.93 to 1.31)	-	Low
4 hours					
1 (Paul et al., 2010)	20 of 20[c]	14 of 20[d]	RR 1.41 (1.05 to 1.90)	-	Moderate
5 hours					
1 (Paul et al., 2010)	20 of 20[c]	12 of 20[d]	RR 1.64 (1.15 to 2.35)	-	Moderate
Afebrile at > 5 to 24 hours					
6 hours					
1 (Paul et al., 2010)	20 of 36[c]	10 of 33[d]	RR 1.95 (1.27 to 3.01)	-	Moderate
1 (Nabulsi et al., 2006)	30 of 36[e]	19 of 33[f]	RR 1.62 (1.25 to 2.11)	-	Low
7 hours					
1 (Nabulsi et al., 2006)	31 of 36[e]	14 of 33[f]	RR 2.03 (1.34 to 3.08)	-	Low
8 hours					
1 (Nabulsi et al., 2006)	29 of 36[e]	11 of 33[f]	RR 2.42 (1.45 to 4.02)	-	Low
Afebrile at > 24 hours					
No data					
Temperature AUC					
No data					
Adverse events					
Diarrhoea					
1 Nabulsi et al., 2006	5 of 37[e]	6 of 37[f]	RR 0.83 (0.28 to 2.49)	-	Very low

NR Not reported, RR relative risk, SD standard deviation, SMD standard mean difference

[a] Alternating acetaminophen (12.5 mg/kg) with ibuprofen (5 mg/kg) every 4 hours

[b] Ibuprofen (5 mg/kg) every 6 hours

[c] 10 mg/kg Ibuprofen and 15 mg/kg acetaminophen single dose

[d] 10 mg/kg Ibuprofen single dose

[e] Ibuprofen 10 mg/kg, followed by acetaminophen 15mg/kg at 4h

[f] Ibuprofen 10 mg/kg, followed by placebo at 4h

Table 9.9 GRADE findings for paracetamol and ibuprofen combined vs. paracetamol and ibuprofen alternating

Number of studies	Number of children		Effect		Quality
	Intervention	Control	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Discomfort at 1 to 2 hours					
No data					
Discomfort at > 2 to 5 hours					
No data					
Discomfort at > 5 to 24 hours					
No data					
Discomfort > 24 hours					
No data					
Mean change temperature at 1 to 2 hours					
No data					
Mean change in temperature at > 2 to 5 hours					
No data					
Mean change in temperature at > 5 to 24 hours					
No data					
Mean change in temperature at > 24 hours					
No data					
Mean temperature at 1 to 2 hours					
1 hour					
1 (Paul et al., 2010)	37.4°C (SD 0.5), n = 20[a]	37.6°C (SD 0.4), n = 20[b]	SMD -0.43 (-1.06 to +0.19)	-	Low
2 hours					
1 (Paul et al., 2010)	37.0°C (SD 0.5), n = 20[a]	37.2°C (SD 0.3), n = 20[b]	SMD -0.48 (-1.10 to +0.15)	-	Low
Mean temperature at > 2 to 5 hours					
3 hours					
1 (Paul et al., 2010)	36.9°C (SD 0.4), n = 20[a]	36.9°C (SD 0.4), n = 20[b]	SMD 0.00 (-0.62 to +0.62)	-	Low
4 hours					
1 (Paul et al., 2010)	36.9°C (SD 0.3), n = 20[a]	36.9°C (SD 0.3), n = 20[b]	SMD 0.00 (-0.62 to +0.62)	-	Low
5 hours					
1 (Paul et al., 2010)	36.9°C (SD 0.5), n = 20[a]	36.8°C (SD 0.3), n = 20[b]	SMD 0.24 (-0.38 to +0.86)	-	Low

Number of studies	Number of children		Effect		Quality
	Intervention	Control	Relative (95% confidence interval)	Absolute (95% confidence interval)	
Mean temperature at > 5 to 24 hours					
6 hours					
1 (Paul et al., 2010)	37.2°C (SD 0.6), n = 20[a]	36.9°C (SD 0.3), n = 20[b]	SMD 0.62 (-0.02 to+1.26)	-	Low
Mean temperature at > 24 hours					
No data					
Afebrile at 1 to 2 hours					
1 hour					
1 (Paul et al., 2010)	18 of 20[a]	16 of 20[b]	RR 1.13 (0.86 to 1.46)	-	Low
2 hours					
1 (Paul et al., 2010)	20 of 20[a]	20 of 20[b]	RR 1.00 (0.91 to 1.10)	-	Moderate
Afebrile at > 2 to 5 hours					
3 hours					
1 (Paul et al., 2010)	20 of 20[a]	20 of 20[b]	RR 1.00 (0.91 to 1.10)	-	Moderate
4 hours					
1 (Paul et al., 2010)	20 of 20[a]	20 of 20[b]	RR 1.00 (0.91 to 1.10)	-	Moderate
5 hours					
1 (Paul et al., 2010)	20 of 20[a]	20 of 20[b]	RR 1.00 (0.91 to 1.10)	-	Moderate
Afebrile at > 5 to 24 hours					
6 hours					
1 (Paul et al., 2010)	19 of 20[a]	20 of 20[b]	RR 0.95 (0.83 to 1.09)	-	Moderate
Afebrile at > 24 hours					
No data					
Temperature AUC					
No data					
Adverse events					
No data					

AUC area under the curve, NR not reported, RR relative risk, SD standard deviation, SMD standard mean difference

[a] 10 mg/kg Ibuprofen and 15 mg/kg acetaminophen single dose

[b] 10 mg/kg Ibuprofen single dose

Evidence statements

A number of calculations have been used in this review. For a relative risk an effect size of 0.25 with the 95% confidence interval not crossing 1 (no effect) was considered a large effect. For standardised mean differences an effect size of 0.1 was considered small, 0.3 medium and 0.5 large.

Paracetamol compared with placebo

One RCT found that quality of life (comfort, activity, alertness, mood and appetite) was improved in children who received paracetamol compared with children who received placebo to treat fever. This finding was a moderate effect size and was statistically significant. The evidence for this finding was of low quality.

Four RCTs found that temperature was reduced more in children who received paracetamol compared with children who received placebo to treat fever. This finding was statistically significant in all the studies. The evidence for this finding ranged from high to very low quality.

A meta-analysis of seven RCTs found more adverse events reported in children who received paracetamol compared with children who received placebo to treat fever. This finding was not statistically significant. The evidence for this finding was of very low quality.

One observational study found that rates of asthma, eczema and rhinoconjunctivitis were higher in children who had used paracetamol in the first year of life or within the past 12 months compared with those who had not. This finding was statistically significant, but these kinds of studies are unable to demonstrate a causal relationship between paracetamol use and long-term conditions. The evidence for this finding was very low quality.

Ibuprofen compared with placebo

No data was found on quality of life.

Three RCTs found that temperature was reduced more in children who received ibuprofen compared with children who received placebo to treat fever. This was a large effect and the finding was statistically significant. The evidence for this finding ranged from moderate to low in quality.

A meta-analysis of seven RCTs found more adverse events reported in children who received ibuprofen compared with children who received placebo to treat fever. This finding was statistically significant. The evidence for this finding was of low quality.

Paracetamol compared with ibuprofen

Two RCTs found that quality of life was improved in children who received ibuprofen compared with children who received paracetamol to treat fever up until 6 hours after treatment, but not thereafter. This finding was a moderate effect size and was statistically significant. The evidence for this finding was of low quality.

One RCT found that there was no difference in quality of life in children who received ibuprofen compared with children who received paracetamol to treat fever from day 1 to 3 of treatment. This finding was statistically significant. The evidence for this finding was of low quality.

Ten RCTs found either no difference or moderate differences in favour of ibuprofen in temperature reduction between 1 and 6 hours after treatment began in children who received ibuprofen compared with children who received paracetamol to treat fever. The evidence for this finding ranged from high to very low in quality.

Three RCTs found that the proportion of afebrile patients was higher in the group of children who received ibuprofen compared with children who received paracetamol to treat fever up until 4 hours after treatment. This evidence was of moderate to very low quality.

A meta-analysis found no difference in the number of adverse events reported in children who received ibuprofen compared with children who received paracetamol to treat fever. The evidence for this finding was of very low quality.

Paracetamol compared with paracetamol and ibuprofen combined

One RCT found no difference in quality of life up to 5 days after treatment began between children who received paracetamol and ibuprofen combined compared with children who received paracetamol only to treat fever. The evidence for this finding was of low quality.

Two RCTs found no difference in temperature reduction in children who received paracetamol and ibuprofen combined compared with children who received paracetamol only to treat fever. The evidence for this finding was of moderate quality.

One RCT found that the proportion of afebrile patients was higher in children who received paracetamol and ibuprofen combined compared with children who received paracetamol alone to treat fever up until 6 hours after treatment, but no difference between groups thereafter. This was a large effect and the finding was statistically significant. The evidence for this finding was of moderate quality.

One RCT found that total time without fever was longer in children who received paracetamol and ibuprofen combined compared with children who received paracetamol to treat fever up until 24 hours after treatment began. This finding was statistically significant. The evidence for this finding was of moderate quality.

One RCT reported no difference in adverse events between paracetamol and ibuprofen combined and paracetamol alone.

Paracetamol compared with paracetamol and ibuprofen alternating

One RCT found quality of life (discomfort) was less up to 3 days after treatment began in children who received alternating paracetamol and ibuprofen compared to children who received only paracetamol to treat fever. This finding was statistically significant. The evidence for this finding was of high quality.

Two RCTs found that temperature was reduced more in children who received alternating paracetamol and ibuprofen compared to children who received only paracetamol to treat fever. This finding was statistically significant at 5 hours, 8 hours and from 1 to 3 days after treatment began. The effect size was moderate to high. The evidence was of high to very low quality.

No adverse events were reported.

Ibuprofen compared with paracetamol and ibuprofen combined

One RCT found no difference in quality of life up to 5 days after treatment began in children who received paracetamol and ibuprofen combined compared with children who received only ibuprofen to treat fever. The evidence for this finding was of low quality.

Three RCTs found no difference in temperature reduction in children who received paracetamol and ibuprofen combined compared with children who received only ibuprofen to treat fever. The evidence for this finding was of moderate quality.

Two RCTs found no difference in the proportion of children who were afebrile when comparing children who received ibuprofen and paracetamol combined with children who received only ibuprofen to treat fever up to 3 hours after treatment began. Between 4 and 8 hours combined therapy had a higher proportion of afebrile children. This finding was statistically significant. The evidence for this finding was of moderate to low quality.

One RCT study found that total time without fever was longer in children who received paracetamol and ibuprofen combined compared with children who received ibuprofen to treat fever up until 24 hours after treatment began. This finding was statistically significant. The evidence for this finding was of moderate quality.

One RCT study reported no difference in adverse events between paracetamol and ibuprofen combined and paracetamol alone.

Ibuprofen compared with paracetamol and ibuprofen alternating

One RCT found discomfort (quality of life) was less up to 3 days after treatment began in children who received alternating paracetamol and ibuprofen compared with children who received ibuprofen only to treat fever. This finding was statistically significant. The evidence for this finding was high quality.

Three RCTs found that temperature was reduced more in children who received alternating paracetamol and ibuprofen compared with children who received only paracetamol to treat fever. This finding was statistically significant at 5 hours and from 1 to 3 days after treatment began. The effect size was moderate to high. The evidence was of high to very low quality.

Two RCTs found a higher proportion of children who were afebrile after they received alternating ibuprofen and ibuprofen combined compared with children who received only ibuprofen. This finding was statistically significant. This effect size was large. The evidence was of low to very low quality.

One RCT reported no difference in adverse events in children who received alternating paracetamol and ibuprofen compared with children who received only paracetamol to treat fever. The evidence was of low quality.

Paracetamol and ibuprofen combined compared with paracetamol and ibuprofen alternating

No data on quality of life was identified.

One RCT found no difference in temperature reduction in children who received alternating paracetamol

and ibuprofen compared with children who received combined paracetamol and ibuprofen taken together to treat fever. The evidence was of low quality.

No evidence was found reporting adverse events.

Health economics profile

No cost effectiveness studies were identified for this question and no additional health economics analysis was undertaken.

Evidence to recommendations

Relative value placed on the outcomes considered

The GDG stated that the overarching aim of the guideline was the early and accurate detection of serious illness in children with fever. This allows for suitable treatment to begin, which will then reduce morbidity and mortality.

For this review, the aim was to assess the effectiveness of antipyretics. The GDG stated 'distress' was the main concern for parents and carers, and for the majority of children with self-limiting viral disease the aim of treatment would be to relieve 'distress'. Therefore, change in a child's level of 'distress' was used as the primary outcome. However, although 'distress' was the primary outcome, the GDG recognised that is was a poorly understood concept and rarely measured in clinical studies, therefore secondary outcomes of change in temperature and time without fever were also used as proxies for 'distress'. Furthermore, both short- and long-term adverse events were assessed.

Consideration of clinical benefits and harms

The GDG members stated that, to their knowledge, all relevant available evidence had been reviewed.

The GDG emphasised that the evidence shows that both ibuprofen and paracetamol reduce temperature in febrile children, and that both also improve some aspects of quality of life. Although there was some evidence that paracetamol was associated with increased risks of asthma, rhinoconjunctivitis and eczema, the GDG recognised that the evidence from this study did not show a causal pathway between use of paracetamol and long-term conditions.

Evidence shows that improvement in quality of life and reduction in temperature was greater with ibuprofen than paracetamol within 4 hours of treatment starting, but that there were no differences over the longer term. No difference was found in the rate of adverse events reported. The GDG concluded that these differences were not clinically important in that either agent is likely to be effective in any individual case. On this basis the GDG concluded that either paracetamol or ibuprofen could be used.

Evidence showed little difference between paracetamol and ibuprofen given alone or given simultaneously to reduce temperature. The GDG recognised that some of the evidence showed a small benefit in reducing temperature when both drugs were given together, but no evidence of a reduction in distress, which was the primary outcome. The GDG recognised that the simultaneous administration of paracetamol and ibuprofen is sometime used by healthcare professionals and carers. However, there is no evidence on effectiveness to support this approach and a lack of data on safety. Furthermore, each drug is known to be effective as a single agent and the potential for confusion and drug administration errors is increased by using more than one drug.

There was limited evidence showing that improvement in quality of life and temperature reduction was greater when paracetamol was alternated with ibuprofen compared with either treatment alone. The GDG recognises an alternating regimen of paracetamol and ibuprofen is sometimes used by healthcare professionals and carers. However, although there was some evidence showing the efficacy of this approach, there was a lack of data on safety outside research settings. The GDG concluded that it would not be unreasonable for healthcare professionals to advise alternating the two agents if they had both been ineffective as standalone treatments.

Healthcare professionals and others involved in the supply of these drugs should ensure that parents understand how to administer them safely, and explain that they are intended for short-term use only. Healthcare professionals should also check whether the child is receiving any other drugs to avoid the risk of drug interactions or inadvertent overdose.

Consideration of health benefits and resource uses

The clinical review reported no evidence of any difference between ibuprofen and paracetamol and some limited evidence of improvement when the regimens were alternated. The benefit of antipyretics lasts a few hours, and the impact on quality of life of the child can be described qualitatively but has not been translated into a meaningful quality of life health state that can be translated into quality adjusted life years (QALYs). There is no evidence that there is a long-term benefit or harm from reducing fever

or that it changes the course of serious bacterial illness. For these reasons a cost effectiveness analysis was not considered to be feasible for this question.

Paracetamol is less expensive than ibuprofen (£1.05 per 200 mL for the oral suspension compared with £2.71 per 200 mL for ibuprofen oral suspension, August 2012). However, the GDG recognised that, in reality, parents and carers have often given their child one antipyretic or the other before seeking medical advice. There is no evidence that switching to the cheaper alternative once a child is on ibuprofen is a cost-effective strategy.

The GDG noted that healthcare professionals routinely advise using paracetamol when discharging children from hospital because it is cheaper than ibuprofen. Parents and carers may have a preference for one preparation over the other based on their past experience with these agents and they may decide to buy their own preferred preparation. The advice to parents not to alternate treatments should dissuade healthcare professionals from offering both paracetamol and ibuprofen on discharge from hospital. It should also reduce the doubling up of NHS prescriptions of both antipyretics where this remains routine practice.

Quality of evidence

A large number of relevant studies were identified for this review. The evidence varied from high to very low in quality depending on the study design and outcome being measured. There was considerable heterogeneity in the treatment regimens used between studies in terms of dosage and timing of administration. In addition, the included populations varied, especially in relation to age and the underlying condition. For a number of studies data had to be extracted from graphs and this is liable to measurement error; this was counted as an imprecision and the quality of the evidence was downgraded in these circumstances. It was for these reasons that meta-analysis was not undertaken.

Other considerations

No inequalities issues were raised in relation to this question.

The GDG was aware that a Cochrane review examining the effectiveness of antipyretics on fever in children was undertaken at the same time as this guideline was under development, but it was unpublished at the time of submission. Discussions with the authors of this review have shown that the same studies have been selected for inclusion.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

10 Advice for home care

Introduction

Feverish illness in children is a normal and common event although it can cause significant anxiety for some parents and carers. Parents may seek support from healthcare services but in most cases the parents can be reassured that the child is best cared for at home. They may need support and advice to do this confidently. The overwhelming majority of children will recover quickly and without problems. However, in a few cases the child's condition may worsen or fail to improve. Parents need information on when and how to seek further advice.

The guidelines development group (GDG) has found evidence to show that administering antipyretics can make a child look better and feel better and therefore make it easier to differentiate those with serious illness from those with non-serious illness. However, there is no evidence to show that it is desirable to administer antipyretics to reduce fever. The desirability of reducing fever is controversial.

Where no evidence was found to answer the questions, the Delphi survey was used. Full details of the survey are available in Appendix A.

10.1 Care at home

The GDG considered subjects that could usefully be included in written or verbal advice for parents and carers following an encounter with the health services regarding a febrile child.

Review question

What advice should be given to parents for further management of a febrile child?

Need to consider:

- hydration
- feeding
- frequency of temperature monitoring
- methods of cooling
- when to attend nursery or school
- appearance of non-blanching rash.

Methods of cooling

Antipyretic interventions are discussed in Chapter 9, and they should be included in advice for parents or carers.

Fluids

One systematic review (SR)[233] reporting that there were no randomised controlled trials (RCTs) assessing the effect of increasing fluid intake in acute respiratory infections found. No further studies were found meeting the inclusion criteria about giving oral fluids and thus the Delphi survey was used.

Delphi statement 1.1

Parents/carers looking after a feverish child at home should be advised to offer the child regular fluids (where a baby or child is breastfed the most appropriate fluid is breast milk).

In round 1 of the survey the rating categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
0	1 (2%)	48 (96%)	1 (2%)	3	50	9

The statement achieved 96% agreement and thus consensus.

Dehydration

A lack of evidence was found about whether to advise the parents/carers to look for signs of dehydration. This then was included in the Delphi survey (see section 3.2).

Delphi statement 1.2

Parents/carers looking after a feverish child at home should be advised how to detect signs of dehydration.

In round 1 of the survey the rating categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
0	6 (12%)	42 (84%)	2 (4%)	3	50	8.5

The statement achieved 84% agreement and thus consensus.

There was some evidence about which features parents and carers should look for. Refer to section 5.4 for symptoms and signs of dehydration for this purpose. The GDG decided that parents or carers should be advised to look for the most sensitive symptoms and signs of dehydration so that cases are not missed, and if signs of dehydration are detected the parents/carers should encourage their child to drink more fluids and consider seeking further advice. The relevant features are:

- sunken fontanelle
- dry mouth
- sunken eyes
- absence of tears
- poor overall appearance.

Checking temperature

A lack of relevant evidence was found about advising parents/carers to regularly measure their child's temperature if the condition is stable. Therefore this was included in the Delphi survey.

Delphi statement 1.3

Parents/carers looking after a feverish child at home should be advised that regular measurement of the child's temperature is not necessary if the child's condition is stable.

In round 1 of the Delphi survey the rating categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
8 (16%)	17 (33%)	24 (47%)	2 (4%)	2	51	7

Consensus was therefore not reached in round 1.

In round 2 the rating categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
9 (18%)	10 (20%)	32 (63%)		1	51	7

As sufficient level of consensus was not achieved, no recommendation could be made about this statement.

There was a lack of evidence to show whether parents/carers looking after a feverish child should check their child during the night. This therefore was included in the Delphi survey.

Delphi statement 1.4

Parents/carers looking after a feverish child at home should be advised to check their child during the night.

In round 1 the rating categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
2 (4%)	11 (22%)	35 (70%)	2 (4%)	3	50	8

Sufficient consensus was not achieved in round 1.

In round 2 the rating categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	5 (10%)	45 (88%)		1	51	8

Therefore sufficient consensus was achieved. As there is no evidence to show how often the parents/carers should check the child during the night, the healthcare professional assessing the child may want to advise on this.

School attendance

The Department for Education and Skills (DfES) has strict policies that emphasise the importance of good school attendance, and that parents should notify their school on the first day of absence through illness, for health and safety reasons. Nevertheless, although there is a document readily available in schools that shows how long a child should be absent if the child has a known infectious disease, there is no evidence that shows how long a child with a fever of unknown origin should be absent from school or nursery and, this was sent to the Delphi panel.

Delphi statement 1.5

The Department for Education and Skills (DfES) has strict policies that emphasise the importance of good school attendance, and that parents should notify their school on the first day of absence through illness, for health and safety reasons. Nevertheless, although there is a document readily available in schools that shows how long a child should be absent if the child has a known infectious disease, there is no evidence that shows how long a child with a fever of unknown origin should be absent from school or nursery and, this was sent to the Delphi panel.

Parents/carers looking after a feverish child at home should be advised to keep their child away from nursery or school while the child's fever persists but to notify the school or nursery of the illness.

In round 1 the ratings categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	5 (10%)	43 (86%)	1 (2%)	3	50	8.5

Consensus was therefore achieved for this statement.

Appearance of non-blanching rash

At the suggestion of a stakeholder, the GDG decided that parents/carers should be told how to identify a non-blanching rash. A non-blanching rash is a feature of meningococcal disease (see section 5.5) and many parents and carers are aware of its significance. Advice centres around the 'tumbler test' in which the rash is found to maintain its colour when glass is pressed on to the skin.

Health economics

The GDG did not identify any health economics issues for the NHS in this section of the guideline.

GDG translation

The GDG accepted that all Delphi statements that achieved consensus should be used to make recommendations about advice for care at home following an encounter with the health services. For clarity, information about the relevant features to look for was added to the recommendation on dehydration.

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

Research recommendations

Number	Research recommendation
	Home-based antipyretic use
RR	The GDG recommends studies on home-based antipyretic use and parental perception of distress caused by fever. [new 2013].
	Why this is important
	The current guideline recommends the use of antipyretics to relieve distress in children. However, the concept of 'distress' and how parents act on it is little understood. Therefore, the GDG recommends that a study is undertaken to investigate 'distress' in children with feverish illness. The study should include parents' and carers' interpretation of this, including: help-seeking behaviour, what triggers presentation to a healthcare professional, what triggers the decision to give a dose of antipyretic, and what triggers the decision to change from one antipyretic to another.

10.2 When to seek further help

In addition to advice about how to care for their febrile child at home, parents and carers also need advice about when they should seek further attention from the health services. This should allow them to take appropriate action if their child deteriorates or does not recover as expected.

Review question

In children with fever at home following a clinical encounter, what indications should direct the parents or carers to seek further advice?

A lack of evidence was found about when parents should seek further advice following a contact with a healthcare professional. Therefore the following statements were included in the Delphi survey.

Fits

Delphi statement 3.1a

Following contact with a healthcare professional, parents/carers who are looking after their feverish child at home should seek further advice if the child suffers a fit.

The first round consensus rating categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
0	0	52 (98%)	1 (2%)		53	9

Consensus was therefore achieved for this statement.

Less well

Delphi statement 3.1b

Following contact with a healthcare professional, parents/carers who are looking after their feverish child at home should seek further advice if the parent/carer feels that child is less well than when they

previously sought advice.

The first round ratings categories for this statement were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
0	2 (4%)	50 (94%)	1 (2%)		53	8

Consensus was therefore achieved for this statement.

Increased parental concern

Delphi statement 3.1c

Following contact with a healthcare professional, parents/carers who are looking after their feverish child at home should seek further advice if they are more worried than when they previously sought advice.

The first round consensus rating categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
0	9 (17%)	43 (81%)	1 (2%)		53	8

Consensus was therefore achieved for this statement.

Length of fever

Delphi statement 3.1d

Following contact with a healthcare professional, parents/carers who are looking after their feverish child at home should seek further advice if the fever lasts longer than 48 hours.

The first round survey ratings categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
4 (8%)	14 (27%)	33 (63%)	1 (2%)	1	52	7

As no consensus was achieved, it went to round 2 where the ratings categories were**:**

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
2 (4%)	9 (17%)	40 (77%)	1 (2%)		52	7

Consensus was therefore achieved for this statement.

Delphi statement 3.1e

Following contact with a healthcare professional, parents/carers who are looking after their feverish child at home should seek further advice if the fever lasts longer than 5 days.

The first round ratings categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	0	50 (96%)	1 (2%)	1	52	9

Consensus was therefore achieved for this statement.

Parental distress and unable to cope

Delphi statement 3.1f

Following contact with a healthcare professional, parents/carers who are looking after their feverish child at home should seek further advice if the parent/carer is very distressed or unable to cope with their child's illness.

The first round ratings categories were:

1 to 3	4 to 6	7 to 9	Don't know	Missing	Total	Median
1 (2%)	5 (9%)	46 (87%)	1 (2%)		53	9

Consensus was therefore achieved for this statement.

Non-blanching rash

After suggestions from stakeholders, the GDG also decided that parents and carers should seek further advice if the child develops a non-blanching rash.

Health economics

The GDG did not identify any issues that required cost-effectiveness analysis for this question.

GDG translation

The GDG decided to include all but one of the Delphi statements that had achieved consensus as recommendations in the guideline. The exception was the statement about seeking further advice if the fever lasts for more than 48 hours. The GDG unanimously decided not to include this statement because they had found evidence on the predictive value of duration of fever after the statement had been put to the Delphi panel. This evidence, which is detailed in section 5.5, suggests that a duration of fever of around 1–2 days is not predictive of serious illness. The GDG considered that it would therefore be contradictory to advise carers to seek medical attention if the fever lasts longer than 48 hours. The statement on seeking advice if the fever lasted longer than 5 days was retained because the GDG considered this situation to be unusual and because a fever of 5 days duration could be a marker of Kawasaki disease or other serious illnesses such as pneumonia or urinary tract infection (UTI).

Recommendations

The current recommendations can be found at www.nice.org.uk/guidance/ng143

11 Health economics

11.1 Cost analysis of thermometers for use in children and infants with fever

Introduction

A cost analysis of the various types of thermometers available in the UK was undertaken for the 2007 guideline in order to demonstrate the range of costs associated with thermometers. The prices for each type of thermometer were obtained from a review of clinical thermometers in the UK market published by the Medicines and Healthcare products Regulatory Agency (MHRA).[27] This review provided an overview of the clinical and procurement issues for each reported thermometer.

The report showed that the price of 'stand-alone' thermometers is highly variable. Prices range from 7p each for disposable chemical thermometers to £400 for some models of electronic contact thermometers. Given this variation, it is important to take into account a range of issues before determining which device is the best choice and achieves best practice.

Apart from the cost of purchasing it is necessary to consider the cost associated with the use of them. For instance, the manufacturers of some thermometers recommend the use of specific disposable covers to help to reduce the risk of cross-infection for those devices that cannot be adequately cleaned. Also, in some cases it may be necessary to take into account the cost of training for the clinical staff. The clinical risk from incorrect readings may be reduced by the staff undertaking competency-based training programmes. Some electronic thermometers are battery powered so the cost of battery replacement should be included in a detailed costing analysis of thermometers. Also, the cost of recalibration and the cost of maintenance are important elements of cost for some specific types of thermometer.

11.2 Description of the costing analysis

In general, thermometry can be categorised by the type of the instrument used and by the site at which the temperature is read. Mercury in glass, electronic and chemical dot thermometers can be used sublingually (orally), in the axilla (under arm) or rectally. Temperature assessment accuracy is critically important. False high readings may lead to expensive and unnecessary painful diagnostic tests and medical interventions. False low readings may lead to greater morbidity and mortality.

Accuracy of body temperature measurement depends not only on the type of thermometer but also on the site of measurement. Given that the site of measurement is a clinically important decision, the classification of the thermometers for this cost analysis was based on the site of measurement. Some types of thermometers cannot provide readings from all the sites of measurements. For instance, chemical thermometers cannot give rectal measurements.

Methods

The structure of the cost analysis and the assumptions in it are based on that devised by Crawford et al.[27] The analysis includes three types of thermometer: chemical, electronic and infrared sensing, and classified according to two different sites of measurements: axilla and tympanic.

The thermometers were subdivided into subcategories of electronic and chemical thermometers since there are cost differences between them. The category of electronic thermometers was split into contact/electronic and contact/compact electronic thermometers.

A robust cost comparison between different technologies should ideally encompass all the contributory costs over a prescribed period: in this case, a 10 year time horizon was used, discounted

at 3.5%. The analysis calculated both the most expensive and the least costly model of each category of thermometer in order to demonstrate the range of costs for each type and how the costs might overlap depending on which model is chosen.

This economic assessment only includes the direct costs of purchase price and, where applicable, the costs of consumables (e.g. disposable covers, sterilised alcohol-impregnated wipes and replacement batteries). Cleaning, maintenance, repair, and calibration costs, although important, were not included here owing to time constraints in collecting the data for the guideline. However, they are not considered to have an important influence on the relative costs of each model compared with its alternatives.

Device-specific costs were obtained from the MHRA.[27] The same assumptions were used as a basis for the calculation of the costs as were used by Crawford et al.[85] Table 11.1 summarises the assumptions used in the costing model.

Axilla measurements can be provided by electronic and chemical thermometers. Tympanic measurements are by specialised infrared sensing thermometers only. Chemical thermometers supplied by different companies use different chemicals. Some change permanently when the temperature is raised (e.g. 3M Tempadot) and others change colour for only a short while when placed in contact with a hot object and then return to the original colour (e.g. Insight Nextemp). Both may be labelled single use, but the second type can be used again on the same patient (providing that it is kept clean with alcohol wipes), and is considered to be a reusable model in this analysis.

The cost of staff time required to measure temperature using each type of thermometer was included in the analysis. Each thermometer has an average time to reading, which gives a total number of hours required to read the thermometer per year, which was then calculated up to the 10 year time horizon used in the cost analysis. This average time to reading is based on best guesses and not on empirical data. These times are indicative only since they exclude any time to locate the device, clean the device or fit and remove probe covers. Also, it does not take into account that nurses may be undertaking other tasks while waiting for a reading for thermometers where this may take more than a few seconds. For some adhesive chemical thermometers (e.g. Insight Traxit), the time to reading changed depending on whether it was a first measurement or subsequent measurement since the thermometer was already in position and at the correct temperature. Therefore the average time per patient episode was calculated to be 180 seconds plus 85 seconds (17 × 5) for the 18 measurements, giving a total of 265 seconds.

Table 11.1 Assumptions used in the costing model

	Contact/chemical	**Electronic contact**	**Compact contact electronic**	**Infrared sensing (tympanic)**
Number purchased	One per measurement (1,550,000)	One per unit (450)	One per hospital bed (2205)	One per unit (450)
Consumables	Alcohol wipes may be required if single-patient-use devices are used	Probe covers	Alcohol wipes	Probe covers
Battery replacement	No	Yes	Yes	Yes
Replacement	Each patient or each measurement, depending on the model	0%	10% per annum	0%
Approximate readings per inpatient episode	18	18	18	18
Inpatient episodes per year	86,000	86,000	86,000	86,000

The nursing cost per hour (£22) was the hourly cost for a staff nurse on a 24 hour ward published in the Unit Costs of Health and Social Care for 2006,[242] which was based on the Agenda for Change salaries for the April 2005 scale at the midpoint for Band 5 (with qualifications).

It should be noted that the analysis did not take into account the additional staff time to change batteries and undertake basic performance checks, although it was recognised that for some models the manufacturers recommend (at least annual) performance and accuracy checks using specialised equipment that can be arranged when a battery needs replacement.

The costs of calibration (a specialised accuracy check) and warranty are not included in the analysis, which is a limitation of the model.

The cost of cleaning (alcohol wipes) is included where these are required after each measurement. For the contact/chemical thermometers used on a single patient, alcohol wipes are not required. For the contact/compact electronic thermometers (axilla using disposable covers), alcohol cleaning of the thermometer body is only required 'when needed' and this is unlikely to be after every measurement. Therefore it was assumed that an alcohol wipe was used after every 50 measurements.

An approximation of 18 readings per inpatient episode was estimated by dividing the estimated number of measurements per year by the number of inpatient episodes per year, and rounding up to the nearest whole number.

Using the above assumptions, the overall cost for each type of thermometer was calculated for those which can provide axilla and ear measurements. The total cost for each type of thermometer for 10 years was calculated using for each site of measurement the minimum and maximum price of the thermometers.

The clinical accuracy of the thermometers is assumed to be the same for all models of thermometer and in all measurement sites in this analysis. This is due to the lack of data on comparative accuracy or ability to detect fever by different models of thermometer, and the lack of data on the impact of temperature accuracy on time to correct diagnosis and initiation of clinical management in children with suspected serious bacterial infection. The assumption is that, used correctly, all the thermometers considered in this analysis can detect a clinically important rise in temperature.

Results

Axilla measurements

Tables 11.2 and 11.3 show the results of the cost analysis for axilla measurement showing the comparative costs over 10 years using maximum and minimum prices for each type of thermometer.

Table 11.2 indicates that, in an acute care setting, using the least cost models available on axilla sites and including the cost of staff, the compact contact electronic thermometer is the best value for money, followed by the reusable contact/chemical thermometer, although this is four times more expensive. The cheapest electronic contact and the single-use chemical thermometers are more than 12 times more expensive than the cheapest contact/electronic thermometer. The large difference in staff time required to take a temperature (5 seconds versus 3 minutes) account for much of the large difference in cost between these types of thermometer.

Table 11.3 shows that using the most expensive models of reusable chemical thermometers in terms of initial purchase price can be less costly over 10 years than the cheaper models. The total cost of the high-priced model including staff time was more than 12 times less than the total cost using the cheapest priced reusable chemical thermometer because the expensive model took only 5 seconds to read after the first initial 3 minute reading. Overall, the results suggest that, in an acute care setting, the best option for a top of the range thermometer was the reusable chemical model, followed by the compact contact electronic model. The worst option was the single-use chemical thermometer which cost over £20 million over 10 years (£14 million when discounted by 3.5%), which was over 14 times more expensive than the next most expensive, which was the electronic contact model (undiscounted).

Table 11.2 Comparative cost of thermometers that can provide axilla measurements in a large teaching

	Type of thermometer			
	Single-measurement contact/chemical (phase change)	**Reusable contact/chemical (phase change)**	**Electronic contact**	**Compact contact electronic**
Model used	3M Tempadot	EzeTemp	Sure Temp. Plus	Microlife MT 1671
Supply of thermometers	One per measurement	One per patient episode	One per ward	One per bed
Purchase cost	£0.07	£0.14	£150.00	£3.36
Price of consumables items and ongoing costs (per item)				
Covers			£0.0275	
Battery life (readings)			5,000	3,000
Cost of batteries			£0.75	£0.2200
Cost of cleaning (alcohol wipes)		£0.008	£0.008	£0.008
Annual cost of consumables and ongoing costs calculated using the assumptions stated in Table 11.1				
Initial purchase cost	£108,500	£12,040	£67,500	£7,409
Replacement cost per year (10%)				£741
Number of batteries/year			310	517
Cost of batteries /year			£233	£114
Cost of alcohol wipes/year		£12,400	£248	£12,400
Cost of covers/year			£42,625	
Total cost consumables		£12,400	£43,416	£13,771
Time to reading (seconds)	180	180	6	60
Seconds on reading/year	279,000,000	279,000,000	9,300,000	93,000,000
Hours on reading/year	77,500	77,500	2,583	25,833
Annual staff costs	£1,705,000	£1,705,000	£56,833	£568,333

	Type of thermometer			
	Single-measurement contact/chemical (phase change)	**Reusable contact/chemical (phase change)**	**Electronic contact**	**Compact contact electronic**
Recurring costs per year (consumables, replacement, staff)	£1,813,500	£1,729,440	£100,249	£582,845
Recurring costs per year (consumables and replacement)	£108,500	£24,440	£110,916	£14,512
Total undiscounted 10 year cost (with staff costs)	**£18,135,000**	**£17,294,400**	**£1,069,988**	**£5,835,863**
Discounted at 3.5%	£12,856,243	£12,260,326	£758,535	£4,137,153
Total undiscounted 10 year cost (without staff costs)	**£1,085,000**	**£244,400**	**£1,176,655**	**£152,530**
Discounted at 3.5%	£769,177	£173,260	£834,153	£108,131

Table 11.3 Comparative cost of thermometers that can provide axilla measurements in a large teaching hospital for 10 years – maximum prices

	Type of thermometer			
	Single-measurement contact/chemical (phase change)	**Reusable contact/chemical (phase change)**	**Electronic contact**	**Contact/compact electronic**
Model used	Insight NexTemp	Insight Traxit	Ivac Temp. Plus II	Proact ST 714
Supply of thermometers	One per measurement	One per patient episode	One per ward	One per bed
Initial purchase cost	0.24	£0.61	£400.00	£13.95
Price of consumables items and ongoing costs (per item)				
Covers			£0.047	£0.045
Battery life (readings)			2,000	1, 800
Cost of batteries			£0.95	£0.5900
Cost of cleaning/ alcohol wipes			£0.008	

	Type of thermometer			
	Single-measurement contact/chemical (phase change)	**Reusable contact/chemical (phase change)**	**Electronic contact**	**Contact/compact electronic**
Annual cost of consumables and ongoing costs calculated using the assumptions stated in Table 11.1				
Initial purchase cost	£372,000	£52,460	£180,000	£30,760
Replacement cost per year (10%)				£3,076
Number of batteries/year			775	861
Cost of batteries/year			£736	£508
Cost of alcohol wipes/year			£12,400	
Cost of covers/year			£72,850	£69,750
Total cost consumables			£85,986	£70,258
Time to first reading (seconds)	180	180	4	5
Time to subsequent readings, if different (seconds)		5		
Seconds on reading/year	279 000 000	7 310 180	6 200 000	7 750 000
Hours on reading/year	77 500	2 031	1 722	2 153
Annual staff costs	£1,705,000	£44,673	£37,889	£47,361
Recurring costs per year (consumables, replacement, staff)	£2,077,000	£97,133	£123,875	£120,695
Recurring costs per year (consumables and replacement)	£372,000	£52,460	£85,986	£73,334
Total undiscounted 10 year cost (with staff costs)	**£20,770,000**	**£971,333**	**£1,418,751**	**£1,237,711**
Discounted at 3.5%	£14,724,244	£688,596	£1,005,780	£877,437
Total undiscounted 10 year cost (without staff costs)	**£3,720,000**	**£524,600**	**£1,039,863**	**£764,100**
Discounted at 3.5%	£2,637,178	£371,899	£737,178	£541,685

Tympanic measurements

Tympanic measurements can be provided by infrared sensing thermometers only, so there is no comparative analysis by different types of thermometer, only by the least and most expensive type of infrared sensing model. The total cost of using exclusively the least costly model and the most expensive model of infrared sensing thermometer was calculated (Table 11.4).

Table 11.4 shows that the lowest purchase price model (the infrared sensing thermometer) has a higher overall cost than the highest priced thermometer because of the increased cost of con- sumables (nearly double the price) which contribute to the total cost. The cost of covers is lower in the most expensive model. The recurring costs per year (consumables and staff) are more than £50,000 more per year for the cheaper model, which outweighs the higher initial purchase price of the most expensive model. The results also indicate that time to reading is not an important cost driver for tympanic measurement since the assumption is that it takes only 2 seconds to make a temperature reading. The (discounted) cost over 10 years including staff costs is in the range £732,000 to £1,064,000, which is the same order of magnitude of costs as the thermometers used for axilla measurement, except that of the single-use chemical thermometer.

Table 11.4 Ten-year costs for infrared sensing thermometers, discounted at 3.5%: summary results for tympanic measurements – minimum and maximum prices

	Model of infrared sensing thermometer (tympanic)	
	TB-100 (thermo Buddy)	**First Temp. Genius**
Purchase cost	£18.32	£249.49
Supply of thermometers	One per ward	One per ward
Price of consumable items and ongoing costs (per item)		
Probe covers	£0.0760	£0.047
Battery life (readings)	6000	5000
Cost of batteries	£0.68	£0.950
Cost of cleaning (alcohol wipes)	£0.008	
Annual cost of consumables and ongoing costs calculated using the assumptions stated in Table 11.1		
Initial purchase cost	£8,244	£112,271
Number of batteries/year	258	310
Cost of batteries/year	£176	£295
Cost of alcohol wipes/year	£12,400	
Cost of covers/year	£117,800	£72,850
Total cost consumables	£130,376	£73,145
Time to reading (seconds)	2	2
Hours on reading/year	861	861
Annual staff costs	£18,944	£18,944
Recurring costs per year (consumables, replacement, staff)	£149,320	£92,089
Recurring costs per year (consumables and replacement)	£130,376	£73,145
Total undiscounted 10-year cost (with staff costs)	**£1,501,445**	**£1,033,160**
Discounted at 3.5%	£1,064,403	£732,427

	Model of infrared sensing thermometer (tympanic)	
	TB-100 (thermo Buddy)	**First Temp. Genius**
Total undiscounted 10-year cost (without staff costs)	**£1,312,001**	**£843,716**
Discounted at 3.5%	£930,102	£598,126

Comparison of costs of axilla and tympanic measurement

Table 11.5 shows the combined results for all types of thermometer used in axilla and tympanic measurement. It indicates that the relative cost of each type of thermometer changes depending on whether an expensive or a cheap model is used and whether staff time is included in the cost, as the time required to read the temperature is an important driver of total cost.

Conclusions

The cost analysis undertaken here is based on the use of thermometers on a ward of an acute hospital. The study[85] on which this analysis is based suggests that staff time is an important driver in determining which thermometer should be used. The analysis presented here supports this hypothesis. The 10 year cost of a (high- and low-priced) thermometer including staff time includes ranges between approximately £600,000 and £1,000,000 for all types of thermometers, except for the single-use chemical thermometer which is far more expensive. The analysis incorporates a number of assumptions about time to reading for accurate measurements, but it suggests that the initial purchase price of thermometers can be misleading as the total cost of using a specific model of thermometer depends on the number of uses, the cost of consumables and the staff time needed to make an accurate reading. Clearly different clinical settings will give different results and may change the relative cost between thermometers, making it more cost-effective to choose one type of thermometer in a low-volume clinical setting and another in a high-volume setting. This analysis shows that those in charge of purchasing thermometers need to consider staff costs and consumables as well as initial purchase price when considering bulk purchases.

Table 11.5 10 year costs by thermometer, with and without staff costs, discounted at 3.5%: summary results for both axilla and tympanic measurements

	10 year cost by type of thermometer				
	Chemical (single use)	**Chemical (reusable)**	**Electronic contact**	**Compact contact electronic**	**Infrared sensing (tympanic)**
Minimum priced model (with staff costs)	£12,856,243	£12,260,326	£758,535	£4,137,153	£1,064,403
Maximum priced model (with staff costs)	£14,724,244	£688,596[a]	£1,005,780	£877,437[a]	£732,427[a]
Minimum priced model (without staff costs)	£769,177	£173,260	£834,153	£108,131	£930,102
Maximum priced model (without staff costs)	£2,637,178	£371,899	£737,178	£541,865	£598,126

[a] Indicates a lower total discounted 10 year cost than the least expensive version of the model due either to higher cost of staff time or consumables.

11.3 Economics of referral to a specialist paediatric team of a child with fever without source, analysis undertaken for the 2007 guideline

Background

One of the key areas where the 2007 guideline that had important resource-use implications is in its impact on changes in referral patterns. Some recommendations in the guideline may lead to a change in current referral practice from general 'first-line' medical care to specialist paediatric services (that is, from primary care, or an emergency department, or following a telephone call to NHS Direct to either hospital-based or community-based paediatricians).

The recommendations in the guideline that may change referral patterns are for a child considered to have an immediately life-threatening illness to be transferred without delay to the care of a paediatric specialist. All children with 'red' features will need to be referred to specialist care, and all children with 'red' or 'amber' features need to be seen within 2 hours if referred from remote assessment.

It was envisaged that the clinical guideline would include an economic analysis of the impact of changing referral patterns. Time was set aside in guideline development group (GDG) meetings to develop a decision tree to analyse the costs and outcomes of such a change.

The decision tree is presented in Figure 11.1. The aim was to undertake a threshold analysis to evaluate the additional costs (or savings) associated with one additional case of serious bacterial illness (SBI) detected.

Structure of the decision model

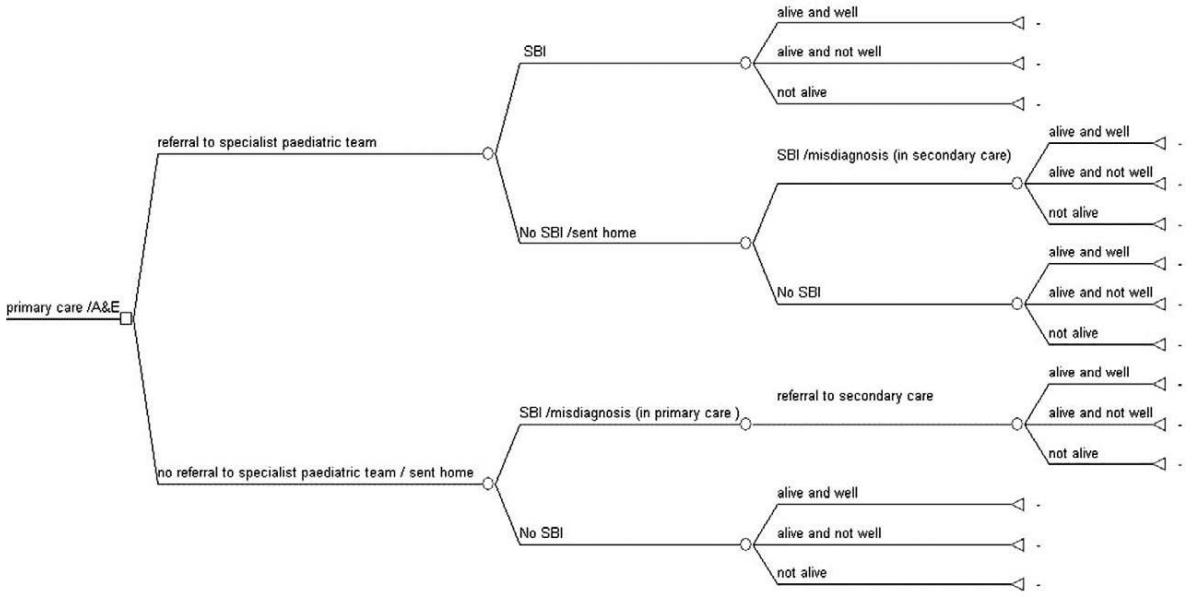
An outline of the pathways of the decision tree is presented in Figure 11.1. The model starts with a population (say, of an average GP practice) of which a proportion of children per year present to 'first-line' services with signs or symptoms of undifferentiated fever.

The first decision (the first split in the pathway) in the model is whether or not to refer the child to specialist paediatric services. If a child is referred, there is a chance that the child has an SBI or they do not. There is a chance that the child may have SBI confirmed through diagnostic tests and subsequently be treated for SBI, and there is a chance that no SBI is confirmed and the child is sent home.

If a child is sent home following referral to a specialist paediatric team, they will improve without treatment if they have no SBI. If they have an untreated SBI, their condition will worsen at home. They will consequently either be sent to hospital (usually as an emergency) or not be sent to hospital. Of those children not sent to hospital, a proportion will improve and be well at home, a proportion will deteriorate but remain unwell, and a proportion will die at home.

If a child is not referred to a specialist paediatric service, there is a chance that they do not have an SBI and would improve without treatment, and a chance that they have an SBI. If they have an SBI, they will either be referred again to a specialist paediatric team for a second time, or not. The structure of the pathway of children referred for a second time to a specialist paediatric team was the same as for children referred the first time, except that it was assumed that a child would not be sent home after a second referral. All children referred to hospital a second time with the same episode of fever without source would be diagnosed and treated for SBI in hospital. This is an assumption and not based on any clinical evidence that could be identified.

Figure 11.1 Decision tree for analysing the impact of changing referral patterns for a child with fever without source



Data required for the model

In order to make this analysis viable, the decision tree required specific data which the GDG thought might be available in some form, through either the published literature or in unpublished data such as national (or even local) audit data. A table with all the key model parameters was circulated among the GDG members to try to locate this data. At the same time, the GDG members were asked whether they could arrive at some consensus about the values required for the model from their collective expert opinion.

As the discussion progressed, it was agreed that the meaningful comparison of referral patterns required other data that would be very hard to obtain either from published sources or from GDG consensus.

A number of key assumptions in the model could not be agreed upon. The first was that the outcomes of care would be worse if treatment was delayed by sending a child home, either from primary care or from secondary care with undiagnosed SBI. Nor was it clear that the costs of care would be substantially different if there were a delay in treatment. It was not possible to estimate the impact that such a delay would have on final outcomes (the death rate) or costs because of the uncertainty around the natural history of specific serious bacterial diseases such as meningitis. Also, it was not possible to agree upon the proportion of children with fever that are currently referred for primary care.

It became apparent after two GDG meetings that it was not possible to reach a consensus on the data required to populate the model, especially because the model considers all forms of SBI and no one specific diagnosis, such as meningitis or pneumonia. Also, since the guideline focused on diagnosis and initial management of SBI only, it would be difficult to obtain reliable data on the number of children alive and well or not alive following detection and initial management of SBI, without looking at treatment and longer term outcomes.

A further problem was the lack of baseline data on the underlying prevalence of SBI in the population. The most uncertain data of all was the estimate of the proportion of cases of SBI that might be missed by sending children home without further tests, in both primary or specialist care settings.

Some data were available from two published studies, one American[243] and one from the UK.[121] Table 11.6 below indicates the data that could be used in the model (part I) and the gaps where no data could be found (part II).

Table 11.6 Data required to complete the economic model for referral of children to specialist paediatric services of children with fever without source

Parameter	Data
Part I: Values where some data were identified	
Primary care	
Proportion of children under 5 referred to a specialist paediatric team (secondary or community care setting) from first-line services (primary care and A&E)	96% secondary care referrals, 4% tertiary referrals[121]
Specialist paediatric care	
In specialist paediatric setting, the proportion of children presenting with undifferentiated fever who screen positive for SBI	62% (460/747 infants)[244]
In specialist paediatric setting, the proportion of children with undifferentiated fever who screened negative for SBI	38%[244]
OR	
In specialist paediatric setting, the proportion of children tested positive for suspected SBI and treated	29% (41/141 infants)[121]
In specialist paediatric setting, the proportion of children screened positive for SBI with a confirmed diagnosis	14% (64/460 infants), 8.7% of all infants admitted (64/747)[244]
In specialist paediatric setting, the proportion of children with no suspected SBI who are admitted for review and go on to develop confirmed SBI	0.68% (1 patient)[244]
In specialist paediatric setting, the proportion of children with no suspected SBI who are sent home (managed as outpatients or under observation at home, with review), who subsequently are admitted to hospital with confirmed SBI	0%[244]
Part II: Values where no data were identified	
Number of children (per year) presenting in primary care with undifferentiated fever (e.g. by region/PCT/GP practice)	
Proportion of children in primary care not referred to specialist paediatric care (no signs/symptoms) who are sent home and subsequently develop SBI	
Proportion of children referred to specialist paediatric care who are sent home and subsequently develop SBI	
Additional healthcare resource use of children sent home from primary care who go on to develop SBI	
Additional healthcare resource use of children sent home from specialist paediatric care who go on to develop SBI	
Outcomes (although outside the scope of the guideline)	
Prognosis/outcome for children who are referred immediately from primary to a specialist paediatric team for suspected SBI: • with confirmed SBI treated in hospital • sent home with no confirmed SBI which subsequently develops into SBI • no subsequently confirmed SBI	Differentiate between: • alive and well • alive and not well • not alive
Prognosis/outcome for children who are NOT referred immediately to a specialist paediatric team for suspected SBI: • who go on to develop SBI • with no SBI	Differentiate between: • alive and well • alive and not well • not alive

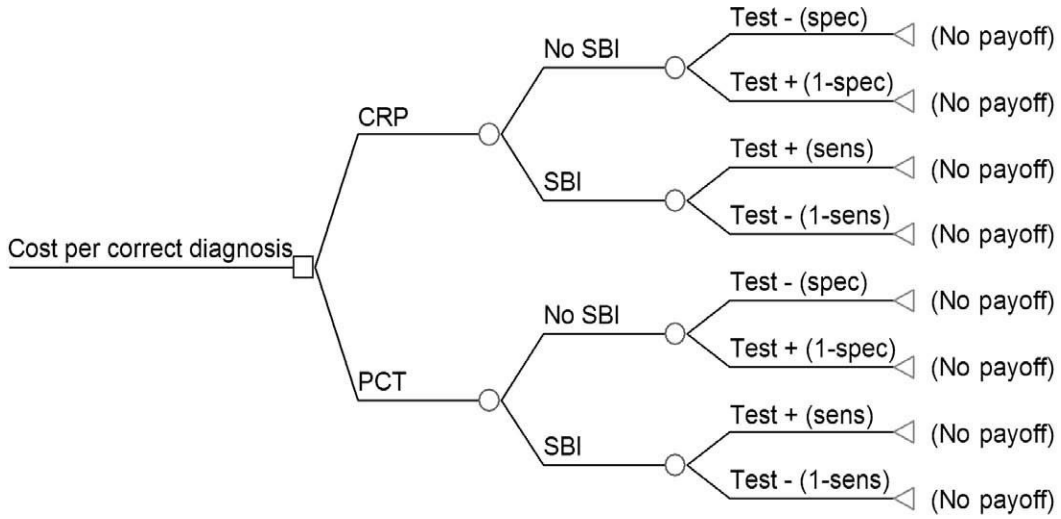
11.4 Economic evaluation of C-reactive protein versus procalcitonin – analysis undertaken for the 2007 guideline

Fever without localising signs in young children remains a diagnostic problem. There is evidence that procalcitonin (PCT) may be more effective in terms of sensitivity than commonly used C-reactive protein (CRP). However, the evidence on diagnostic accuracy is not robust. An economic evaluation approach was adopted to assess the cost-effectiveness of using different estimates of specificity and sensitivity of these tests from the published data.

A simple decision-analytic model was constructed which incorporated both the sensitivity and specificity of each test. Additional correct diagnosis was the outcome used. The model is based on limited information on PCT in children with fever without apparent source (FWS) and in other situations PCT may perform better than CRP.

Figure 11.2 is a schematic representation of the decision tree used in the analysis. Before investigations, febrile children were assumed to have one of two health states: either with no serious bacterial illness (SBI) or with SBI. After the investigations, febrile children were assigned a true positive or negative diagnosis, or a false positive or negative diagnosis. The model covers only the initial diagnosis and not the cost of treatment of SBI. The term SBI for this guideline includes seven potential types of serious infection. Each type of infection would require a different pathway. The description of this pathway and its potential outcomes was beyond the scope of this guideline.

Figure 11.2 Cost-effectiveness of PCT versus CRP decision tree



Methods

Clinical effectiveness

'Correct diagnosis' was identified as the outcome of the analysis. This can take into account both sensitivity and specificity in order to derive the precise levels of correctly diagnosed cases for each type of investigation.

Correct diagnosis = true positive + true negative diagnosis

Data used in the model

Diagnostic accuracy

Estimates of the diagnostic accuracy are taken from the systematic review of the clinical evidence presented in this guideline. Specifically, there are two studies which provide clinical effectiveness for the model. Table 11.7 summarises the data on diagnostic accuracy of PCT and CRP presented in

these studies of children with FWS. The levels of specificity and sensitivity from the most recent study are used as baseline parameters for the model.

Table 11.7 Source of effectiveness data from the existing published studies

	CRP	PCT	Source
Sensitivity	0.79	0.93	Galetto-Lacour et al. (2003)[178]
Specificity	0.79	0.74	
Sensitivity	0.89	0.93	Lacour et al. (2001)[245]
Specificity	0.75	0.78	

CRP C-reactive protein, PCT procalcitonin

Prevalence of SBI for children with fever without localising signs is a key parameter of the model. However, no accurate prevalence data for the UK could be identified. Therefore, an estimate of 5% was used in the first instance based on GDG expert opinion, which is a strong assumption of the analysis. Table 11.8 summarises all the clinical data used as baseline parameters in the model.

Table 11.8 Baseline parameters for the effectiveness data

	CRP	PCT	Source
Prevalence	0.05	0.05	GDG expert opinion
Sensitivity	0.79	0.93	Galetto-Lacour et al. (2003)[178]
Specificity	0.79	0.74	

CRP C-reactive protein, PCT procalcitonin

Costs

The perspective adopted by the economic analysis was that of the NHS, and prices are for 2006. The cost of the test included the cost per investigation only. It was assumed that the price of the investigation reflects the cost of reagents and the cost of labour as well. The cost of CRP could be identified by the GDG members from their local services. However, the cost of PCT was more difficult to estimate since a published price, including all associated costs, could not be identified from the sources available. One GDG member provided the price for a PCT assay. Table 11.9 shows the cost of each type of investigation and the source of the cost data. The potential cost of SBI treatment is not included in the analysis.

Table 11.9 Baseline parameters for the cost data

	CRP	PCT	Source
Cost per investigation	£1.50	£9.00	GDG

CRP C-reactive protein, PCT procalcitonin

Results

A cohort of 1000 febrile children without localising signs for each type of investigation was assumed. The results of the economic analysis are presented as cost per correct diagnosis. Using baseline data, CRP appears to be a significantly less costly and possibly more accurate diagnostic test than PCT in terms of correctly diagnosed cases (Table 11.10). Taking into account only the levels of sensitivity, PCT is a better diagnostic test than CRP as it manages to capture more SBI (more true positives). However, PCT may have a lower level of specificity than CRP which means that PCT identifies fewer true negative results than CRP. Also, the decrease in the correctly diagnosed cases having no SBI is higher than the increase in the correctly diagnosed cases having SBI and for this reason the final number of correctly diagnosed cases is lower for PCT than CRP.

Table 11.10 Additional cost per additional correct diagnosis detected of PCT over CRP

Investigation	Cost	Effectiveness (correct diagnoses)	Incremental cost (additional cost of PCT over CRP)	Incremental effectiveness (additional correct diagnosis)	Additional cost per additional correct diagnosis
CRP	£1,500	790			
PCT	£9,000	750	£17,500	−41	Dominated (more costly, less effective)

CRP C-reactive protein, PCT procalcitonin

Sensitivity analysis

Both one-way and two-way sensitivity analyses were undertaken. One-way sensitivity analysis involves altering the value of a single parameter while holding all the others constant, to determine how robust the conclusion is to the values used in the model. Two-way sensitivity analysis means that two parameters are changed simultaneously.

1. Varying the prevalence of SBI in the population

Given that there is lack of published evidence with regard to the prevalence of SBI for the febrile children, sensitivity analysis was conducted by varying the levels of prevalence in order to assess the extent to which the final results are dependent on change in this parameter. CRP dominated PCT until the prevalence reached 27% in the population. However, the additional cost per additional correct diagnosis was £5,769.

2. Diagnostic accuracy of CRP and PCT

Sensitivity analysis was conducted by using various estimates of the diagnostic accuracy of the tests. Data from an older study conducted by the same authors[245] was inputted into the cost analysis. Table 11.11 shows that, using different data for diagnostic accuracy, the additional cost per additional correct diagnosis by switching from using CRP to PCT to detect SBI may be up to £246 per test.

Table 11.11 Results of sensitivity analysis using levels of diagnostic accuracy from the second study[245]

Investigation	Cost	Effectiveness (correct diagnoses)	Incremental cost (additional cost of PCT over CRP)	Incremental effectiveness (additional correct diagnosis)	Additional cost per additional correct diagnosis
CRP	£1,500	757			
PCT	£9,000	788	£7,500	31	£246

CRP C-reactive protein, PCT procalcitonin

3. Sensitivity of the diagnostic tests

One-way sensitivity analysis was conducted to test the robustness of the final results by varying the levels of sensitivity of the tests only. CRP still dominated PCT when the level of sensitivity for PCT was increased to 1.00 (maximum). Also, CRP still dominated PCT even after decreasing significantly the level for CRP. This means that the CRP was still more cost-effective than PCT even when changing only the levels of sensitivity of PCT and CRP.

4. Specificity of the diagnostic tests

Sensitivity analysis was undertaken to check the robustness of the results with regard to the levels of specificity. The final results were sensitive to the level of specificity of the tests. By increasing the level of specificity from 0.74 to 0.79, the PCT became more effective than CRP. However, the additional cost per additional correct diagnosis was £1,071 per test.

Limitations

The economic analysis of PCT versus CRP was based on the best available evidence, which was completely absent for prevalence of SBI. Also, the sensitivity and specificity data were from a very limited number of studies of children with FWS. Generally, PCT performs better than CRP in other situations, so FWS data may not be reliable.

Therefore, great care is needed when interpreting and deriving the final results of this analysis, as there are some limitations. Sensitivity analysis shows that the final results are sensitive to the prevalence of SBI and to the levels of diagnostic accuracy at a cost per test of £1.50 and £9.00 for CRP and PCT, respectively (cost data was from GDG members and not published data). This indicates that the validity of the results depends considerably on the quality of the data which are used in order to derive the levels of correct diagnosis.

Another caveat of the model is the choice of outcome measure. The preferred methodology according to the NICE technical manual is to present outcomes in terms of the quality-adjusted life year (QALY). Given the range of SBIs under consideration, and the associated range of treatment pathways, it was impossible to estimate the cost per QALY for these diagnostic tests. This may have some influence over the results, as some children may undergo unnecessary treatment, while others will not be given required treatment, based on false results following diagnosis. By measuring the results in cost per correct diagnosis, the model may not reflect the true long-term costs and outcomes associated with each diagnostic method.

Conclusions

Using the strong baseline assumptions, CRP appears to be both less costly and to provide more correct diagnoses than PCT. However, this result was highly sensitive to test accuracies, which were different in the two studies that reported data for diagnosing SBI in children with fever without localising signs. PCT became more effective than CRP even with small changes in specificity but this increase in effectiveness is associated with higher cost per correct diagnosis.

Without conversion to QALYs, it is not possible to assess whether this additional cost is ‘worth’ the additional benefits of PCT.

Given current published evidence, this economic analysis does not support the replacement of CRP with PCT in routine practice.

11.5 Hour time limit for an urgent face-to-face consultation following remote assessment: GDG reasoning and justification in the absence of data to inform a formal economic analysis – analysis undertaken for the 2007 guideline

Background

The GDG was asked to produce a guideline to aid healthcare professionals in identifying children with serious bacterial illness (SBI) in an attempt to reduce mortality and morbidity in young children. During the guideline development process, the GDG identified evidence-based symptoms and signs that indicate whether a child has a high risk of having SBI. It also identified symptoms and signs that indicate that a child is at very low risk of SBI and can be looked after at home. Current practice is not evidence based and is variable. It is likely that referral patterns from some healthcare providers will change when the guideline is implemented. It is anticipated that some children who would previously not always have been recognised as needing specialist attention (a very small proportion of children who present with fever) will in the future be referred for consultation with a specialist. Furthermore, a number of children for whom referral is not indicated (the far larger proportion) and who would previously have been referred for consultation or unnecessary investigations, will now not be referred unnecessarily under this new guidance. The focus of the guideline is that the right children should be getting the right treatment at the right time and adverse health outcomes (including death) will therefore be avoided. The GDG noted the evidence that problem-based guidelines with carepathways for children with medical problems reduce invasive investigations, and lead to more appropriate treatment and reduced time spent in accident and emergency (A&E) services.[246]

GDG justification of the 2 hour waiting time for an urgent referral

An important feature of this clinical guideline on children with feverish illness is the introduction of a 'traffic light' system to identify children with varying degrees of risk of serious illness. The guideline makes clear recommendations on which children are unlikely to require medical attention beyond information and reassurance (children with 'green' features) and who can thus be confidently managed at home. The guideline identifies children who require an urgent face-to-face consultation with a healthcare professional ('red') and those who may require a face-to-face consultation or require a healthcare 'safety net' to be put in place ('amber').

Because of the limited information that can be obtained from a remote assessment, the GDG originally recommended that all children with 'red' or 'amber' features should be referred for urgent face-to-face assessment. The GDG felt it was necessary to make a recommendation on the maximum time a child should have to wait to be first assessed by a healthcare professional if they were classified as requiring an urgent consultation during a remote assessment. The aim of this was to recommend a time frame within which action taken will make a difference to the outcome for the child.

Despite an extensive search of the published and grey literature, no clinical data could be identified to define this limit. The GDG debated the issue among themselves and decided that it was such an important question that wider consensus was required. Accordingly, the question went out as part of the Delphi consultation exercise as agreed in the guideline methods protocol. A high level of agreement was reached for a maximum wait of 2 hours following referral for urgent face-to-face assessment (83% agreement). 2 hours was chosen as one of the time periods for the Delphi exercise because it is an existing Department of Health standard for urgent referrals for out-of-hours health care.[247]

It was recognised by the GDG that children with one or more 'amber' signs included children who may not require an urgent referral. It was agreed to make a recommendation on specific waiting times only for children with 'red' features, and to recommend that a child with one or more 'amber' features is seen face-to-face by a healthcare professional, but that the timing of the consultation for these children could be carried out within a longer time frame which could be based on the clinical judgement of the person carrying out the initial remote assessment.

The GDG believes that a 2 hour maximum wait for an urgent consultation does not represent an uplift in care and is a cost-effective use of NHS resources. The reasons for this conclusion are outlined here. First, there is audit data to suggest that this is already accepted routine practice for children at a high risk of SBI. Second, the GDG strongly believes that a wait longer than 2 hours could potentially increase mortality and morbidity. Finally, the GDG believes that by using a traffic light system to classify children according to their risk of having a serious illness, healthcare professionals will have a clearer indication as to which children do genuinely require an assessment by a healthcare professional within 2 hours. By excluding the children with 'green' features and most of the children with 'amber' features from this urgent referral group, the GDG believes the number of children who are referred for a face-to-face assessment by a healthcare professional within 2 hours will be reduced.

Evidence was presented to the GDG to show that the Department of Health has already set a national standard for response to urgent calls as part of the *National Quality Requirements in the Delivery of Out-of-Hours Services*.[247] This specifies a maximum 2 hour wait for a face-to face urgent consultation for out-of-hours care: 'Face-to-face consultations (whether in a centre or in the patient's place of residence) must be started within the following timescales, after the definitive clinical assessment has been completed:

- Emergency: Within 1 hour.
- Urgent: Within 2 hours.
- Less urgent: Within 6 hours'.

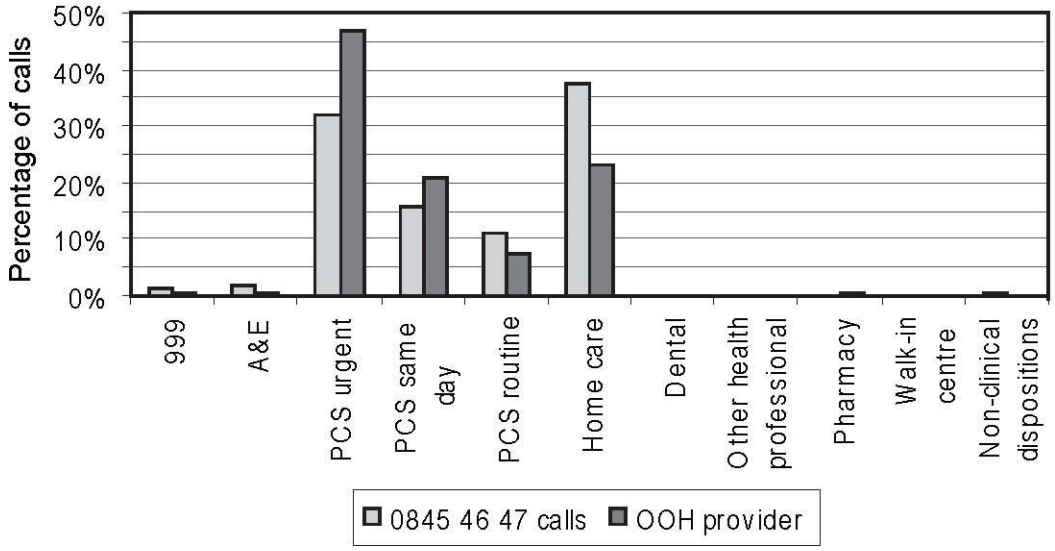
Further evidence was presented from NHS Direct that, in line with the out-of-hours Quality Requirements, currently recommends a time frame of less than 2 hours for a child requiring an urgent face-to-face assessment. Audit data from NHS Direct was presented to the GDG to show that, of those who contact NHS Direct via the 0845 telephone number, 31.8% of children under 5 years with a primary diagnosis of fever were referred on for an urgent face-to-face clinical assessment within 2 hours, following detailed nurse assessment (Figure 11.3). Also, 47% of out-of-hours calls for the same patient group were referred for a face-to-face clinical assessment within 2 hours. (It is important to note that during the course of these assessments a focus for the fever may be identified which in itself justified the referral within this time period.)

One stakeholder comment suggested that a 2 hour time limit for an urgent referral would be very difficult to implement in an A&E care setting where the 4 hour waiting time directive is the current target for the NHS. The guideline is clear that primary care should continue to be the first point of contact for a child with fever (as validated by the NHS Direct data presented here showing that children with fever are

referred to the GP within 2 hours, 6 hours or for a routine appointment). The GDG clarified that the new recommendation means that a child with 'red' features should be offered an *initial* assessment (for example, by an A&E triage nurse) within 2 hours, and that the current target of 4 hours for A&E is the time limit for initial assessment, treatment and discharge. The promise to patients derived from the NHS Plan in 2000 set out in *Your Guide to the NHS* stated that, on arrival in A&E, 'you should be assessed by a nurse or doctor, depending on how urgent your case is, within 15 minutes of arrival …'.[248]

These two waiting time targets are therefore compatible and in keeping with the Department of Health NHS Plan and Quality Requirements. Other stakeholders who commented on the 2 hour time frame felt that it was too long a wait for children requiring an urgent referral.

Figure 11.3 NHS Direct audit data covering the period 1 January 2006 to 31 December 2006; this data equates to a coverage of the whole of the population of England for the 0845 46 47 calls and a population coverage of 708,500 for the out-of-hours calls



The GDG believes that, if the traffic light system is adhered to, the recommendation for a 2 hour urgent referral will apply to a smaller but more relevant proportion of children with fever than are currently referred for an urgent assessment. A GDG member who is a GP presented evidence to the GDG from a survey of children presenting with fever as their predominant symptom and the prevalence of 'amber' features in this patient group. The practice has 9518 patients, with 633 children aged 5 years and under.

There were 157 consultations in this age group, involving 77 children with 83 episodes of acute fever with no other symptoms that worried the parent. Fifty-three episodes were telephone triage, and in 24 of these cases a face-to-face consultation was advised (45.2%). In thirteen of these cases, an 'amber' symptom was noted. The rest (104) were all face-to-face consultations without telephone triage, and in 18 consultations, 'amber' symptoms were recorded, with a diagnosis made in nine cases. Six of these children were referred for a paediatric assessment unit for specialist advice, which represents 3.8% of children presenting with fever as their primary symptom. During the period of the survey, there were no children who would have been classified as 'red' under the traffic light system.

Only 13 of those assessed remotely and 18 of those assessed face-to-face showed 'amber' features, and thus potentially none of these children fell into the urgent referral group. The absence of either 'red' or 'amber' features would have allowed at least some of these children to be confidently managed at home, and those with 'amber' features only could have been referred within a longer time frame of safety netting, which could have been put into place following face-to-face assessment. The data suggests that the proportion of children who require an urgent face-to-face referral following remote assessment would potentially be reduced and is very small compared with the far greater number of children who have either 'amber' symptoms and require assessment within a longer time frame by a healthcare professional or have self-limiting illness (who can be confidently managed at home).

Having reviewed the data and based on their own experience, the GDG consensus was that an individual GP in a group practice such as the one surveyed would be unlikely to see more than one or two cases of SBI a year, and for some of the more rare conditions would be unlikely to see one case in their professional career. During the period of the survey there were no children who would have been classified as 'red' under the traffic light system. This is because urgent referrals would only be needed for children with 'red' features and a proportion of children with 'amber' features. This assertion is

supported by the data in the GP survey referred to above where no children were classified as 'red' and 19% were 'amber'.

Further evidence of the number of children likely to present to secondary care with 'red' symptoms was considered. An American study of 6611 febrile children presenting to an emergency department found that 3.3% of children had a Yale Observation Score greater than 10.[101] A YOS score of 10 means the child has symptoms that are 'red' signs and symptoms on the proposed traffic light system. It is important to note that the 3.3% is a small fraction of the total number of children with fever but it still may be an overestimate because the data do not indicate how many of the 3.3% of children with a YOS score over 10 have other symptoms which are 'red' features in the traffic light system. Also, the study was done in a hospital setting and it is based on the American healthcare system. Furthermore, the GDG's recommendation would only apply to children referred from remote assessment in this context and not all children with 'red' symptoms, many of whom will present for a face-to-face clinical assessment as their first point of healthcare contact.

Cost-effectiveness of a 2 hour referral for face-to-face assessment

The GDG did not identify any data on the likely cost or cost savings from recommending a 2 hour time limit for an urgent face-to-face assessment or the likelihood of this leading to an increase in referrals to specialist care. The issue was discussed in detail during a number of GDG meetings. The main point that was agreed was that the GDG believes that the guideline's recommendations will support the identification of those children requiring urgent assessment, referral and initiation of management which in some cases will be life-saving and certainly prevent unnecessary long-term morbidity. There is a cost-effectiveness threshold under which any intervention that saves lives or prevents serious morbidity is generally seen to be cost-effective. If we assume that a life-saving intervention that prevents one death in a very young child is worth around 25 QALYs (75 years discounted at 3.5%), then an intervention that costs £500,000 (25 × £20,000) and saves one life is within the threshold for cost-effectiveness.

The GDG found it impossible to guess how many children with 'red' symptoms who were seen face- to-face urgently from a remote assessment (within 2 hours) would be saved from death or serious morbidity. The argument for cost-effectiveness is that £500,000 (to save one child's life) could be spent on additional face-to-face assessments for it to be cost-effective if it saved one life. The cost of additional face-to-face assessment is hard to estimate if it is within surgery hours, but it costs around £35–40 for an out-of-hours consultation [1, 249] or £70 for a home visit.[250] Therefore if an additional 7,100 (£500,000/£70) patients could be seen for face-to-face assessment, this would be cost-effective if it saved one additional child's life.

This does not take into account the potential savings from preventing the health and social care costs of serious morbidity in children which would make the intervention more cost-effective. Nor does it take into account that the carers of children with 'red' symptoms will contact health services somehow, and the guideline emphasises the fact that this should almost always be primary care in the first instance. This is a less expensive option than A&E services which cost £77–105 per visit for 2005/06, depending on the cost of investigations.[250]

This very brief analysis of cost-effectiveness assumes that at least three children's deaths are prevented every year in the district general hospital by putting in place a 2 hour assessment in a population of 250,000, and there are children are currently at risk of death and serious morbidity who are not currently being urgently assessed and referred for specialist advice. It also assumes that all children at risk of death from SBI are seen eventually by a healthcare professional, and do not die at home without any health service contact. It is assumed that deaths can be prevented by more timely referral to specialist services for those children who urgently need it, and that the cost of investigations and initial management once reaching a specialist care unit would be the same at whatever stage they were referred (that is, a standard package of investigations and management of a child with suspected SBI would be initiated).

Clearly there are costs around diagnosis and initial management of a child with suspected SBI once they reach specialist services, but the GDG was not clear that these would be any different (whether higher costs if a child is referred urgently or higher if referred after a delay of more than 2 hours). Without empirical data, these assumptions cannot be verified, but the GDG members believe that these are conservative assumptions that reflect the real world closely enough to make the assertion that the 2 hour face-to-face referral is very likely to be cost-effective.

Conclusion

[1] Annual cost or provision of out-of-hours care in England was £316 million in 2004–05, and the number of people using the service in England was 9 million.

The aim of this guideline is to improve the identification of those children who are genuinely at a high risk of serious illness and require urgent assessment and treatment to prevent death and serious morbidity. Using the traffic light system, those children in the ‘red’ category have been identified as being at a high risk of serious illness and the GDG believes that it is already established best clinical and cost-effective practice for this small group to be seen urgently within 2 hours and this guidance will reinforce that practice. The guideline will also reduce unnecessary assessment (urgent and routine) and diagnostic testing of children who are at low risk of serious illness.

12 References

This section was updated in 2013.

References [2013]

Akpede et al., 1992

Akpede,G.O., Sykes,R.M., Abiodun,P.O., Indications for lumbar puncture in children presenting with convulsions and fever of acute onset: experience in the Children's Emergency Room of the University of Benin Teaching Hospital, Nigeria, Annals of Tropical Paediatrics, 12, 385-389, 1992

Alpert et al., 1990

Alpert,G., Hibbert,E., Fleisher,G.R., Case-control study of hyperpyrexia in children, Pediatric Infectious Disease Journal, 9, 161-163, 1990

Andreola et al., 2007

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, Pediatric Infectious Disease Journal, 26, 672-677, 2007

Andreola et al., 2007a

Andreola,B., Bressan,S., Callegaro,S., Liverani,A., Plebani,M., Da,DaltL, Procalcitonin and C-reactive protein as diagnostic markers of severe bacterial infections in febrile infants and children in the emergency department, Pediatric Infectious Disease Journal, 26, 672-677, 2007

Autret et al., 1994

Autret,E., Breart,G., Jonville,A.P., Courcier,S., Lassale,C., Goehrs,J.M., Comparative efficacy and tolerance of ibuprofen syrup and acetaminophen syrup in children with pyrexia associated with infectious diseases and treated with antibiotics, European Journal of Clinical Pharmacology,Eur.J.Clin.Pharmacol., 46, 197-201, 1994

Autret et al., 1997

Autret,E., Reboul-Marty,J., Henry-Launois,B., Laborde,C., Courcier,S., Goehrs,J.M., Languillat,G., Launois,R., Evaluation of ibuprofen versus aspirin and paracetamol on efficacy and comfort in children with fever, European Journal of Clinical Pharmacology, 51, 367-371, 1997

Baker et al., 1987

Baker,M.D., Fosarelli,P.D., Carpenter,R.O., Childhood fever: correlation of diagnosis with temperature response to acetaminophen, Pediatrics, 80, 315-318, 1987

Baker et al., 1989

Baker,R.C., Seguin,J.H., Leslie,N., Gilchrist,M.J., Myers,M.G., Fever and petechiae in children, Pediatrics, 84, 1051-1055, 1989

Baker et al., 1989a

Baker,R.C., Tiller,T., Bausher,J.C., Bellet,P.S., Cotton,W.H., Finley,A.H., Lenane,A.M., McHenry,C., Perez,K.K., Shapiro,R.A., Severity of disease correlated with fever reduction in febrile infants, Pediatrics, 83, 1016-1019, 1989

Baker et al., 1990

Baker,M.D., Avner,J.R., Bell,L.M., Failure of infant observation scales in detecting serious illness in febrile, 4- to 8-week-old infants, Pediatrics, 85, 1040-1043, 1990

Bang & Chaturvedi, 2009

Bang,A., Chaturvedi,P., Yale Observation Scale for prediction of bacteremia in febrile children, Indian Journal of Pediatrics, 76, 599-604, 2009

Baskin et al., 1992

Baskin,M.N., O'Rourke,E.J., Fleisher,G.R., Outpatient treatment of febrile infants 28 to 89 days of age with intramuscular administration of ceftriaxone, Journal of Pediatrics, 120, 22-27, 1992

Batra et al., 2011

Batra,P., Gupta,S., Gomber,S., Saha,A., Predictors of meningitis in children presenting with first febrile seizures, Pediatric Neurology, 44, 35-39, 2011

Beasley et al., 2008

Beasley,R., Clayton,T., Crane,J., von,Mutius E., Lai,C.K., Montefort,S., Stewart,A., ISAAC Phase Three Study Group., Association between paracetamol use in infancy and childhood, and risk of asthma, rhinoconjunctivitis, and eczema in children aged 6-7 years: analysis from Phase Three of the ISAAC programme, Lancet, 372, 1039-1048, 2008

Berger et al., 1996

Berger,R.M., Berger,M.Y., van Steensel-Moll,H.A., Dzoljic-Danilovic,G., rksen-Lubsen,G., A predictive model to estimate the risk of serious bacterial infections in febrile infants, European Journal of Pediatrics, 155, 468-473, 1996

BergerRM et al., 1996

Berger RM, Berger MY, van Steensel-Moll HA, Dzoljic-Danilovic G, Derksen-Lubsen G., A predictivemodel to estimate the risk of serious bacterial infections in febrile infants., European Journal of Pediatrics, 155, 468-73, 1996

Bin et al., 2010

Bin,Salleeh H., McGillivray,D., Martin,M., Patel,H., Duration of fever affects the likelihood of a positive bag urinalysis or catheter culture in young children, Journal of Pediatrics, 156, 629-633, 2010

Bleeker et al., 2001

Bleeker,S.E., Moons,K.G., rksen-Lubsen,G., Grobbee,D.E., Moll,H.A., Predicting serious bacterial infection in young children with fever without apparent source, Acta Paediatrica, 90, 1226-1232, 2001

Bleeker et al., 2007

Bleeker,S.E., rksen-Lubsen,G., Grobbee,D.E., Donders,A.R., Moons,K.G., Moll,H.A., Validating and updating a prediction rule for serious bacterial infection in patients with fever without source, Acta Paediatrica, 96, 100-104, 2007

Bonadio et al., 1994

Bonadio,W.A., Smith,D.S., Sabnis,S., The clinical characteristics and infectious outcomes of febrile infants aged 8 to 12 weeks, Clinical Pediatrics, 33, 95-99, 1994

Brent et al., 2011

Brent,A.J., Lakhanpaul,M., Thompson,M., Collier,J., Ray,S., Ninis,N., Levin,M., MacFaul,R., Risk score to stratify children with suspected serious bacterial infection: Observational cohort study, Archives of Disease in Childhood, 96, 361-367, 2011

Brent et al., 2011a

Brent,A.J., Lakhanpaul,M., Ninis,N., Levin,M., MacFaul,R., Thompson,M., Evaluation of temperature-pulse centile charts in identifying serious bacterial illness: observational cohort study, Archives of Disease in Childhood, 96, 368-373, 2011

Brewer et al.1968

Brewer,E.J.,Jr., A comparative evaluation of indomethacin, acetaminophen and placebo as antipyretic agents in children, Arthritis and Rheumatism, 11, 645-651, 1968

Byington et al., 2002

Byington,C.L., Spencer,L.Y., Johnson,T.A., Pavia,A.T., Allen,D., Mason,E.O., Kaplan,S., Carroll,K.C., Daly,J.A., Christenson,J.C., Samore,M.H., An epidemiological investigation of a sustained high rate of pediatric parapneumonic empyema: risk factors and microbiological associations, Clinical Infectious Diseases, 34, 434-440, 2002

Chen et al., 2009

Chen,C.J., Lo,Y.F., Huang,M.C., Chung,R.L., Tang,R.B., Wu,K.G., A model for predicting risk of serious bacterial infection in febrile infants younger than 3 months of age, Journal of the Chinese Medical Association: JCMA, 72, 521-526, 2009

Craig et al., 2010

Craig,J.C., Williams,G.J., Jones,M., Codarini,M., Macaskill,P., Hayen,A., Irwig,L., Fitzgerald,D.A., Isaacs,D., McCaskill,M., 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, 340, c1594-, 2010

Crain & Shelov, 1982

Crain,E.F., Shelov,S.P., Febrile infants: predictors of bacteremia, Journal of Pediatrics, 101, 686-689, 1982

Crocker et al., 1985

Crocker,P.J., Quick,G., McCombs,W., Occult bacteremia in the emergency department: diagnostic criteria for the young febrile child, Annals of Emergency Medicine, 14, 1172-1177, 1985

Davies & Maconochie, 2009

Davies,P., Maconochie,I., The relationship between body temperature, heart rate and respiratory rate in children, Emergency Medicine Journal, 26, 641-643, 2009

Doran et al., 1989

Doran,T.F., De,Angelis C., Baumgardner,R.A., Mellits,E.D., Acetaminophen: more harm than good for chickenpox?, Journal of Pediatrics, 114, 1045-1048, 1989

Dubos et al., 2008

Dubos,F., Hue,V., Grandbastien,B., Catteau,B., Martinot,A., Bacterial skin infections in children hospitalized with varicella: a possible negative impact of non-steroidal anti-inflammatory drugs?, Acta Dermato-Venereologica, 88, 26-30, 2008

Enarson MC, Samina A, Vandermeer B, et al. Beliefs and expectations of Canadian parents who bring febrile children for medical care. Pediatrics 2012: 130 (4): 905-912

Erlewyn-Lajeunesse et al., 2006

Erlewyn-Lajeunesse,M.D.S., Coppens,K., Hunt,L.P., Chinnick,P.J., Davies,P., Higginson,I.M., Benger,J.R., Randomised controlled trial of combined paracetamol and ibuprofen for fever, Archives of Disease in Childhood, 91, 414-416, 2006

Factor et al., 2001

Factor,S.H., Schillinger,J.A., Kalter,H.D., Saha,S., Begum,H., Hossain,A., Hossain,M., Dewitt,V., Hanif,M., Khan,N., Perkins,B., Black,R.E., Schwartz,B., Diagnosis and management of febrile children using the WHO/UNICEF guidelines for IMCI in Dhaka, Bangladesh, Bulletin of the World Health Organization, 79, 1096-1105, 2001

Figueras et al., 2002

Figueras,Nadal C., Garcia de Miguel,M.J., Gomez,Campdera A., Pou,Fernandez J., Alvarez,Calatayud G., Sanchez,Bayle M., Paediatric Fever Co-operative Group from the Spanish Paediatric Association, Effectiveness and tolerability of ibuprofen-arginine versus paracetamol in children with fever of likely infectious origin, Acta Paediatrica,Acta Paediatr., 91, 383-390,2002

Fouzas et al., 2010

Fouzas,S., Mantagou,L., Skylogianni,E., Varvarigou,A., Reactive thrombocytosis in febrile young infants with serious bacterial infection, Indian Pediatrics, 47, 937-943, 2010

Francois et al., 2010

Francois,P., Desrumaux,A., Cans,C., Pin,I., Pavese,P., Labarere,J., Prevalence and risk factors of suppurative complications in children with pneumonia, Acta Paediatrica, International Journal of Paediatrics, 99, 861-866, 2010

Galetto-Lacour et al., 2003

Galetto-Lacour,A., Zamora,S.A., Gervaix,A., Bedside procalcitonin and C-reactive protein tests in children with fever without localizing signs of infection seen in a referral center, Pediatrics, 112, 1054-1060, 2003

Ghotbi & Shiva, 2009

Ghotbi,F., Shiva,F., An assessment of the necessity of lumbar puncture in children with seizure and fever, JPMA - Journal of the Pakistan Medical Association, 59, 292-295, 2009

Gomez et al., 2010

Gomez,B., Mintegi,S., Benito,J., Egireun,A., Garcia,D., Astobiza,E., Blood culture and bacteremia predictors in infants less than three months of age with fever without source, Pediatric Infectious Disease Journal, 29, 43-47, 2010

Gomez et al., 2012

Gomez,B., Mintegi,S., Rubio,M.C., Garcia,D., Garcia,S., Benito,J., Clinical and analytical characteristics and short-term evolution of enteroviral meningitis in young infants presenting with fever without source, Pediatric Emergency Care, 28, 518-523, 2012

Guen et al., 2007

Guen,C.G.-L., Delmas,C., Launay,E., Caillon,J., Loubersac,V., Picherot,G., Roze,C.J., Contribution of procalcitonin to occult bacteraemia detection in children, Scandinavian Journal of Infectious Diseases, 39, 157-159, 2007

Gupta et al., 2007

Gupta,H., Shah,D., Gupta,P., Sharma,K.K., Role of paracetamol in treatment of childhood Fever: a double-blind randomized placebo controlled trial, Indian Pediatrics, 44, 903-911, 2007

Haddon et al., 1999

Haddon,R.A., Barnett,P.L., Grimwood,K., Hogg,G.G., Bacteraemia in febrile children presenting to a paediatric emergency department., Medical Journal of Australia,Med.J.Aust., 170, 475-478, 1999

Hanna & Greenes, 2004

Hanna,Colleen M., Greenes,David S., How much tachycardia in infants can be attributed to fever?, Annals of emergency medicine,Ann Emerg Med, 43, 699-705, 2004

Hay et al., 2009

Hay,A.D., Redmond,N.M., Costelloe,C., Montgomery,A.A., Fletcher,M., Hollinghurst,S., Peters,T.J., Paracetamol and ibuprofen for the treatment of fever in children: the PITCH randomised controlled trial, Health Technology Assessment (Winchester, England), 13, 1-163, 2009

Hewson et al., 2000

Hewson,P., Poulakis,Z., Jarman,F., Kerr,J., McMaster,D., Goodge,J., Silk,G., Clinical markers of serious illness in young infants: a multicentre follow-up study, Journal of Paediatrics and Child Health, 36, 221-225, 2000

Hollinghurst et al., 2008

Hollinghurst,S., Redmond,N., Costelloe,C., Montgomery,A., Fletcher,M., Peters,T.J., Hay,A.D., Paracetamol plus ibuprofen for the treatment of fever in children (PITCH): economic evaluation of a randomised controlled trial, BMJ, 337, a1490-, 2008

Hsiao et al., 2006

Hsiao,A.L., Chen,L., Baker,M.D., Incidence and predictors of serious bacterial infections among 57- to 180-day-old infants, Pediatrics, 117, 1695-1701, 2006

Hugenholtz M, Bröer C, Van Daalen R. Apprehensive parents: a qualitative study of parents seeking immediate primary care for their children. British Journal of General Practice 2009:59 (560):173-9

Isaacman & Burke, 2002

Isaacman,D.J., Burke,B.L., Utility of the serum C-reactive protein for detection of occult bacterial infection in children, Archives of Pediatrics and Adolescent Medicine, 156, 905-909, 2002

Joffe et al., 1983

Joffe,A., McCormick,M., DeAngelis,C., Which children with febrile seizures need lumbar puncture? A decision analysis approach, American Journal of Diseases of Children, 137, 1153-1156, 1983

Kauffman et al., 1992

Kauffman,R.E., Sawyer,L.A., Scheinbaum,M.L., Antipyretic efficacy of ibuprofen vs acetaminophen, American Journal of Diseases of Children, 146, 622-625, 1992

Kai J. What worries parents when their preschool children are acutely ill, and why: a qualitative study. British Medical Journal 1996: 313:983-6

Kai J. Parents' difficulties and information needs in coping with acute illness in preschool children: A qualitative study. British Medical Journal 1996: 313 (7063): 987-990

Karwowska A, Nijssen-Jordan C, Johnson D, et al. Parental and health care provider understanding of childhood fever: a Canadian perspective. Canadian Journal of Emergency Medicine 2002;4(6):394–400.

Kramer et al., 2008

Kramer,L.C., Richards,P.A., Thompson,A.M., Harper,D.P., Fairchok,M.P., Alternating antipyretics: antipyretic efficacy of acetaminophen versus acetaminophen alternated with ibuprofen in children, Clinical Pediatrics, 47, 907-911, 2008

Lacour et al., 2001

Lacour,A.G., Gervaix,A., Zamora,S.A., Vadas,L., Lombard,P.R., Dayer,J.M., Suter,S., Procalcitonin, IL-6, IL-8, IL-1 receptor antagonist and C-reactive protein as identificators of serious bacterial infections in children with fever without localising signs, European Journal of Pediatrics, 160, 95-100, 2001

Luaces-Cubells et al., 2012

Luaces-Cubells,C., Mintegi,S., Garcia-Garcia,J.J., Astobiza,E., Garrido-Romero,R., Velasco-Rodriguez,J., Benito,J., Procalcitonin to detect invasive bacterial infection in non-toxic-appearing infants with fever without apparent source in the emergency department, Pediatric Infectious Disease Journal, 31, 645-647, 2012

Mandl et al., 1997

Mandl,K.D., Stack,A.M., Fleisher,G.R., Incidence of bacteremia in infants and children with fever and petechiae, Journal of Pediatrics, 131, 398-404, 1997

Maniaci et al., 2008

Maniaci,V., Dauber,A., Weiss,S., Nylen,E., Becker,K.L., Bachur,R., Procalcitonin in young febrile infants for the detection of serious bacterial infections, Pediatrics, 122, 701-710, 2008

Manzano et al., 2011

Manzano,S., Bailey,B., Gervaix,A., Cousineau,J., Delvin,E., Girodias,J.B., Markers for bacterial infection in children with fever without source, Archives of Disease in Childhood, 96, 440-446, 2011

Mazur et al., 1989

Mazur,L.J., Jones,T.M., Kozinetz,C.A., Temperature response to acetaminophen and risk of occult bacteremia: a case-control study, Journal of Pediatrics, 115, 888-891, 1989

McCarthy et al., 1980

McCarthy,P.L., Jekel,J.F., Stashwick,C.A., Spiesel,S.Z., Dolan,T.F.,Jr., History and observation variables in assessing febrile children, Pediatrics, 65, 1090-1095, 1980

McCarthy et al., 1981

McCarthy,P.L., Jekel,J.F., Stashwick,C.A., Spiesel,S.Z., Dolan,T.F., Sharpe,M.R., Forsyth,B.W., Baron,M.A., Fink,H.D., Rosenbloom,M.L., Etkin,T., Zelson,J.H., Further definition of history and observation variables in assessing febrile children, Pediatrics, 67, 687-693, 1981

McCarthy et al., 1982

McCarthy,P.L., Sharpe,M.R., Spiesel,S.Z., Dolan,T.F., Forsyth,B.W., DeWitt,T.G., Fink,H.D., Baron,M.A., Cicchetti,D.V., Observation scales to identify serious illness in febrile children, Pediatrics, 70, 802-809, 1982

McCarthy et al., 1985

McCarthy,P.L., Lembo,R.M., Baron,M.A., Fink,H.D., Cicchetti,D.V., Predictive value of abnormal physical examination findings in ill-appearing and well-appearing febrile children, Pediatrics, 76, 167-171, 1985

McIntyre & Hull, 1996

McIntyre,J., Hull,D., Comparing efficacy and tolerability of ibuprofen and paracetamol in fever, Archives of Disease in Childhood, 74, 164-167, 1996

Mikaeloff et al., 2008

Mikaeloff,Y., Kezouh,A., Suissa,S., Nonsteroidal anti-inflammatory drug use and the risk of severe skin and soft tissue complications in patients with varicella or zoster disease.[Erratum appears in Br J Clin Pharmacol. 2010 Jun;69(6):722], British Journal of Clinical Pharmacology, 65, 203-209, 2008

Morris et al., 2007

Morris,C.M., Tefuarani,N., Ripa,P., Laki,R., Vince,J.D., Urinary tract infection in infants and young children presenting with fever without a focus in Port Moresby, Papua New Guinea Medical Journal, 50, 145-151, 2007

Nabulsi et al., 2006

Nabulsi,M.M., Tamim,H., Mahfoud,Z., Itani,M., Sabra,R., Chamseddine,F., Mikati,M., Alternating ibuprofen and acetaminophen in the treatment of febrile children: a pilot study [ISRCTN30487061], BMC Medicine, 4, 4-, 2006

Nademi et al., 2001

Nademi,Z., Clark,J., Richards,C.G., Walshaw,D., Cant,A.J., The causes of fever in children attending hospital in the north of England, Journal of Infection, 43, 221-225, 2001

Nahata et al., 1992

Nahata,M.C., Powell,D.A., Durrell,D.E., Miller,M.A., Gupta,N., Efficacy of ibuprofen in pediatric patients with fever, International Journal of Clinical Pharmacology, Therapy, and Toxicology, 30, 94- 96, 1992

Newman et al., 2002

Newman,T.B., Bernzweig,J.A., Takayama,J.I., Finch,S.A., Wasserman,R.C., Pantell,R.H., Urine testing and urinary tract infections in febrile infants seen in office settings: the Pediatric Research in Office Settings' Febrile Infant Study, Archives of Pediatrics and Adolescent Medicine, 156, 44-54, 2002

Nielsen et al., 2001

Nielsen,H.E., Andersen,E.A., Andersen,J., Bottiger,B., Christiansen,K.M., Daugbjerg,P., Larsen,S.O., Lind,I., Nir,M., Olofsson,K., Diagnostic assessment of haemorrhagic rash and fever, Archives of Disease in Childhood, 85, 160-165, 2001

Nijman et al., 2011

Nijman,R.G., Zwinkels,R.L., van,Veen M., Steyerberg,E.W., van der,Lei J., Moll,H.A., Oostenbrink,R., Can urgency classification of the Manchester triage system predict serious bacterial infections in febrile children?, Archives of Disease in Childhood, 96, 715-722, 2011

Offringa et al., 1992

Offringa,M., Beishuizen,A., rksen-Lubsen,G., Lubsen,J., Seizures and fever: can we rule out meningitis on clinical grounds alone?, Clinical Pediatrics, 31, 514-522, 1992

Olaciregui et al., 2009

Olaciregui,I., Hernandez,U., Munoz,J.A., Emparanza,J.I., Landa,J.J., Markers that predict serious bacterial infection in infants under 3 months of age presenting with fever of unknown origin, Archives of Disease in Childhood, 94, 501-505, 2009

Owusu-Ofori et al., 2004

Owusu-Ofori,A., Agbenyega,T., Ansong,D., Scheld,W.M., Routine lumbar puncture in children with febrile seizures in Ghana: should it continue?, International Journal of Infectious Diseases, 8, 353- 361, 2004

Pantell et al., 2004

Pantell,R.H., Newman,T.B., Bernzweig,J., Bergman,D.A., Takayama,J.I., Segal,M., Finch,S.A., Wasserman,R.C., Management and outcomes of care of fever in early infancy, JAMA, 291, 1203- 1212, 2004

Pashapour et al., 2009

Pashapour,N., Macooei,A.A., Golmobammadlou,S., Alternating ibuprofen and acetaminophen in the treatment of febrile hospitalized children aged 9-24 months, Iranian Journal of Pediatrics, 19, 164- 168, 2009

Paul et al., 2010

Paul,I.M., Sturgis,S.A., Yang,C., Engle,L., Watts,H., Berlin,C.M.,Jr., Efficacy of standard doses of Ibuprofen alone, alternating, and combined with acetaminophen for the treatment of febrile children, Clinical Therapeutics, 32, 2433-2440, 2010

Pierce & Voss, 2010

Pierce,C.A., Voss,B., Efficacy and safety of ibuprofen and acetaminophen in children and adults: a meta-analysis and qualitative review. [93 refs], Annals of Pharmacotherapy, 44, 489-506, 2010

Pratt & Attia, 2007

Pratt,A., Attia,M.W., Duration of fever and markers of serious bacterial infection in young febrile children, Pediatrics International, 49, 31-35, 2007

Pulliam et al., 2001

Pulliam,P.N., Attia,M.W., Cronan,K.M., C-reactive protein in febrile children 1 to 36 months of age with clinically undetectable serious bacterial infection, Pediatrics, 108, 1275-1279, 2001

Rabasa & Gofama, 2009

Rabasa,A.I., Gofama,M.M., Urinary tract infection in febrile children in Maiduguri north eastern Nigeria, Nigerian Journal of Clinical Practice, 12, 124-127, 2009

Rudinsky et al., 2009

Rudinsky,S.L., Carstairs,K.L., Reardon,J.M., Simon,L.V., Riffenburgh,R.H., Tanen,D.A., Serious bacterial infections in febrile infants in the post-pneumococcal conjugate vaccine era, Academic Emergency Medicine, 16, 585-590, 2009

Sands R, Shanmugavadivel D, Stephenson T, et al. Medical problems presenting to paediatric emergency departments: 10 years on. Emergency Medical Journal 2011: 29 (5): 379-82

Sarrell et al., 2006

Sarrell,E.M., Wielunsky,E., Cohen,H.A., Antipyretic treatment in young children with fever: acetaminophen, ibuprofen, or both alternating in a randomized, double-blind study, Archives of Pediatrics and Adolescent Medicine, 160, 197-202, 2006

Schwartz et al., 2009

Schwartz,S., Raveh,D., Toker,O., Segal,G., Godovitch,N., Schlesinger,Y., A week-by-week analysis of the low-risk criteria for serious bacterial infection in febrile neonates, Archives of Disease in Childhood, 94, 287-292, 2009

Shaw et al., 1998

Shaw,K.N., Gorelick,M., McGowan,K.L., Yakscoe,N.M., Schwartz,J.S., Prevalence of urinary tract infection in febrile young children in the emergency department, Pediatrics, 102, e16-, 1998

Shettigar et al., 2011

Shettigar,C.G., Rao,D., Hegde,P., Soans,S., Routine urine culture in febrile young children, Journal of Clinical and Diagnostic Research, 5, 452-455, 2011

Shin et al., 2009

Shin,S.H., Choi,C.W., Lee,J.A., Kim,E.K., Choi,E.H., Kim,H.S., Kim,B.I., Choi,J.H., Risk factors for serious bacterial infection in febrile young infants in a community referral hospital, Journal of Korean Medical Science, 24, 844-848, 2009

Sidler et al., 1990

Sidler,J., Frey,B., Baerlocher,K., A double-blind comparison of ibuprofen and paracetamol in juvenile pyrexia, British Journal of Clinical Practice, Supplement. 70, 22-25, 1990

Singhi et al., 1992

Singhi,S., Kohli,V., Ayyagiri,A., Bacteremia and bacterial infections in highly febrile children without apparent focus, Indian Pediatrics,Indian Pediatr., 29, 1285-1289, 1992

Southey et al., 2009

Southey,E.R., Soares-Weiser,K., Kleijnen,J., Systematic review and meta-analysis of the clinical safety and tolerability of ibuprofen compared with paracetamol in paediatric pain and fever. [77 refs], Current Medical Research and Opinion, 25, 2207-2222, 2009

Stanley et al., 2005

Stanley,R., Pagon,Z., Bachur,R., Hyperpyrexia among infants younger than 3 months, Pediatric Emergency Care, 21, 291-294, 2005

Stathakis et al., 2007

Stathakis,T., Acworth,J.P., Barnett,A.G., Prediction tool for bacteraemia in children aged 3-36 months, Emergency Medicine Australasia, 19, 353-358, 2007

Sugimura et al., 1994

Sugimura,T., Fujimoto,T., Motoyama,H., Maruoka,T., Korematu,S., Asakuno,Y., Hayakawa,H., Risks of antipyretics in young children with fever due to infectious disease, Acta Paediatrica Japonica, 36, 375-378, 1994

Tal et al., 1997

Tal,Y., Even,L., Kugelman,A., Hardoff,D., Srugo,I., Jaffe,M., The clinical significance of rigors in febrile children, European Journal of Pediatrics, 156, 457-459, 1997

Taveras EM, Durousseau S, Flores G. Parents' beliefs and practices regarding childhood fever: a study of a multiethnic and socioeconomically diverse sample of parents. Pediatric Emergency Care 2004: 20 (9): 579-587

Taylor et al., 1995

Taylor,J.A., Del,Beccaro M., Done,S., Winters,W., Establishing clinically relevant standards for tachypnea in febrile children younger than 2 years, Archives of Pediatrics and Adolescent Medicine, 149, 283-287, 1995

Teach & Fleisher, 1995

Teach,S.J., Fleisher,G.R., Efficacy of an observation scale in detecting bacteremia in febrile children three to thirty-six months of age, treated as outpatients. Occult Bacteremia Study Group, Journal of Pediatrics,J.Pediatr., 126, 877-881, 1995

Teach & Fleisher, 1997

Teach,S.J., Fleisher,G.R., Duration of fever and its relationship to bacteremia in febrile outpatients three to 36 months old. The Occult Bacteremia Study Group, Pediatric Emergency Care, 13, 317-319, 1997

Teele et al., 1975

Teele,D.W., Pelton,S.I., Grant,M.J., Herskowitz,J., Rosen,D.J., Allen,C.E., Wimmer,R.S., Klein,J.O., 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, Journal of Pediatrics, 87, 227-230, 1975

Thayyil et al., 2005

Thayyil,S., Shenoy,M., Hamaluba,M., Gupta,A., Frater,J., Verber,I.G., Is procalcitonin useful in early diagnosis of serious bacterial infections in children?, Acta Paediatrica, 94, 155-158, 2005

Thompson et al., 2009

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?, Archives of Disease in Childhood, 94, 888-893, 2009

Thompson et al., 2009a

Thompson,M., Harnden,A., Perera,R., Mayon-White,R., Smith,L., McLeod,D., Mant,D., Deriving temperature and age appropriate heart rate centiles for children with acute infections, Archives of Disease in Childhood, 94, 361-365, 2008

Thompson et al., 2009b

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?, Archives of Disease in Childhood, 94, 888-893, 2009

Thompson at al., 2012

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, Systematic review and validation of prediction rules for identifying children with serious infections in emergency departments and urgent- access primary care, Health Technol Assess;16(15):1–100, 2012

Torrey et al., 1985

Torrey,S.B., Henretig,F., Fleisher,G., Goldstein,R.M., Ardire,A., Ludwig,S., Ruddy,R., Temperature response to antipyretic therapy in children: relationship to occult bacteremia, American Journal of Emergency Medicine, 3, 190-192, 1985

Trautner et al., 2006

Trautner,B.W., Caviness,A.C., Gerlacher,G.R., Demmler,G., Macias,C.G., 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, 118, 34-40, 2006

Ulukol et al., 1999

Ulukol,B., Koksal,Y., Cin,S., Assessment of the efficacy and safety of paracetamol, ibuprofen and nimesulide in children with upper respiratory tract infections, European Journal of Clinical Pharmacology, 55, 615-618, 1999

utret-Leca et al., 2007

utret-Leca,E., Gibb,I.A., Goulder,M.A., Ibuprofen versus paracetamol in pediatric fever: objective and subjective findings from a randomized, blinded study, Current Medical Research and Opinion, 23, 2205-2211, 2007

Van et al., 1984

Van,Nguyen Q., Nguyen,E.A., Weiner,L.B., Incidence of invasive bacterial disease in children with fever and petechiae, Pediatrics, 74, 77-80, 1984

Van et al., 1995

Van,Esch A., Van Steensel-Moll,H.A., Steyerberg,E.W., Offringa,M., Habbema,J.D., rksen-Lubsen,G., Antipyretic efficacy of ibuprofen and acetaminophen in children with febrile seizures, Archives of Pediatrics and Adolescent Medicine, 149, 632-637, 1995

Vauzelle-Kervroedan et al., 1997

Vauzelle-Kervroedan,F., d'Athis,P., Pariente-Khayat,A., Debregeas,S., Olive,G., Pons,G., Equivalent antipyretic activity of ibuprofen and paracetamol in febrile children, Journal of Pediatrics, 131, 683- 687, 1997

Walson et al., 1989

Walson,P.D., Galletta,G., Braden,N.J., Alexander,L., Ibuprofen, acetaminophen, and placebo treatment of febrile children, Clinical Pharmacology and Therapeutics, 46, 9-17, 1989

Walson et al., 1992

Walson,P.D., Galletta,G., Chomilo,F., Braden,N.J., Sawyer,L.A., Scheinbaum,M.L., Comparison of multidose ibuprofen and acetaminophen therapy in febrile children, American Journal of Diseases of Children,Am.J.Dis.Child., 146, 626-632, 1992

Walson, 1990

Walson,P.D., Ibuprofen versus paracetamol for the treatment of fever in children, British Journal of Clinical Practice, Supplement. 70, 19-21, 1990

Weber et al., 2003

Weber,M.W., Carlin,J.B., Gatchalian,S., Lehmann,D., Muhe,L., Mulholland,E.K., WHO Young Infants Study Group., Predictors of neonatal sepsis in developing countries, Pediatric Infectious Disease Journal, 22, 711-717, 2003

Weisse et al., 1987

Weisse,M.E., Miller,G., Brien,J.H., Fever response to acetaminophen in viral vs. bacterial infections, Pediatric Infectious Disease Journal, 6, 1091-1094, 1987

Wells et al., 2001

Wells,L.C., Smith,J.C., Weston,V.C., Collier,J., Rutter,N., The child with a non-blanching rash: how likely is meningococcal disease?, Archives of Disease in Childhood, 85, 218-222, 2001

Wilson et al., 1991

Wilson,J.T., Brown,R.D., Kearns,G.L., Eichler,V.F., Johnson,V.A., Bertrand,K.M., Lowe,B.A., Single-dose, placebo-controlled comparative study of ibuprofen and acetaminophen antipyresis in children, Journal of Pediatrics, 119, 803-811, 1991

Woelker et al., 2012

Woelker,J.U., Sinha,M., Christopher,N.C., Powell,K.R., Serum procalcitonin concentration in the evaluation of febrile infants 2 to 60 days of age, Pediatric Emergency Care, 28, 410-415, 2012

Wong et al., 2001

Wong,A., Sibbald,A., Ferrero,F., Plager,M., Santolaya,M.E., Escobar,A.M., Campos,S., Barragan,S., De Leon,Gonzalez M., Kesselring,G.L., Fever Pediatric Study Group, Antipyretic effects of dipyrone versus ibuprofen versus acetaminophen in children: results of a multinational, randomized, modified double-blind study, Clinical Pediatrics,Clin.Pediatr., 40, 313-324, 2001

Yamamoto et al., 1987

Yamamoto,L.T., Wigder,H.N., Fligner,D.J., Rauen,M., Dershewitz,R.A., Relationship of bacteremia to antipyretic therapy in febrile children, Pediatric Emergency Care, 3, 223-227, 1987

Yeboah-Antwi et al., 2008

Yeboah-Antwi,K., ddo-Yobo,E., du-Sarkodie,Y., Carlin,J.B., Plange-Rhule,G., Osei,Akoto A., Weber,M.W., Hamer,D.H., Clinico-epidemiological profile and predictors of severe illness in young infants (0-59 days) in Ghana, Annals of Tropical Paediatrics, 28, 35-43, 2008

Young Infants Clinical Signs Study Group, 2008

Young Infants Clinical Signs Study Group, Clinical signs that predict severe illness in children under age 2 months: a multicentre study, Lancet, 371, 135-142, 2008

Zarkesh et al., 2011

Zarkesh,M., Hashemian,H., Momtazbakhsh,M., Rostami,T., Assessment of febrile neonates according to low risk criteria for serious bacterial infection, Iranian Journal of Pediatrics, 21, 436-440, 2011

Zorc et al., 2005

Zorc,J.J., Levine,D.A., Platt,S.L., Dayan,P.S., Macias,C.G., Krief,W., Schor,J., Bank,D., Shaw,K.N., Kuppermann,N., Multicenter RSV-SBI Study Group of the Pediatric Emergency Medicine Collaborative Research Committee of the American Academy of Pediatrics., Clinical and demographic factors associated with urinary tract infection in young febrile infants, Pediatrics, 116, 644-648, 2005

References [2007]

1. Hay AD. The prevalence of symptoms and consultations in pre-school children in the Avon Longitudinal Study of Parents and Children (ALSPAC): A prospective cohort study. *Family Practice* 2005;22(4):367–74.

2. Singhal A, Morley R, Abbott R, et al. Clinical safety of iron-fortified formulas. *Pediatrics* 2000;105:e38.

3. Saxena S, Majeed A, Jones M. Socio-economic differences inchildhood consultation rates in general practice in England and Wales: prospective cohort study. *British Medical Journal* 1999;318:642–6.

4. McCormick A, Fleming D, Charlton J. Morbidity statistics from general practice: fourth national

study 1991–1992. London: HMSO; 1995.

5. Dale J, Crouch R, Lloyd D. Primary care: nurse-led telephone triage and advice out-of-hours. *Nursing Standard* 1998;12(47):41–5.

6. Armon K, Stephenson T, Gabriel V, et al. Audit: Determining the common medical presenting problems to an accident and emergency department. *Archives of Disease in Childhood* 2001;84(5):390–2.

7. Stewart M, Werneke U, MacFaul R, et al. Medical and social factors associated with the admission and discharge of acutely ill children. *Archives of Disease in Childhood* 1998;79(3):219–24.

8. Ishimine I. Fever without source in children 0 to 36 months of age. *Pediatric Clinics of North America* 2006;53:167–94.

9. Baraff LJ. Management of fever without source in infants and children.*Annals of Emergency Medicine* 2000;36(6):602–14.

10. Herz AM, Greenhow TL, Alcantara J, et al. Changing epidemiology of outpatient bacteremia in 3 to 36 month old children after the introduction of the heptavalent-conjugated pneumococcal vaccine. *Pediatric Infectious Disease Journal* 2006;25(4):293–300.

11. Ninis N, Phillips C, Bailey L, et al. The role of healthcare delivery in the outcome of meningococcal disease in children: case–control study of fatal and non-fatal cases. *British Medical Journal* 2005;330(7506):1475.

12. Heyderman RS, Ben-Shlomo Y, Brennan CA, *et al.* The incidence and mortality for meningococcal disease associated with area deprivation: an ecological study of hospital episode statistics. *Archives of Disease in Childhood* 2004;89:1064–8.

13. Kai J. What worries parents when their preschool children are acutely ill, and why: a qualitative study. *British Medical Journal* 1996;313(7063):983–6.

14. Karwowska A, Nijssen-Jordan C, Johnson D, *et al.* Parental and health care provider understanding of childhood fever: a Canadian perspective. *Canadian Journal of Emergency Medicine* 2002;4(6):394–400.

15. Department of Health. *National Service Framework for Children, Young People and Maternity Services – Core Standards*. London: Department of Health; 2004.

16. Mandell GL, Bennett JE, Dolin R, editors.*Mandell, Douglas and Bennett's Principles and Practice of Infectious Diseases,* 6th ed. Philadelphia: Churchill Livingstone; 2004.

17. National Institute for Clinical Excellence. *Guideline Development Methods: Information for National Collaborating Centres and Guideline Developers*. London: National Institute for Clinical Evidence; 2005.

18. Oxman AD, Sackett DL, Guyatt GH. Users' guide to the medical literature. I. How to get started. *JAMA: the journal of the American Medical Association* 1993;270(17):2093–5.

19. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group.*JAMA: the journal of the American Medical Association* 1993;270(21):2598–601.

20. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group.*JAMA: the journal of the American Medical Association* 1994;271(1):59–63.

21. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group.*JAMA: the journal of the American Medical Association* 1994;271(5):389–91.

22. 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: the journal of the American Medical Association* 1994;271(9):703–7.

23. Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. *Evidence-Based Medicine.How to Practice and Teach EBM*. 3rd ed. Edinburgh: Churchill Livingstone; 2005.

24. Scottish Intercollegiate Guidelines Network. *SIGN 50: A Guideline developers' handbook*. No. 50. Edinburgh: Scottish Intercollegiate Guideline Network; 2001.

25. Wilczynski NL, Haynes RB. Developing optimal search strategies for detecting clinically sound prognostic studies in MEDLINE: an analytic survey.*BMC Medicine* 2004;2:23.

26. Drummond MF, Sculpher M, Torrance GW, O'Brien BJ, Stoddart GL. *Methods for the Economic Evaluation of Health Care Programmes*. 3rd ed. Oxford University Press; 2005.

27. Crawford D, Greene N, Wentworth S. *Thermometer Review: UK Market Survey*. No. MHRA 04144. 2005.

28. Health and Safety Executive. *Control of Substances Hazardous to Health*. 2007.

29. Craig JV, Lancaster GA, Williamson PR, *et al.* Temperature measured at the axilla compared with rectum in children and young people: Systematic review. *British Medical Journal* 2000;320(7243):1174–8.

30. Craig JV, Lancaster GA, Taylor S, *et al.* Infrared ear thermometry compared with rectal thermometry in children: A systematic review. *Lancet* 2002;360(9333):603–9.

31. El-Radhi AS, Barry W. Thermometry in paediatric practice. *Archives of Disease in Childhood* 2006;91(4):351–6.

32. Pickersgill J, Fowler H, Boothman J, *et al.* Temperature taking: children's preferences. *Paediatric Nursing* 2003;15(2):22–5.

33. Banco L, Jayashekaramurthy S, Graffam J. The inability of a temperature-sensitive pacifier to identify fevers in ill infants.*American Journal of Diseases of Children* 1988;142(2):171–2.

34. Beckstrand RL, Wilshaw R, Moran S, *et al.*Supralingual temperatures compared to tympanic and rectal temperatures. *Pediatric Nursing* 1996;22(5):436–8.

35. Morley CJ, Hewson PH, Thornton AJ, *et al.* Axillary and rectal temperature measurements in infants. *Archives of Disease in Childhood* 1992;67(1):122–5.

36. Bliss-Holtz J. Comparison of rectal, axillary, and inguinal temperatures in full-term newborn infants. *Nursing Research* 1989;38(2):85–7.

37. Shann F, Mackenzie A. Comparison of rectal, axillary, and forehead temperatures. *Archives of Pediatrics and Adolescent Medicine* 1996;150(1):74–8.

38. Saxena AK, Topp SS, Heinecke A, *et al.* Application criteria for infrared ear thermometers in pediatric surgery. *Technology and Health Care* 2001;9(3):281–5.

39. Osinusi K, Njinyam MN. Comparison of body temperatures taken at different sites and the reliability of axillary temperature in screening for fever.*African Journal of Medicine and Medical Sciences* 1997;26(3–4):163–6.

40. Muma BK, Treloar DJ, Wurmlinger K, *et al.* Comparison of rectal, axillary, and tympanic membrane temperatures in infants and young children. *Annals of Emergency Medicine* 1991;20(1):41–4.

41. Chaturvedi D, Vilhekar KY, Chaturvedi P, *et al.* Comparison of axillary temperature with rectal or oral temperature and determination of optimum placement time in children. *Indian Pediatrics* 2004;41(6):600–3.

42. Anagnostakis D, Matsaniotis N, Grafakos S, *et al.* Rectal-axillary temperature difference in febrile and afebrile infants and children. *Clinical Pediatrics* 1993;32(5):268–72.

43. Jirapaet V, Jirapaet K. Comparisons of tympanic membrane, abdominal skin, axillary, and rectal temperature measurements in term and preterm neonates. *Nursing and Health Sciences* 2000;2(1):1–8.

44. Falzon A, Grech V, Caruana B, *et al.* How reliable is axillary temperature measurement? *ActaPaediatrica* 2003;92(3):309–13.

45. Zengeya ST, Blumenthal I. Modern electronic and chemical thermometers used in the axilla are inaccurate. *European Journal of Pediatrics* 1996;155(12):1005–8.

46. Akinbami FO, Sowunmi A. Body temperature in the Nigerian neonate – comparison of axillary and rectal temperatures. *African Journal of Medicine and Medical Sciences* 1991;20(1):49–52.

47. Haddock BJ, Merrow DL, Swanson MS. The falling grace of axillary temperatures.*Pediatric Nursing* 1996;22(2):121–5.

48. Lodha R, Mukerji N, Sinha N, *et al.* Is axillary temperature an appropriate surrogate for core temperature? *Indian Journal of Pediatrics* 2000;67(8):571–4.

49. Buntain WL, Pregler M, O'Brien PC, *et al.* Axillary versus rectal temperature: a comparative study. *Journal of the Louisiana State Medical Society* 1977;129(1):5–8.

50. Upadhyay M, Singhi S, Murlidharan J, *et al.* Randomized evaluation of fluid resuscitation with crystalloid (saline) and colloid (polymer from degraded Gelatin in saline) in pediatric septic shock. *Indian Pediatrics* 2005;42(3):223–31.

51. Barrus DH. A comparison of rectal and axillary temperatures by electronic thermometer measurement in preschool children.*Pediatric Nursing* 1983;9(6):424–5.

52. Weisse ME, Reagen MS, Boule L, *et al.* Axillary vs. rectal temperatures in ambulatory and hospitalized children. *Pediatric Infectious Disease Journal* 1991;10(7):541–2.

53. Brown PJ, Christmas BF, Ford RP. Taking an infant's temperature: axillary or rectal thermometer? *New Zealand Medical Journal* 1992;105(939):309–11.

54. Jean-Mary MB, Dicanzio J, Shaw J, *et al.* Limited accuracy and reliability of infrared axillary and aural thermometers in a pediatric outpatient population. *Journal of Pediatrics* 2000;141(5):671–6.

55. Childs C, Harrison R, Hodkinson C. Tympanic membrane temperature as a measure of core temperature. *Archives of Disease in Childhood* 1999;80(3):262–6.

56. Ogren JM. The inaccuracy of axillary temperatures measured with an electronic thermometer. *American Journal of Diseases of Children* 1990;144(1):109–11.

57. Messmer PR, Rodriguez S, Adams J, *et al.* Effect of kangaroo care on sleep time for neonates. *Pediatric Nursing* 1997;23(4):408–14.

58. Postma CT, de BT, Roelofs A, *et al.* [The ear thermometer: not a good replacement for the rectal thermometer] [Dutch]. *NederlandsTijdschriftvoorGeneeskunde* 1999;143(4):222–3.

59. Morley C, Murray M, Whybrew K. The relative accuracy of mercury, Tempa-DOT and FeverScan thermometers.*Early Human Development* 1998;53(2):171–8.

60. Leick-Rude MK, Bloom LF. A comparison of temperature-taking methods in neonates.*Neonatal Network – Journal of Neonatal Nursing* 1998;17(5):21–37.

61. Pejaver RK, Nisarga R, Gowda B. Temperature monitoring in newborns using thermospot. *Indian Journal of Pediatrics* 2004;71(9):795–6.

62. Scholefield JM, Gerber MA, Dwyer P. Liquid crystal forehead temperature strips.A clinical appraisal.*American Journal of Diseases of Children* 1982;136(3):198–201.

63. Valadez JJ, Elmore-Meegan M, Morley D. Comparing liquid crystal thermometer readings and mercury thermometer readings of infants and children in a traditional African setting. Implications for community-based health. Tropical and Geographical Medicine 1995;47(3):130–3.

64. Dart RC, Lee SC, Joyce SM, et al. Liquid crystal thermometry for continuous temperature measurement in emergency department patients. Annals of Emergency Medicine 1985;14(12):1188–90.

65. Fawcett J. The accuracy and reliability of the tympanic membrane thermometer.A literature review. Emergency Nurse 2001;8(9):13–7.

66. Kenney RD, Fortenberry JD, Surratt SS, et al. Evaluation of an infrared tympanic membrane thermometer in pediatric patients. Pediatrics 1990;85(5):854–8.

67. Akinyinka OO, Omokhodion SI, Olawuyi JF, et al. Tympanic thermometry in Nigerian children. Annals of Tropical Paediatrics 2001;21(2):169–74.

68. Davis K. The accuracy of tympanic temperature measurement in children.Pediatric Nursing 1993;19(3):267–72.

69. Yetman RJ, Coody DK, West MS, *et al.* Comparison of temperature measurements by an aural infrared thermometer with measurements by traditional rectal and axillary techniques. *Journal of Pediatrics* 1993;122(5 Pt 1):769–73.

70. Mayfield SR, Nakamura KT, Bhatia J, *et al.* Tympanic membrane temperature of term and preterm neonates. *Early Human Development* 1984;9(3):241–7.

71. Stewart JV, Webster D. Re-evaluation of the tympanic thermometer in the emergency department. *Annals of Emergency Medicine* 1992;21(2):158–61.

72. Lanham DM, Walker B, Klocke E, *et al.* Accuracy of tympanic temperature readings in children under 6 years of age. *Pediatric Nursing* 1999;25(1):39–42.

73. Talo H, Macknin ML, Medendorp SV. Tympanic membrane temperatures compared to rectal and oral temperatures. *Clinical Pediatrics* 1991;30(4 Suppl):30–3.

74. Rogers J, Curley M, Driscoll J, *et al.* Evaluation of tympanic membrane thermometer for use with pediatric patients. *Pediatric Nursing* 1991;17(4):376–8.

75. Rhoads FA, Grandner J. Assessment of an aural infrared sensor for body temperature measurement in children. *Clinical Pediatrics* 1990;29(2):112–5.

76. Pransky SM. The impact of technique and conditions of the tympanic membrane upon infrared tympanic thermometry.*Clinical Pediatrics* 1991;30(4 Suppl):50–2.

77. Bernardo LM, Clemence B, Henker R, *et al.*A comparison of aural and rectal temperature measurements in children with moderate and severe injuries.*Journal of Emergency Nursing* 1996;22(5):403–8.

78. Selfridge J, Shea SS. The accuracy of the tympanic membrane thermometer in detecting fever in infants aged 3 months and younger in the emergency department setting.*Journal of Emergency Nursing* 1993;19(2):127–30.

79. Brennan DF, Falk JL, Rothrock SG, *et al.* Reliability of infrared tympanic thermometry in the detection of rectal fever in children. *Annals of Emergency Medicine* 1995;25(1):21–30.

80. Loveys AA, Dutko-Fioravanti I, Eberly SW, *et al.* Comparison of ear to rectal temperature measurements in infants and toddlers. *Clinical Pediatrics* 1999;38(8):463–6.

81. Petersen-Smith A, Barber N, Coody DK, *et al.* Comparison of aural infrared with traditional rectal temperatures in children from birth to age three years. *Journal of Pediatrics* 1994;125(1):83–5.

82. Sehgal A, Dubey NK, Jyothi MC, *et al.* Comparison of tympanic and rectal temperature in febrile patients. *Indian Journal of Pediatrics* 2002;69(4):305–8.

83. El-Rahdi AS, Patel S. An evaluation of tympanic thermometry in a paediatric emergency department.*Emergency Medicine Journal* 2006;23(1):40–1.

84. Schuh S, Komar L, Stephens D, *et al.* Comparison of the temporal artery and rectal thermometry in children in the emergency department. *Pediatric Emergency Care* 2004;20(11):736–41.

85. Crawford DC, Hicks B, Thompson MJ. Which thermometer? Factors influencing best choice for intermittent clinical temperature assessment.*Journal of Medical Engineering and Technology* 2006;30(4):199–211.

86. Banco L, Veltri D. Ability of mothers to subjectively assess the presence of fever in their children. *American Journal of Diseases of Children* 1984;138(10):976–8.

87. Hooker EA, Smith SW, Miles T, *et al.* Subjective assessment of fever by parents: comparison with measurement by noncontact tympanic thermometer and calibrated rectal glass mercury thermometer. *Annals of Emergency Medicine* 1996;28(3):313–7.

88. Nwanyanwu OC, Ziba C, Redd SC, *et al.* Palpation as a method of fever determination in Malawian children who are less than 5 years old: how reliable is it? *Annals of Tropical Medicine and Parasitology* 1997;91(4):359–63.

89. Singhi S, Sood V. Reliability of subjective assessment of fever by mothers. *Indian Pediatrics* 1990;27(8):811–5.

90. Ernst TN, Philp M. Temperature assessment by parental palpation. *American Journal of Diseases of Children* 1985;139(6):546–7.

91. BezerraAlves JG, De Barros CJ. Ability of mothers to assess the presence of fever in their children without using a thermometer.*Tropical Doctor* 2002;32(3):145–6.

92. Biarent D, Bingham R, Richmond S, *et al.* European Resuscitation Council guideline for resuscitation 2005. Section 6.Paediatric life support.*Resuscitation* 2005;67S1:S97-S133.

93. Hewson P, Poulakis Z, Jarman F, *et al.* Clinical markers of serious illness in young infants: a multicentre follow-up study. *Journal of Paediatrics and Child Health* 2000;36(3):221–5.

94. Baker RC, Seguin JH, Leslie N, *et al.* Fever and petechiae in children. *Pediatrics* 1989;84(6):1051–5.

95. Crain EF, Shelov SP. Febrile infants: Predictors of bacteremia. *Journal of Pediatrics* 1982;101(5):686–9.

96. Mandl KD, Stack AM, Fleisher GR. Incidence of bacteremia in infants and children with fever and petechiae. *Journal of Pediatrics* 1997;131(3):398–404.

97. Hiew TM, Tan AM, Cheng HK. Clinical features and haematological indices of bacterial infections in young infants.*Singapore Medical Journal* 1992;33(2):125–30.

98. Weber MW, Carlin JB, Gatchalian S, *et al.* Predictors of neonatal sepsis in developing countries. *Pediatric Infectious Disease Journal* 2003;22(8):711–7.

99. Ronfani L, Vilarim JNA, Dragovich D, *et al.* Signs of severe bacterial infection in neonates. *Journal of Tropical Pediatrics* 1999;45(1):48–51.

100. McCarthy PL, Sharpe MR, Spiesel SZ. Observation scales to identify serious illness in febrile children. *Pediatrics* 1982;70(5):802–9.

101. Teach SJ, Fleisher GR. Efficacy of an observation scale in detecting bacteremia in febrile children three to thirty-six months of age, treated as outpatients. Occult Bacteremia Study Group. *Journal of Pediatrics* 1995;126(6):877–81.

102. McCarthy PL, Lembo RM, Fink HD. Observation, history, and physical examination in diagnosis of serious illnesses in febrile children <=24 months. *Journal of Pediatrics* 1987;110(1):26–30.

103. McCarthy PL, Lembo RM, Baron MA, *et al.* Predictive value of abnormal physical examination findings in ill-appearing and well-appearing febrile children. *Pediatrics* 1985;76(2):167–71.

104. 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(6):1040–3.

105. Jamuna R, Srinivasan S, Harish BN. Factors predicting occult bacteremia in young children. *Indian Journal of Pediatrics* 2000;67(10):709–11.

106. Bonadio WA, Hennes H, Smith D, *et al.* Reliability of observation variables in distinguishing infectious outcome of febrile young infants. *Pediatric Infectious Disease Journal* 1993;12(2):111–4.

107. Bonadio WA, Hagen E, Rucka J, *et al.* Efficacy of a protocol to distinguish risk of serious bacterial infection in the outpatient evaluation of febrile young infants. *Clinical Pediatrics* 1993;32(7):401–4.

108. Dagan R, Powell KR, Hall CB, *et al.* Identification of infants unlikely to have serious bacterial infection although hospitalized for suspected sepsis. *Journal of Pediatrics* 1985;107(6):855–60.

109. Jaskiewicz JA, McCarthy CA, Richardson AC, *et al.* Febrile infants at low risk for serious bacterial infection--an appraisal of the Rochester criteria and implications for management. Febrile Infant Collaborative Study Group.*Pediatrics* 1994;94(3):390–6.

110. Resuscitation Council (UK). *Resuscitation Guidelines 2005*. London: Resuscitation Council (UK); 2005.

111. Hanna CM, Greenes DS. How much tachycardia in infants can be attributed to fever? *Annals of Emergency Medicine* 2004;43(6):699–705.

112. Dark P, Woodford M, Vail A, Mackway-Jones K, Yates D, Lecky F. Systolic hypertension and the response to blunt trauma in infants and children. *Resuscitation* 2002;54(3):245–53.

113. Leonard PA, Beattie TF. Is measurement of capillary refill time useful as part of the initial assessment of children? *European Journal of Emergency Medicine* 2004;11(3):158–63.

114. Gorelick MH, Shaw KN, Murphy KO, *et al.* Effect of fever on capillary refill time. *Pediatric Emergency Care* 1997;13(5):305–7.

115. Otieno H, Were E, Ahmed I, *et al.* Are bedside features of shock reproducible between different observers? *Archives of Disease in Childhood* 2004;89(10):977–9.

116. Tibby SM, Hatherill M, Murdoch IA. Capillary refill and core-peripheral temperature gap as indicators of haemodynamic status in paediatric intensive care patients. *Archives of Disease in Childhood* 1999;80(2):163–6.

117. Steiner MJ, DeWalt DA, Byerley JS. Is this child dehydrated? *JAMA: the journal of the American Medical Association* 2004;291(22):2746–54.

118. Wells LC, Smith JC, Weston VC, *et al.* The child with a non-blanching rash: How likely is meningococcal disease? *Archives of Disease in Childhood* 2001;85(3):218–22.

119. Riordan FAI, Marzouk O, Thomson APJ, *et al.* Prospective validation of the glasgow meningococcal septicaemia prognostic score. Comparison with other scoring methods.*European Journal of Pediatrics* 2002;161(10):531–7.

120. Pantell RH, Newman TB, Bernzweig J, *et al.* Management and outcomes of care of fever in early infancy. *JAMA: the journal of the American Medical Association* 2004;291(10):1203–12.

121. Nademi Z, Clark J, Richards CG, *et al.*The causes of fever in children attending hospital in the north of England.*Journal of Infection* 2001;43:(4)221–5.

122. Teach SJ, Fleisher GR. Duration of fever and its relationship to bacteremia in febrile outpatients three to 36 months old. The Occult Bacteremia Study Group.*Pediatric Emergency Care* 1997;13(5):317–9.

123. Haddon RA, Barnett PL, Grimwood K, *et al.* Bacteraemia in febrile children presenting to a pae3iatric emergency department. *Medical Journal of Australia* 1999;170(10):475–8.

124. Teele DW, Pelton SI, Grant MJ, *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. *Journal of Pediatrics* 1975;87(2):227–30.

125. Caspe WB, Chamudes O, Louie B. The evaluation and treatment of the febrile infant.*Pediatric Infectious Disease* 1983;2(2):131–5.

126. Hsiao AL, Chen L, Baker MD. Incidence and predictors of serious bacterial infections among 57- to 180-day-old infants.*Pediatrics* 2006;117(5):1695–701.

127. Singhi S, Kohli V, Ayyagiri A. Bacteremia and bacterial infections in highly febrile children without apparent focus. *Indian Pediatrics* 1992;29(10):1285–9.

128. Ray JG. Screening and active management reduced perinatal complications more than

routine care in gestational diabetes. *ACP Journal Club* 2005;143(3):65.

129. Berger RM, Berger MY, Van Steensel-Moll HA, *et al.*A predictive model to estimate the risk of serious bacterial infections in febrile infants.*European Journal of Pediatrics* 1996;155(6):468–73.

130. Trautner BW, Caviness AC, Gerlacher GR, *et al.* 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(1):34–40.

132. Nielsen HE, Andersen EA, Andersen J, *et al.* Diagnostic assessment of haemorrhagic rash and fever. *Archives of Disease in Childhood* 2001;85(2):160–5.

133. Thompson MJ, Ninis N, Perera R, *et al.* Clinical recognition of meningococcal disease in children and adolescents. *Lancet* 2006;367(9508):397–403.

134. Walsh-Kelly C, Nelson DB, Smith DS, *et al.* Clinical predictors of bacterial versus aseptic meningitis in childhood. *Annals of Emergency Medicine* 1992;21(8):910–4.

135. Oostenbrink R, Moons KG, Derksen-Lubsen AG, Grobbee DE, Moll HA. A diagnostic decision rule for management of children with meningeal signs.[erratum appears in *Eur J Epidemiol* 2004;19(12):1137 Note: Moons, Carl GM [corrected to Moons, Karel GM]]. *European Journal of Epidemiology* 2004;19(2):109–16.

136. Tunkel AR, Scheld WM. Acute bacterial meningitis. *Lancet* 1995;346(8991–8992):1675–80.

137. Chin RF, Neville BG, Peckham C, *et al.* Incidence, cause, and short-term outcome of convulsive status epilepticus in childhood: prospective population-based study. *Lancet* 2006;368(9531):222–9.

138. Kennedy PG. A retrospective analysis of forty-six cases of herpes simplex encephalitis seen in Glasgow between 1962 and 1985. *Quarterly Journal of Medicine* 1988;68(255):533–40.

139. Mahabee-Gittens EM, Grupp-Phelan J, Brody AS, *et al.* Identifying children with pneumonia in the emergency department. *Clinical Pediatrics* 2005;44(5):427–35.

140. Taylor JA, Del BM, Done S, *et al.* Establishing clinically relevant standards for tachypnea in febrile children younger than 2 years. *Archives of Pediatrics and Adolescent Medicine* 1995;149(3):283–7.

141. Lucero MG, Tupasi TE, Gomez ML, *et al.* Respiratory rate greater than 50 per minute as a clinical indicator of pneumonia in Filipino children with cough. *Reviews of Infectious Diseases* 1990;12(Suppl 8):S1081–3.

142. Gupta D, Mishra S, Chaturvedi P. Fast breathing in the diagnosis of pneumonia – a reassessment. *Journal of Tropical Pediatrics* 1996;42(4):196–9.

143. Shamo'on H, Hawamdah A, Haddadin R, *et al.* Detection of pneumonia among children under six years by clinical evaluation. *Eastern Mediterranean Health Journal* 2004;10(4–5):482–7.

144. Redd SC, Vreuls R, Metsing M, *et al.* Clinical signs of pneumonia in children attending a hospital outpatient department in Lesotho. *Bulletin of the World Health Organization* 1994;72(1):113–8.

145. Kocher MS, Mandiga R, Zurakowski D, *et al.* Validation of a clinical prediction rule for the differentiation between septic arthritis and transient synovitis of the hip in children. *Journal of Bone and Joint Surgery – American Volume* 2004;86-A:(8)1629–35.

146. Kao HC, Huang YC, Chiu CH, *et al.* Acute hematogenous osteomyelitis and septic arthritis in children. *Journal of Microbiology, Immunology and Infection* 2003;36:(4)260–5.

147. Razak M, Ismail MM, Omar A. A review of haematogenous osteomyelitis in children in Kuala Lumpur Hospital.*Medical Journal of Malaysia* 1998;53Suppl A:83–5.

148. Akinyoola AL, Obiajunwa PO, Oginni LM. Septic arthritis in children.*West African Journal of Medicine* 2006;25(2):119–23.

149. Tseng C-F, Fu Y-C, Fu L-S, *et al.* Clinical spectrum of Kawasaki disease in infants. *Chinese Medical Journal* 2001;64(3):168–73.

150. Huang GY, Ma XJ, Huang M, *et al.* Epidemiologic pictures of Kawasaki disease in Shanghai from 1998 through 2002. *Journal of Epidemiology* 2006;16(1):9–14.

151. Neighbour R. The inner consultation. 1st edn. Lancaster, MTP; 1987.

152. Swingler GH, Zwarenstein M. Chest radiograph in acute respiratory infections in children.

(Cochrane Review). In: *Cochrane Database of Systematic Reviews*, Issue 3, 2005. Oxford: Update Software.

153. Feder HM Jr. Occult pneumococcal bacteremia and the febrile infant and young child. *Clinical Pediatrics*1980;19(7):457–62.

154. Goh PL, Lee SW, Wong EH. Predictors of serious bacterial infection in children aged 3 to 36 months with fever without source. *Singapore Medical Journal* 2006;47(4):276–80.

155. Rothrock SG, Harper MB, Green SM, *et al.* Do oral antibiotics prevent meningitis and serious bacterial infections in children with *streptococcus pneumoniae* occult bacteremia? A meta-analysis.*Pediatrics* 1997; 99(3):438–44.

156. Rothrock SG, Green SM, Harper MB, *et al.* Parenteral vs oral antibiotics in the prevention of serious bacterial infections in children with *Streptococcus pneumoniae* occult bacteremia: a meta-analysis. *Academic Emergency Medicine* 1998;5(6):599–606.

157. Damoiseaux RA, van Balen FA, Hoes AW, *et al.* Primary care based randomised, double blind trial of amoxicillin versus placebo for acute otitis media in children aged under 2 years. *British Medical Journal* 2000;320(7231):350–4.

158. Seppala H, Klaukka T, Vuopio-Varkila J, *et al.* The effect of changes in the consumption of macrolide antibiotics on erythromycin resistance in group A streptococci in Finland. *New England Journal of Medicine* 1997;337(7):441–7.

159. Hahne SJ, Charlett A, Purcell B, *et al.* Effectiveness of antibiotics given before admission in reducing mortality from meningococcal disease: systematic review. *British Medical Journal* 2006;332(7553):1299–303.

160. Harnden A, Ninis N, Thompson M, *et al.* Parenteral penicillin for children with meningococcal disease before hospital admission: case–control study. *British Medical Journal* 2006;332(7553):1295–8.

161. Martin D, Kieft C, Miller J. The epidemiology of meningococcal disease in New Zealand in 1998. 1999.

162. Baker RC, Tiller T, Bausher JC, *et al.* Severity of disease correlated with fever reduction in febrile infants. *Pediatrics* 1989;83(6):1016–9.

163. Baraff LJ, Oslund SA, Schriger DL, *et al.* Probability of bacterial infections in febrile infants less than three months of age: A meta-analysis. *Pediatric Infectious Disease Journal* 1992;11(4):257–65.

164. Dagan R, Sofer S, Philip M, *et al.* Ambulatory care of febrile infants younger than 2 months of age classified as being at low risk for having serious bacterial infections. *Journal of Pediatrics* 1988;112(3):355–60.

165. van RossumAM, Wulkan RW, Oudesluys-Murphy AM. Procalcitonin as an early marker of infection in neonates and children.*Lancet Infectious Diseases* 2004;4(10):620–30.

166. Carrol ED, Newland P, Riordan FAI, *et al.*Procalcitonin as a diagnostic marker of meningococcal disease in children presenting with fever and a rash. *Archives of Disease in Childhood* 2002;86(4):282–5.

167. Thayyil S, Shenoy M, Hamaluba M, *et al.* Is procalcitonin useful in early diagnosis of serious bacterial infections in children? *ActaPaediatrica* 2005;94(2):155–8.

168. Kohli V, Singhi S, Sharma P, *et al.* Value of serum C-reactive protein concentrations in febrile children without apparent focus. *Annals of Tropical Paediatrics* 1993;13(4):373–8.

169. Pulliam PN, Attia MW, Cronan KM. C-reactive protein in febrile children 1 to 36 months of age with clinically undetectable serious bacterial infection. *Pediatrics* 2001;108(6):1275–9.

170. Isaacman DJ, Burke BL. Utility of the serum C-reactive protein for detection of occult bacterial infection in children. *Archives of Pediatrics and Adolescent Medicine* 2002;156(9):905–9.

171. Fernandez LA, Luaces CC, Garcia Garcia JJ, *et al.*Procalcitonin in pediatric emergency departments for the early diagnosis of invasive bacterial infections in febrile infants: Results of a multicenter study and utility of a rapid qualitative test for this marker. *Pediatric Infectious Disease Journal* 2003;22(10 Suppl):895–903.

172. Gendrel D, Raymond J, Coste J, *et al.* Comparison of procalcitonin with C-reactive protein, interleukin 6 and interferon-alpha for differentiation of bacterial vs. viral infections. *Pediatric Infectious Disease Journal* 1999;18(10):875–81.

173. Lembo RM, Marchant CD. Acute phase reactants and risk of bacterial meningitis among febrile infants and children.*Annals of Emergency Medicine* 1991;20(1):36–40.

174. Moulin F, Raymond J, Lorrot M, *et al.*Procalcitonin in children admitted to hospital with community acquired pneumonia. *Archives of Disease in Childhood* 2001;84(4):332–6.

175. Lee GM, Harper MB. Risk of bacteremia for febrile young children in the post-*Haemophilusinfluenzae* type b era.*Archives of Pediatrics and Adolescent Medicine* 1998;152(7):624–8.

176. Bachur R, Perry H, Harper MB. Occult pneumonias: empiric chest radiographs in febrile children with leukocytosis. *Annals of Emergency Medicine* 1999;33(2):166–73.

178. Galetto-Lacour A, Zamora SA, Gervaix 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(5):1054–60.

179. Swingler GH. Radiologic differentiation between bacterial and viral lower respiratory infection in children: a systematic literature review. *Clinical Pediatrics* 2000;39(11):627–33.

180. Virkki R, Juven T, Rikalainen H, *et al.* Differentiation of bacterial and viral pneumonia in children. *Thorax* 2002;57(5):438–41.

181. Titus MO, Wright SW. Prevalence of serious bacterial infections in febrile infants with respiratory syncytial virus infection. *Pediatrics* 2003;112(2):282–4.

182. Smitherman HF, Caviness AC, Macias CG. Retrospective review of serious bacterial infections in infants who are 0 to 36 months of age and have influenza A infection. *Pediatrics* 2005;115(3):710–8.

183. Purcell K, Fergie J. Concurrent serious bacterial infections in 912 infants and children hospitalized for treatment of respiratory syncytial virus lower respiratory tract infection. *Pediatric Infectious Disease Journal* 2004;23(3):267–9.

184. Weisse ME, Miller G, Brien JH. Fever response to acetaminophen in viral vs. bacterial infections.*Pediatric Infectious Disease Journal* 1987;6(12):1091–4.

185. Torrey SB, Henretig F, Fleisher G, *et al.* Temperature response to antipyretic therapy in children: relationship to occult bacteremia. *American Journal of Emergency Medicine* 1985;3(3):190–2.

186. Yamamoto LT, Wigder HN, Fligner DJ, *et al.* Relationship of bacteremia to antipyretic therapy in febrile children.*Pediatric Emergency Care* 1987;3(4):223–7.

187. Baker MD, Fosarelli PD, Carpenter RO. Childhood fever: correlation of diagnosis with temperature response to acetaminophen. *Pediatrics* 1987;80(3):315–8.

188. Richardson AC, Roghmann KJ, White KC. Use of clinical observation scales following antipyretic therapy to predict serious illness in febrile children. *American Journal of Diseases of Children* 1999;144:435.

189. Filicori M, Flamigni C, Dellai P, *et al.* Treatment of anovulation with pulsatile gonadotropin-releasing hormone: prognostic factors and clinical results in 600 cycles. *Journal of Clinical Endocrinology and Metabolism* 1994;79(4):1215–20.

190. Oates-Whitehead RM, Maconochie I, Baumer H, Stewart MER. Fluid therapy for acute bacterial meningitis. (Cochrane Review). In: *Cochrane Database of Systematic Reviews*, Issue 3, 2006. Oxford: Update Software.

191. Anonymous. Human albumin administration in critically ill patients: systematic review of randomised controlled trials. Cochrane Injuries Group Albumin Reviewers.*British Medical Journal* 1998;317(7153):235–40.

192. Finfer S, Bellomo R, Boyce N, *et al.*A comparison of albumin and saline for fluid resuscitation in the intensive care unit.*New England Journal of Medicine* 2004;350(22):2247–56.

193. Duke T, Mokela D, Frank D, *et al.* Management of meningitis in children with oral fluid restriction or intravenous fluid at maintenance volumes: a randomised trial. *Annals of Tropical Paediatrics* 2002;22(2):145–57.

194. Han YY, Carcillo JA, Dragotta MA, *et al.* Early reversal of pediatric-neonatal septic shock by community physicians is associated with improved outcome. *Pediatrics* 2003;112(4):793–9.

195. van de Beek D, de Gans J, McIntyre P, Prasad K. Corticosteroids for acute bacterial meningitis. (Cochrane Review). In: *Cochrane Database of Systematic Reviews*, Issue 3, 2006.

Oxford: Update Software.

196. Ibrahim EH, Sherman G, Ward S, *et al.*The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting.*Chest* 2000;118(1):146–55.

197. Whitley RJ, Alford CA, Hirsch MS, *et al.*Vidarabine versus acyclovir therapy in herpes simplex encephalitis. *New England Journal of Medicine* 1986;314(3):144–9.

198. Whitley R, Arvin A, Prober C, *et al.*A controlled trial comparing vidarabine with acyclovir in neonatal herpes simplex virus infection.*New England Journal of Medicine* 1991;324(7):444–9.

199. Kimberlin D, Lin CY, Jacobs RF, *et al.* Safety and efficacy of high-dose intravenous acyclovir in the management of neonatal herpes simplex virus infections. *Pediatrics* 2001;108(2):230–8.

200. Raschilas F, Wolff M, Delatour F, *et al.* Outcome of and prognostic factors for herpes simplex encephalitis in adult patients: results of a multicenter study. *Clinical Infectious Diseases* 2002;35(3):254–60.

201. Osman O, Brown D, Beattie T, *et al.* Management of febrile children in a paediatric emergency department. *Health Bulletin* 2002;60(1):33–9.

202. Laundy M, jayi-Obe E, Hawrami K, *et al.* Influenza A community-acquired pneumonia in East London infants and young children. *Pediatric Infectious Disease Journal* 2003;22(10 Suppl):S223–7.

203. Richardson MP, Reid A, Tarlow MJ, *et al.* Hearing loss during bacterial meningitis [erratum appears in *Arch Dis Child* 1997 Apr;76(4):386]. *Archives of Disease in Childhood* 1997;76(2):134–8.

204. Department of Health. *Hospital Episode Statistics* [www.hesoline.nhs.uk]. 2006.

205. Boulant JA. Thermoregulation. In: Mackowik PA, ed. *Fever, Basic Mechanisms and Management*. Philadelphia: Lippincott-Raven; 1997. p. 35–58.

206. Crocetti M, Moghbeli N, Serwint J. Fever phobia revisited: have parental misconceptions about fever changed in 20 years? *Pediatrics* 2001;107(6):1241–6.

207. Nesse RM, Williams GC. *Evolution and Healing*. London: Phoenix; 1994.

208. Axelrod P. External cooling in the management of fever. *Clinical Infectious Diseases* 2000;31(Suppl 5):S224–9.

209. Purssell E. Physical treatment of fever. *Archives of Disease in Childhood* 2000;82(3):238–9.

210. Perrott DA, Piira T, Goodenough B, *et al.* Efficacy and safety of acetaminophen vs ibuprofen for treating children's pain or fever: A Meta-analysis. *Archives of Pediatrics and Adolescent Medicine* 2004;158(6):521–6.

211. Kinmonth AL, Fulton Y, Campbell MJ. Management of feverish children at home.*British Medical Journal* 1992;305(6862):1134–6.

212. Lenhardt R, Negishi C, Sessler DI, *et al.* The effects of physical treatment on induced fever in humans. *American Journal of Medicine* 1999;106:550–5.

213. Meremikwu M, Oyo-Ita A. Physical methods for treating fever in children. (Cochrane Review). In: *Cochrane Database of Systematic Reviews*, Issue 1, 2007. Oxford: Update Software.

214. Purssell E. Treating fever in children: paracetamol or ibuprofen? *British Journal of Community Nursing* 2002;7(6):316–20.

215. Wong A, Sibbald A, Ferrero F, *et al.* Antipyretic effects of dipyrone versus ibuprofen versus acetaminophen in children: results of a multinational, randomized, modified double-blind study. *Clinical Pediatrics* 2001;40(6):313–24.

216. Figueras NC, Garcia de Miguel MJ, Gomez CA, *et al.* Effectiveness and tolerability of ibuprofen-arginine versus paracetamol in children with fever of likely infectious origin. *ActaPaediatrica* 2002;91(4):383–90.

217. Walson PD, Galletta G, Chomilo F, *et al.* Comparison of multidose ibuprofen and acetaminophen therapy in febrile children. *American Journal of Diseases of Children* 1992;146(5):626–32.

218. Autret E, Breart G, Jonville AP, *et al.* Comparative efficacy and tolerance of ibuprofen syrup and acetaminophen syrup in children with pyrexia associated with infectious diseases and treated with antibiotics. *European Journal of Clinical Pharmacology* 1994;46(3):197–201.

219. Heubi JE, Barbacci MB, Zimmerman HJ. Therapeutic misadventures with acetaminophen: hepatoxicity after multiple doses in children. *Journal of Pediatrics* 1998;132(1):22–7.

220. Mayoral CE, Marino RV, Rosenfeld W, *et al.* Alternating antipyretics: is this an alternative? *Pediatrics* 2000;105(5):1009–12.

221. Erlewyn-Lajeunesse MD, Coppens K, Hunt LP, *et al.* Randomised controlled trial of combined paracetamol and ibuprofen for fever. *Archives of Disease in Childhood* 2006;91(5):414–6.

222. Lal A, Gomber S, Talukdar B. Antipyretic effects of nimesulide, paracetamol and ibuprofen-paracetamol. *Indian Journal of Pediatrics* 2000;67(12):865–70.

223. Sarrell EM, Wielunsky E, Cohen HA. Antipyretic treatment in young children with fever: acetaminophen, ibuprofen, or both alternating in a randomized, double-blind study. *Archives of Pediatrics and Adolescent Medicine* 2006;160(2):197–202.

224. Nabulsi MM, Tamim H, Mahfoud Z, *et al.* Alternating ibuprofen and acetaminophen in the treatment of febrile children: a pilot study. *BMC Medicine* 2006;4:4.

225. Del Vecchio MT, Sundel ER. Alternating antipyretics: is this an alternative? *Pediatrics* 2001;108(5):1236–7.

226. Brandts CH, Ndjave M, Graninger W, *et al.* Effect of paracetamol on parasite clearance time in *Plasmodium falciparum* malaria. *Lancet* 1997;350(9079):704–9.

227. Doran TF, De Angelis C, Baumgardner RA, *et al.* Acetaminophen: more harm than good for chickenpox? *Journal of Pediatrics* 1989;114(6):1045–8.

228. Kramer MS, Naimark LE, Roberts-Brauer R, *et al.* Risks and benefits of paracetamolantipyresis in young children with fever of presumed viral origin. *Lancet* 1991;337(8741):591–4.

230. Verity CM, Greenwood R, Golding J. Long-term intellectual and behavioral outcomes of children with febrile convulsions. *New England Journal of Medicine* 1998;338(26):1723–8.

231. Andre P, Thebaud B, Guibert M, *et al.* Maternal-fetal staphylococcal infections: A series report. *American Journal of Perinatology* 2000;17(8):423–8.

232. Meremikwu M, Oyo-Ita A. Paracetamol for treating fever in children. (Cochrane Review). In: *Cochrane Database of Systematic Reviews*, Issue 3, 2005. Oxford: Update Software.

233. Guppy MPB, Mickan SM, Del Mar CB. Advising patients to increase fluid intake for treating acute respiratory infections. (Cochrane Review). In: *Cochrane Database of Systematic Reviews*, Issue 3, 2006. Oxford: Update Software.

234. NICE. Making group decisions and reaching consensus. In: NICE. *The Guidelines Manual*. London: National Institute for Health and Clinical Excellence; 2006. ch. 9.

235. Murphy MK, Black NA, Lamping DL, *et al.* Consensus development methods and their use in clinical guideline development. *Health Technology Assessment* 1998;2(3):1–88.

236. Black N, Murphy M, Lamping D, *et al.* Consensus development methods: a review of best practice in creating clinical guidelines. *Journal of Health Services Research & Policy* 1999;4:236–48.

237. Raine R, Sanderson C, Black N. Improving clinical guideline development: a challenge to current methods. *BMJ* 2005;331:631–3.

238. Fitch K, Bernstein S, Aguilar MS, Burnand B, LaCalle JR, Van het Loo M, McDonnell J, *et al.The RAND/UCLA Appropriateness Method User's Manual*. MR-1269-DG-XII/RE. Rand Corporation;2000;.

239. Raine R, Hutchings A. A systematic review of factors affecting the judgments produced by formal consensus development methods in health care.*Journal of Health Services Research & Policy* 2006;11(3):172–9.

240. Elwyn G, Greenhalgh T, Macfarlane F. Groups: A guide to small groups. In: *Healthcare, Management, Education and Research*. Abingdon: Radcliffe Medical Press; 2001.

241. Elwyn G, O'Connor A, Stacey D, *et al.*Developing a quality criteria framework for patient decision aids: online international Delphi consensus process. *BMJ* 2006:333;417–22.

242. Curtis L, Netten A. *Unit Costs of Health & Social Care 2005*. PSSRU University of Kent; 2005.

243. Boeckx W, Gordts S, Buysse K, *et al.* Reversibility after female sterilization. *BJOG: An International Journal of Obstetrics & Gynaecology* 1986;93(8):839–42.

244. Baraff LJ. Outpatient management of fever in selected infants.*New England Journal of Medicine* 1994;330(13):938–9.

245. Lacour AG, Gervaix A, Zamora SA, *et al.*Procalcitonin, IL-6, IL-8, IL-1 receptor antagonist and C-reactive protein as identificators of serious bacterial infections in children with fever without localising signs. *European Journal of Pediatrics* 2001;160(2):95–100.

246. Armon K, MacFaul R, Hemingway P, *et al.* The impact of presenting problem based guidelines for children with medical problems in an accident and emergency department. *Archives of Disease in Childhood* 2004;89(2):159–64.

247. Department of Health. *National Quality Requirements in the Delivery of Out-of-Hours Services*.Gateway no. 3776. 2004.

248. NHS. *Your Guide to the NHS*. 2001.

249. National Audit Office. *Report on the provision of out of hours care. The Provision of Out-of-Hours Care in England*. Report by the Comptroller and Auditor General. No. HC1041. 2006.

250. Curtis L, Netten A. *Unit Costs of Health and Social Care*. Personal and Social Services Research Unit University of Kent at Canterbury; 2006.

251. Press S, Quinn BJ. The pacifier thermometer.Comparison of supralingual withrectal temperatures in infants and young children.*Archives of Pediatrics and Adolescent Medicine* 1997;151(6):551–4.

252. Dodd SR, Lancaster GA, Craig JV, *et al.* In a systematic review, infrared ear thermometry for fever diagnosis in children finds poor sensitivity. *Journal of Clinical Epidemiology* 2006;59(4):354–7.

254. Nadal D, Leppert D, Frei K, *et al.* Tumour necrosis factor-alpha in infectious meningitis. *Archives of Disease in Childhood* 1989;64(9):1274–9.

257. Oray-Schrom P, Phoenix C, St MD, *et al.* Sepsis workup in febrile infants 0–90 days of age with respiratory syncytial virus infection. *Pediatric Emergency Care* 2003;19(5):314–19.

258. Rushton HG. Nocturnal enuresis: epidemiology, evaluation, and currently available treatment options. *Journal of Pediatrics* 1989;114(4 Pt 2):691–6.

259. Garra G, Cunningham SJ, Crain EF. Reappraisal of criteria used to predict serious bacterial illness in febrile infants less than 8 weeks of age. *Academic Emergency Medicine* 2005;12(10):921–5.

260. Jose TE, Mohiudheen H, Patel C, *et al.* Direct radionuclide cystography by supra-pubic puncture: Comparison with conventional voiding cystourethrography. *Nuclear Medicine Communications* 2004;25(4):383–5.

261. McCarthy PL, Jekel JF, Stashwick CA, *et al.* History and observation variables in assessing febrile children. *Pediatrics* 1980;65(6):1090–5.

262. Swingler GH, Hussey GD, Zwarenstein M. Randomised controlled trial of clinical outcome after chest radiograph in ambulatory acute lower-respiratory infection in children. *Lancet* 1998;351(9100):404–8.

263. Duke T, Blaschke AJ, Sialis S, Bonkowsky JL. Hypoxaemia in acute respiratory and non-respiratory illnesses in neonates and children in a developing country.*Archives of Disease in Childhood* 2002;86(2):108–12.

264. Gadomski AM, Aref GH, Hassanien F, el Ghandour S, El-Mougi M. Caretaker recognition of respiratory signs in children: correlation with physical examination findings, x-ray diagnosis and pulse oximetry. *International Journal of Epidemiology* 1993;22:1166–73.

265. Mower WR, Sachs C, Nicklin EL, Baraff LJ. Pulse oximetry as a fifth pediatric vital sign. *Pediatrics* 1997;99:681–6.

266. Kibirige MS, Edmond K, Kibirige JI, Rahman S. A seven year experience of medical emergencies in the assessment unit.*Archives of Disease in Childhood* 2003;88(2):125–9.

267. John M, Raj IS, Macaden R, Raghuveer TS, Yeswanth M, Meundi DM. Cerebrospinal fluid C-reactive protein measurement – a bedside test in the rapid diagnosis of bacterial meningitis. *Journal of Tropical Pediatrics* 1990;36:213–17.

268. Casado-Flores J, Blanco-Quiros A, Nieto M, Asensio J, Fernandez C. Prognostic utility of the semi-quantitative procalcitonin test, neutrophil count and C-reactive protein in meningococcal infection in children. *European Journal of Pediatrics* 2006;165:26–9.

269. Korppi M, Remes S. Serum procalcitonin in pneumococcal pneumonia in children. *European Respiratory Journal* 2001;17:623–7.

270. Gendrel D, Raymond J, Assicot M, Moulin F, Iniguez JL, Lebon P, Bohuon C. Measurement

of procalcitonin levels in children with bacterial or viral meningitis. *Clinical Infectious Diseases* 1997;24:1240–2.

271. Gendrel D, Bohuon C. Procalcitonin in pediatrics for differentiation of bacterial and viral infections. *Intensive Care Medicine* 2000;26:s178–81.

272. Korppi M, Remes S, Heiskanen-Kosma T. Serum procalcitonin concentrations in bacterial pneumonia in children: a negative result in primary healthcare settings. *Pediatric Pulmonology* 2003;35:56–61.

273. Ballot DE, Perovic O, Galpin J, Cooper PA. Serum procalcitonin as an early marker of neonatal sepsis.*South African Medical Journal Suid-AfrikaanseTydskrifVirGeneeskunde* 2004;94:851–4.

274. Wafula E, Tindybwa DB, Onyango F. The diagnostic value of various features for acute lower respiratory infection among under fives. *East African Medical Journal* 1989;66:678–84.

275. Stoll ML, Rubin L. Incidence of Occult Bacteremia among highly febrile young children in the era of the Pneumococcal conjugate vaccine: A study from a Children's Hospital emergency Department and Urgent Care Centre. *Archives of Pediatric and Adolescent Medicine* 2004;158:671–5.

276. Alpern E, Alessandrini E, Bell L, Shaw K. Occult bacteremia from a pediatric emergency department: current prevalence, time to detection, and outcome. *Pediatrics* 2000;106:505–11.

277. Bandyopadhyay S, Bergholte J, Blackwell C. Risk of serious bacterial infection in children with fever without a source in the post-*Haemophilusinfluenzae* era when antibiotics are reserved for culture-proven bacteremia. *Archives of Pediatric and Adolescent Medicine* 2002;156:512–17.

278. Mazur L, Kline M, Lorin M. Extreme leukocytosis in patients presenting to a pediatric emergency department. *Pediatric Emergency Care* 1991;7:215–18.

279. Hatherill M, Tibby SM, Sykes K, Turner C, Murdoch IA. Diagnostic markers of infection: comparison of procalcitonin with C reactive protein and leucocyte count. *Archives of Disease in Childhood* 1999;81:417–21.

280. Heulitt MJ, Ablow RC, Santos CS, O'Shea T. Febrile infants less than 3 months old: value of chest radiography. *Radiology* 1998;167:135–7.

281. Palmer SR, Corson J, Hall R, Payne S, Ludlow J, Deere B, *et al*. Meningococcal disease in Wales: Clinical features, outcome and public health management. *Journal of Infection* 1992;25:321–8.

282. Aksoylar S, Aksit S, Caglayan S, Yaprak I, Bakiler R, Cetin F. Evaluation of sponging and antipyretic medication to reduce body temperature in febrile children. *Actapaediatrica Japonica; Overseas edition* 1997;39:215–17.

283. Agbolosu NB, Cuevas LE, Milligan P, Broadhead RL, Brewster D, Graham SM. Efficacy of tepid sponging versus paracetamol in reducing temperature in febrile children. *Annals of Tropical Paediatrics* 1997;17:283–8.

284. Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteremia in young febrile children. *Annals of Emergency Medicine* 1998;31(6):679–87.

286. March Mde F, Sant'Anna CC. Signs and symptoms indicative of community-acquired pneumonia in infants under six months. *Brazilian Journal of Infectious Diseases* 2005;9(2):150–5.

287. Brogan PA, Raffles A. The management of fever and petechiae: making sense of rash decisions. *Archives of Disease in Childhood* 2000;83(6):506–7.

288. Advanced Life Support Group. *Advanced Paediatric Life Support: The Practical Approach*. 4th ed. BMJ Books/Blackwells; 2005.

289. McIntosh N, Helms P, Smyth R. *Forfar and Arneil's Textbook of Pediatrics*. 6th ed. Churchill Livingstone; 2003.

13 Abbreviations and glossary

13.1 Abbreviations

AHA	American Heart Association
ANC	absolute neutrophil count
AOR	adjusted odds ratio
APLS	Advanced Paediatric Life Support
ARR	absolute risk reduction
CCT	controlled clinical trial
CER	control event rate (see event rate)
CI	confidence interval
CNS	central nervous system
CRP	C-reactive protein
CRT	capillary refill time
CSF	cerebrospinal fluid
DGH	district general hospital (non-teaching hospital)
ED	emergency department
EER	experimental event rate (see event rate)
EL	evidence level (level of evidence)
ER	emergency room
ESR	erythrocyte sedimentation rate
FWS	fever without (apparent) source
GDG	guideline development group
GP	general practitioner
HES	Hospital Episode Statistics
hpf	high power field
HSE	herpes simplex encephalitis
HTA	Health Technology Appraisal
ICU	intensive care unit
IQR	inter quartile reange
ITU	intensive therapy unit
IV	intravenous
LR	likelihood ratio

LRTI	lower respiratory tract infection
MCD	meningococcal disease
MHRA	Medicines and Healthcare products Regulatory Agency
NCC-WCH	National Collaborating Centre for Women's and Children's Health
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NNH	number needed to harm (see number needed to treat)
NNT	number needed to treat
NPV	negative predictive value
NSAIDs	nonsteroidal anti-inflammatory drugs
OBI	occult bacterial infection
OPD	outpatient department
OR	odds ratio
PCT	procalcitonin
PCT	primary care trust
PGE2	prostaglandin E2
PPIP	Patient and Public Involvement Programme
PPV	positive predictive value
QALY	quality-adjusted life year
RCT	randomised controlled trial
ROC	receiver operating characteristic
RR	relative risk; respiratory rate
RSV	respiratory syncytial virus
SBI	serious bacterial illness/infection
SD	standard deviation
SIGN	Scottish Intercollegiate Guidelines Network
SpO2	transcutaneous oxygen saturation
SR	systematic review
temp.	temperature
TRIP	Turning Research into Practice
UTI	urinary tract infection
WBC	white blood cell count
WHO	World Health Organization
YIOS	Young Infant Observation Score
YOS	Yale Observation Score

13.2 Glossary

Absolute risk	Measures the probability of an event or outcome occurring (e.g. an adverse reaction to the drug being tested) in the group of people under study. Studies that compare two or more groups of patients may report results in terms of the absolute risk reduction.
Absolute risk reduction (ARR)	The ARR is the difference in the risk of an event occurring between two groups of patients in a study – for example, if 6% of patients die after receiving a new experimental drug and 10% of patients die after having the old drug treatment then the ARR is 10% − 6% = 4%. Thus by using the new drug instead of the old drug 4% of patients can be prevented from dying. Here the ARR measures the risk reduction associated with a new treatment. See also absolute risk.
Acute sector	Hospital-based health services which are provided on an inpatient, day case or outpatient basis.
Acute trust	A trust is an NHS organisation responsible for providing a group of healthcare services. An acute trust provides hospital services (but not mental health hospital services, which are provided by a mental health trust).
Allied health professionals	Healthcare professionals other than doctors and nurses directly involved in the provision of health care. Includes several groups such as physiotherapists, occupational therapists and dietitians. (Formerly known as professions allied to medicine or PAMs.)
Ambulatory care	All types of health services provided to patients who are not confined to a hospital bed as inpatients during the time services are rendered. Examples relevant to this guideline would include attendance to a walk-in centre or paediatric assessment unit, or the provision of care by paediatric community nurses.
Antipyretic interventions	Procedures or medications used with the intent of reducing body temperature in patients with fever. The term includes physical cooling methods and antipyretic medication. Paracetamol and ibuprofen are the drugs commonly used for this purpose in the UK.
Applicability	The extent to which the results of a study or review can be applied to the target population for a clinical guideline.
Appraisal of evidence	Formal assessment of the quality of research evidence and its relevance to the clinical question or guideline under consideration, according to predetermined criteria.
Bacteraemia	The presence of bacteria in the blood. In this condition the bacteria are not causing an infection in the bloodstream (cf. septicaemia).
Best available evidence	The strongest research evidence available to support a particular guideline recommendation.
Bias	Influences on a study that can lead to invalid conclusions about a treatment or intervention. Bias in research can make a treatment look better or worse than it really is. Bias can even make it look as if the treatment works when it actually does not. Bias can occur by chance or as a result of systematic errors in the design and execution of a study. Bias can occur at various stages in the research process, e.g. in the collection, analysis, interpretation, publication or review of research data. For examples see selection bias, performance bias, information bias, confounder or confounding factor, publication bias.
Blinding or masking	The practice of keeping the investigators or subjects of a study ignorant of the group to which a subject has been assigned. For example, a clinical trial in which the participating patients or their doctors are unaware of whether they (the patients) are taking the experimental drug or a placebo (dummy treatment). The purpose of 'blinding' or 'masking' is to protect against bias. See

	also double-blind study, single-blind study, triple-blind study.
Capillary refill time (CRT)	A test performed on physical examination in which the skin is pressed until blanched by the clinician's finger and the time taken for the skin to return to its previous colour is measured. Capillary refill time (CRT) can be measured peripherally (on the extremities) or centrally (on the chest wall). A prolonged CRT may be a sign of circulatory insufficiency (e.g. shock) or dehydration.
Case–control study	A study that starts with the identification of a group of individuals sharing the same characteristics (e.g. people with a particular disease) and a suitable comparison (control) group (e.g. people without the disease). All subjects are then assessed with respect to things that happened to them in the past, e.g. things that might be related to getting the disease under investigation. Such studies are also called retrospective as they look back in time from the outcome to the possible causes.
Case report (or case study)	Detailed report on one patient (or case), usually covering the course of that person's disease and their response to treatment.
Case series	Description of several cases of a given disease, usually covering the course of the disease and the response to treatment. There is no comparison (control) group of patients.
Causal relationship	Describes the relationship between two variables whenever it can be established that one causes the other. For example, there is a causal relationship between a treatment and a disease if it can be shown that the treatment changes the course or outcome of the disease. Usually randomised controlled trials are needed to ascertain causality. Proving cause and effect is much more difficult than just showing an association between two variables. For example, if it happened that everyone who had eaten a particular food became sick, and everyone who avoided that food remained well, then the food would clearly be associated with the sickness. However, even if leftovers were found to be contaminated, it could not be proved that the food caused the sickness – unless all other possible causes (e.g. environmental factors) had been ruled out.
Cerebrospinal fluid (CSF)	The watery fluid that surrounds the brain and spinal cord. Samples of CSF can be obtained by lumbar puncture.
Checklist	See study checklist.
Chemical dot thermometer	A thermometer consisting of cells embedded in a plastic strip in which the cells contain a combination of chemicals that change colour in response to changes in temperature. Also known as a chemical phase-change thermometer.
Chest indrawing	The indrawing of the lower chest wall. This is an important distinction from adults as ribs are made of cartilage in young children and form part of the chest wall.
Clinical audit	A systematic process for setting and monitoring standards of clinical care. Whereas 'guidelines' define what the best clinical practice should be, 'audit' investigates whether best practice is being carried out. Clinical audit can be described as a cycle or spiral. Within the cycle there are stages that follow a systematic process of establishing best practice, measuring care against specific criteria, taking action to improve care, and monitoring to sustain improvement. The spiral suggests that as the process continues, each cycle aspires to a higher level of quality.
Clinical effectiveness	The extent to which a specific treatment or intervention, when used under *usual or everyday conditions*, has a beneficial effect on the course or outcome of disease compared with no treatment or other routine care. (Clinical trials that assess effectiveness are sometimes called management trials.) Clinical 'effectiveness' is not the same as efficacy.

Clinical governance	A framework through which NHS organisations are accountable for both continually improving the quality of their services and safeguarding high standards of care by creating an environment in which excellence in clinical care will flourish.
Clinical impact	The effect that a guideline recommendation is likely to have on the treatment, or treatment outcomes, of the target population.
Clinical importance	The importance of a particular guideline recommendation to the clinical management of the target population.
Clinical question	This term is sometimes used in guideline development work to refer to the questions about treatment and care that are formulated in order to guide the search for research evidence. When a clinical question is formulated in a precise way, it is called a focused question.
Clinical trial	A research study conducted with patients which tests out a drug or other intervention to assess its effectiveness and safety. Each trial is designed to answer scientific questions and to find better ways to treat individuals with a specific disease. This general term encompasses controlled clinical trials and randomised controlled trials.
Clinician	A qualified healthcare professional providing patient care, e.g. doctor, nurse, physiotherapist.
Cluster	A group of patients, rather than an individual, used as the basic unit for investigation. See also cluster design, cluster randomisation.
Cluster design	Cluster designs are those where research subjects are not sampled or selected independently, but in a group. For example, a clinical trial where patients in a general practice are allocated to the same intervention; the general practice forming a cluster. See also cluster and cluster randomisation.
Cluster randomisation	A study in which groups of individuals (e.g. patients in a GP surgery or on a hospital ward) are randomly allocated to treatment groups. Take, for example, a smoking cessation study of two different interventions – leaflets and teaching sessions. Each GP surgery within the study would be randomly allocated to administer one of the two interventions. See also cluster and cluster design.
Cochrane Collaboration	An international organisation in which people find, appraise and review specific types of studies called randomised controlled trials. The Cochrane Database of Systematic Reviews contains regularly updated reviews on a variety of health issues and is available electronically as part of the Cochrane Library.
Cochrane Library	The Cochrane Library consists of a regularly updated collection of evidence-based medicine databases including the Cochrane Database of Systematic Reviews (reviews of randomised controlled trials prepared by the Cochrane Collaboration). The Cochrane Library is available on CD-ROM and the internet.
Cohort	A group of people sharing some common characteristic (e.g. patients with the same disease), followed up in a research study for a specified period of time.
Cohort study	An observational study that takes a group (cohort) of patients and follows their progress over time in order to measure outcomes such as disease or mortality rates and make comparisons according to the treatments or interventions that patients received. Thus, within the study group, subgroups of patients are identified (from information collected about patients) and these groups are compared with respect to outcome, e.g. comparing mortality between one group that received a specific treatment and one group that did not (or between two groups that received different levels of treatment). Cohorts can be assembled in the present and followed into the future (a 'concurrent' or 'prospective' cohort study) or identified from past records and followed forward from that time up to the present (a 'historical' or 'retrospective' cohort study). Because patients are not randomly allocated to subgroups, these subgroups

	may be quite different in their characteristics and some adjustment must be made when analysing the results to ensure that the comparison between groups is as fair as possible.
Combined modality	Use of different treatments in combination (e.g. surgery, chemotherapy and radiotherapy used together for cancer patients).
Commercial 'in confidence' material	Information (e.g. the findings of a research project) defined as 'confidential' as its public disclosure could have an impact on the commercial interests of a particular company. (Academic 'in confidence' material is information (usually work produced by a research or professional organisation) that is pending publication.)
Co-morbidity	Co-existence of a disease or diseases in the people being studied in addition to the health problem that is the subject of the study.
Confidence interval (CI)	A way of expressing certainty about the findings from a study or group of studies, using statistical techniques. A confidence interval describes a range of possible effects (of a treatment or intervention) that are consistent with the results of a study or group of studies. A wide confidence interval indicates a lack of certainty or precision about the true size of the clinical effect and is seen in studies with too few patients. Where confidence intervals are narrow they indicate more precise estimates of effects and a larger sample of patients studied. It is usual to interpret a '95%' confidence interval as the range of effects within which we are 95% confident that the true effect lies.
Confounder or confounding factor	Something that influences a study and can contribute to misleading findings if it is not understood or appropriately dealt with. For example, if a group of people exercising regularly and a group of people who do not exercise have an important age difference then any difference found in outcomes about heart disease could well be due to one group being older than the other rather than due to the exercising. Age is the confounding factor here and the effect of exercising on heart disease cannot be assessed without adjusting for age differences in some way.
Consensus development conference	A technique used for the purpose of reaching an agreement on a particular issue. It involves bringing together a group of about ten people who are presented with evidence by various interest groups or experts who are not part of the decision-making group. The group then retires to consider the questions in the light of the evidence presented and attempts to reach a consensus. See also Consensus methods.
Consensus methods	A variety of techniques that aim to reach an agreement on a particular issue. Formal consensus methods include Delphi and nominal group techniques, and consensus development conferences. In the development of clinical guidelines, consensus methods may be used where there is a lack of strong research evidence on a particular topic.
Considered judgement	The application of the collective knowledge of a guideline development group to a body of evidence, to assess its applicability to the target population and the strength of any recommendation that it would support.
Consistency	The extent to which the conclusions of a collection of studies used to support a guideline recommendation are in agreement with each other. See also homogeneity.
Control event rate (CER)	See event rate.
Control group	A group of patients recruited into a study that receives no treatment, a treatment of known effect, or a placebo (dummy treatment), in order to provide a comparison for a group receiving an experimental treatment, such as a new drug.

Controlled clinical trial (CCT)	A study testing a specific drug or other treatment involving two (or more) groups of patients with the same disease. One (the experimental group) receives the treatment that is being tested, and the other (the comparison or control group) receives an alternative treatment, a placebo (dummy treatment) or no treatment. The two groups are followed up to compare differences in outcomes to see how effective the experimental treatment was. A CCT where patients are randomly allocated to treatment and comparison groups is called a randomised controlled trial.
Cost–benefit analysis	A type of economic evaluation where both costs and benefits of healthcare treatment are measured in the same monetary units. If benefits exceed costs, the evaluation would recommend providing the treatment.
Cost-effectiveness	Value for money. A specific healthcare treatment is said to be 'cost-effective' if it gives a greater health gain than could be achieved by using the resources in other ways.
Cost-effectiveness analysis	A type of economic evaluation comparing the costs and the effects on health of different treatments. Health effects are measured in 'health-related units', for example, the cost of preventing one additional heart attack.
Cost–utility analysis	A special form of cost-effectiveness analysis where health effects are measured in quality-adjusted life years. A treatment is assessed in terms of its ability to both extend life and to improve the quality of life.
C-reactive protein (CRP)	A plasma protein that circulates in increased amounts during inflammation and after tissue damage. Measurement of CRP in blood samples is widely used as a marker of infection or inflammation.
Cross-sectional study	The observation of a defined set of people at a single point in time or time period – a snapshot. (This type of study contrasts with a longitudinal study, which follows a set of people over a period of time.)
Data set	A list of required information relating to a specific disease.
Decision analysis	Decision analysis is the study of how people make decisions or how they *should* make decisions. There are several methods that decision analysts use to help people to make better decisions, including decision trees.
Decision tree	A decision tree is a method for helping people to make better decisions in situations of uncertainty. It illustrates the decision as a succession of possible actions and outcomes. It consists of the probabilities, costs and health consequences associated with each option. The overall effectiveness or overall cost-effectiveness of various actions can then be compared.
Declaration of interest	A process by which members of a working group or committee 'declare' any personal or professional involvement with a company (or related to a technology) that might affect their objectivity e.g. if their position or department is funded by a pharmaceutical company.
Delphi method	A technique used for the purpose of reaching an agreement on a particular issue, without the participants meeting or interacting directly. It involves sending participants a series of postal questionnaires asking them to record their views. After the first questionnaire, participants are asked to give further views in the light of the group feedback. The judgements of the participants are statistically aggregated, sometimes after weighting for expertise. See also consensus methods.
Delphi statement	A statement of the advised course of action in relation to a particular clinical topic, based on the collective views of a body of experts by using the Delphi technique.
Diagnostic study	A study to assess the effectiveness of a test or measurement in terms of its ability to accurately detect or exclude a specific disease.

Dominance	A term used in health economics describing when an option for treatment is both less clinically effective and more costly than an alternative option. The less effective and more costly option is said to be 'dominated'.
Double-blind study	A study in which neither the subject (patient) nor the observer (investigator/clinician) is aware of which treatment or intervention the subject is receiving. The purpose of blinding is to protect against bias.
Economic evaluation	A comparison of alternative courses of action in terms of both their costs and consequences. In health economic evaluations the consequences should include health outcomes.
Effectiveness	See clinical effectiveness.
Efficacy	The extent to which a specific treatment or intervention, under *ideally controlled conditions* (e.g. in a laboratory), has a beneficial effect on the course or outcome of disease compared with no treatment or other routine care.
Elective	A term for clinical procedures that are regarded as advantageous to the patient but not urgent.
Empirical	Based directly on experience (observation or experiment) rather than on reasoning alone.
Encephalitis	Inflammation of the substance of the brain. It is usually caused by infection with viruses (e.g. herpes simplex virus).
Epidemiology	The study of diseases within a population, covering the causes and means of prevention.
Event rate	The proportion of patients in a group for whom a specified health event or outcome is observed. Thus, if out of 100 patients, the event is observed in 27, the event rate is 0.27 or 27%. Control event rate (CER) and experimental event rate (EER) are the terms used in control and experimental groups of patients, respectively.
Evidence based	The process of systematically finding, appraising and using research findings as the basis for clinical decisions.
Evidence-based clinical practice	Evidence-based clinical practice involves making decisions about the care of individual patients based on the best research evidence available rather than basing decisions on personal opinions or common practice (which may not always be evidence based). Evidence-based clinical practice therefore involves integrating individual clinical expertise and patient preferences with the best available evidence from research.
Evidence level (EL)	A code (e.g. 1++, 1+) linked to an individual study, indicating where it fits into the hierarchy of evidence and how well it has adhered to recognised research principles. Also called level of evidence.
Evidence table	A table summarising the results of a collection of studies which, taken together, represent the evidence supporting a particular recommendation or series of recommendations in a guideline.
Exclusion criteria	See selection criteria.
Experimental event rate (EER)	See event rate.
Experimental study	A research study designed to test whether a treatment or intervention has an effect on the course or outcome of a condition or disease – where the conditions of testing are to some extent under the control of the investigator. Controlled clinical trials and randomised controlled trials are examples of experimental studies.

Experimental treatment	A treatment or intervention (e.g. a new drug) being studied to see whether it has an effect on the course or outcome of a condition or disease.
External validity	The degree to which the results of a study hold true in non-study situations, for example in routine clinical practice. May also be referred to as the generalisability of study results to non-study patients or populations.
Extrapolation	The application of research evidence based on studies of a specific population to another population with similar characteristics.
Extremities	Medical term for the hands and feet.
Febrile convulsion	A fit caused by high body temperature in young children. Uncomplicated febrile convulsions are not associated with epilepsy in later life or other neurological complications.
Fever	For the purposes of this guideline, fever was defined as 'an elevation of body temperature above the normal daily variation'. See section 1.2 for more information on this definition.
Fever without (apparent) source (FWS)	The condition in which a patient has a fever but no obvious cause or focus of infection can be found on physical examination.
Focal neurological signs	Findings on physical examination that are caused by lesions in a particular area of the central nervous system. Examples include weakness of a limb or a cranial nerve palsy. These signs suggest that a given disease process is focal rather than diffuse.
Focal seizures	An epileptic seizure that originates from one part of the brain. Symptoms depend on which part of the brain is affected. Typically, one part of the body or one side of the body will develop convulsive movements. Focal (or partial) seizures can also include sensory disturbances, such as smelling or hearing things that are not there. In an uncomplicated focal seizure, consciousness is not lost. However, focal seizures can progress to involve the whole brain in a generalised seizure in which consciousness will be lost.
Focus group	A qualitative research technique. It is a method of group interview or discussion of 6–12 people focused around a particular issue or topic. The method explicitly includes and uses the group interaction to generate data.
Focused question	A study question that clearly identifies all aspects of the topic that are to be considered while seeking an answer. Questions are normally expected to identify the patients or population involved, the treatment or intervention to be investigated, what outcomes are to be considered, and any comparisons that are to be made. For example, do insulin pumps (intervention) improve blood sugar control (outcome) in adolescents with type 1 diabetes (population) compared with multiple insulin injections (comparison)? See also clinical question.
Fontanelle	A membrane-covered gap or soft spot between the skull bones on the vertex of an infant's skull. A bulging fontanelle can be a sign of bacterial meningitis.
Forest plot	A graphical display of results from individual studies on a common scale, allowing visual comparison of results and examination of the degree of heterogeneity between studies.
Funnel plot	Funnel plots are simple scatter plots on a graph. They show the treatment effects estimated from separate studies on the horizontal axis against a measure of sample size on the vertical axis. Publication bias may lead to asymmetry in funnel plots.
Generalisability	The extent to which the results of a study hold true for a population of patients beyond those who participated in the research. See also external validity.

Gold standard	A method, procedure or measurement that is widely accepted as being the best available.
Grey literature	Reports that are unpublished or have limited distribution, and are not included in bibliographic retrieval systems.
Grunting	A deep guttural breathing sound that can represent respiratory distress in infants and young children.
Guideline	A systematically developed tool that describes aspects of a patient's condition and the care to be given. A good guideline makes recommendations about treatment and care, based on the best research available, rather than opinion. It is used to assist clinician and patient decision making about appropriate health care for specific clinical conditions.
Guideline recommendation	Course of action advised by the guideline development group on the basis of their assessment of the supporting evidence.
Health economics	A branch of economics that studies decisions about the use and distribution of healthcare resources.
Health technology	Health technologies include medicines, medical devices such as artificial hip joints, diagnostic techniques, surgical procedures, health promotion activities (e.g. the role of diet versus medicines in disease management) and other therapeutic interventions.
Health technology appraisal (HTA)	A health technology appraisal, as undertaken by NICE, is the process of determining the clinical and cost-effectiveness of a health technology. NICE health technology appraisals are designed to provide patients, health professionals and managers with an authoritative source of advice on new and existing health technologies.
Herpes simplex infections	A group of acute infections caused by herpes simplex virus type 1 or type 2 that is characterised by the development of one or more small fluid-filled vesicles with a raised erythematous base on the skin or mucous membrane. Occasionally the viruses can cause more serious infections such as encephalitis in young children.
Heterogeneity	Or lack of homogeneity. The term is used in meta-analyses and systematic reviews when the results or estimates of effects of treatment from separate studies seem to be very different – in terms of the size of treatment effects or even to the extent that some indicate beneficial and others suggest adverse treatment effects. Such results may occur as a result of differences between studies in terms of the patient populations, outcome measures, definition of variables or duration of follow-up.
Hierarchy of evidence	An established hierarchy of study types, based on the degree of certainty that can be attributed to the conclusions that can be drawn from a well-conducted study. Well-conducted randomised controlled trials (RCTs) are at the top of this hierarchy. (Several large statistically significant RCTs which are in agreement represent stronger evidence than say one small RCT.) Well-conducted studies of patients' views and experiences would appear at a lower level in the hierarchy of evidence.
Homogeneity	This means that the results of studies included in a systematic review or meta-analysis are similar and there is no evidence of heterogeneity. Results are usually regarded as homogeneous when differences between studies could reasonably be expected to occur by chance. See also consistency.
Leucocyte count	The number of white blood cells per unit volume in venous blood. A differential leucocyte count measures the relative numbers of the different types of white cell.

Ill appearance	An ill-looking child is an overall impression the assessing healthcare professional can make when presented with a child. This impression is formed not only from objective measurements but also from subjective feelings about how the child looks/reacts. If a healthcare professional's subjective instinct is to describe the child as ill looking then the child is most likely at high risk of serious illness. Healthcare professionals should be confident to follow their impressions of a child's wellbeing.
Inclusion criteria	See selection criteria.
In-depth interview	A qualitative research technique. It is a face-to-face conversation between a researcher and a respondent with the purpose of exploring issues or topics in detail. It does not use pre-set questions, but is shaped by a defined set of topics or issues.
Infant	A child that is under the age of 12 months.
Information bias	Pertinent to all types of study and can be caused by inadequate questionnaires (e.g. difficult or biased questions), observer or interviewer errors (e.g. lack of blinding), response errors (e.g. lack of blinding if patients are aware of the treatment they receive) and measurement errors (e.g. a faulty machine).
Intention-to-treat (ITT) analysis	An analysis of a clinical trial where patients are analysed according to the group to which they were initially randomly allocated, regardless of whether or not they had dropped out, fully complied with the treatment, or crossed over and received the alternative treatment. Intention-to-treat analyses are favoured in assessments of clinical effectiveness as they mirror the non-compliance and treatment changes that are likely to occur when the treatment is used in practice.
Internal validity	Refers to the integrity of the study design.
Intervention	Healthcare action intended to benefit the patient, for example drug treatment, surgical procedure, psychological therapy, etc.
Interventional procedure	A procedure used for diagnosis or treatment that involves making a cut or hole in the patient's body, entry into a body cavity or using electromagnetic radiation (including X-rays or lasers). The National Institute for Health and Clinical Excellence (NICE) has the task of producing guidance about whether specific interventional procedures are safe enough and work well enough for routine use.
Kawasaki disease	A condition consisting of prolonged fever, a rash, changes to the extremities and mucous membranes, and enlargement of lymph glands in the neck. The exact cause is unknown but the condition is thought to be caused by a microbiological toxin. Kawasaki disease can cause aneurysms in the coronary arteries unless it is treated promptly.
Level of evidence	See evidence level.
Literature review	A process of collecting, reading and assessing the quality of published (and unpublished) articles on a given topic.
Longitudinal study	A study of the same group of people at more than one point in time. (This type of study contrasts with a cross-sectional study which observes a defined set of people at a single point in time.)
Lumbar puncture	A procedure in which cerebrospinal fluid is obtained by inserting a hollow needle into the space between vertebrae in the lumbar region of the spine. The procedure is used to diagnose bacterial meningitis and encephalitis.
Masking	See blinding.
Meningitis	Inflammation of the meninges, the membranes that lie between the surface of the brain and the inside of the skull. Meningitis is usually caused by infection

	with bacteria or viruses. Bacterial meningitis is a serious condition associated with appreciable mortality and significant neurological complications.
Meningococcal disease	Any of a number of infections caused by the bacterium *Neisseria meningitidis* (also known as the meningococcus). In young children meningococcal disease usually manifests as septicaemia, meningitis or a combination of the two. Meningococcal septicaemia is the leading infectious cause of death in childhood in the UK.
Meta-analysis	Results from a collection of independent studies (investigating the same treatment) are pooled, using statistical techniques to synthesise their findings into a single estimate of a treatment effect. Where studies are not compatible, for example because of differences in the study populations or in the outcomes measured, it may be inappropriate or even misleading to statistically pool results in this way. See also systematic review and heterogeneity.
Methodological quality	The extent to which a study has conformed to recognised good practice in the design and execution of its research methods.
Methodology	The overall approach of a research project, for example the study will be a randomised controlled trial, of 200 people, over 1 year.
Multicentre study	A study where subjects were selected from different locations or populations, for example a cooperative study between different hospitals or an international collaboration involving patients from more than one country.
Nasal flaring	An enlargement of the nostrils during breathing. Nasal flaring can indicate that increased work is required for breathing.
Negative predictive value (NPV)	The proportion of people with a negative test result who do not have the disease (where not having the disease is indicated by the gold standard test being negative).
Neonate	A newly born child aged up to and including 28 days.
NHS Direct	NHS Direct is a service that provides 24 hour confidential health advice and information. NHS Direct can help people who are feeling ill, are unsure what to do, would like to find out more about a condition or treatment, or need details of local health services. The service can be accessed by: • visiting www.nhsdirect.nhs.uk • going to NHS Direct Interactive on digital satellite TV (by pressing the interactive button on the remote control)calling 0845 4647.
Nominal group technique	A technique used for the purpose of reaching an agreement on a particular issue. It uses a variety of postal and direct contact techniques, with individual judgements being aggregated statistically to derive the group judgement. See also consensus methods.
Non-experimental study	A study based on subjects selected on the basis of their availability, with no attempt having been made to avoid problems of bias.
Non-paediatric practitioner	The term non-paediatric practitioner refers to a healthcare professional who has not had specific training and does not have recognised expertise in the management of children and their illnesses (cf. paediatric specialist). The term is mainly used to refer to healthcare professionals working in primary care but it may also apply to healthcare professionals in many general emergency departments.
Non-systematic review	See review.
Number needed to treat (NNT)	This measures the impact of a treatment or intervention. It states how many patients need to be treated with the treatment in question in order to prevent an event which would otherwise occur. For example, if the NNT = 4, then four patients would have to be treated to prevent one bad outcome. The closer the

	NNT is to 1, the better the treatment is. Analogous to the NNT is the number needed to harm (NNH), which is the number of patients that would need to receive a treatment to cause one additional adverse event. For example, if the NNH = 4, then four patients would have to be treated for one bad outcome to occur.
Objective measure	A measurement that follows a standardised procedure that is less open to subjective interpretation by potentially biased observers and study participants.
Observation	Observation is a research technique used to help understand complex situations. It involves watching, listening to and recording behaviours, actions, activities and interactions. The settings are usually natural, but they can be laboratory settings, as in psychological research.
Observational study	In research about diseases or treatments, this refers to a study in which nature is allowed to take its course. Changes or differences in one characteristic (e.g. whether or not people received a specific treatment or intervention) are studied in relation to changes or differences in other(s) (e.g. whether or not they died), without the intervention of the investigator. There is a greater risk of selection bias than in experimental studies.
Odds ratio (OR)	Odds are a way of representing probability, especially familiar for betting. In recent years odds ratios have become widely used in reports of clinical studies. They provide an estimate (usually with a confidence interval) for the effect of a treatment. Odds are used to convey the idea of 'risk' and an odds ratio of 1 between two treatment groups would imply that the risks of an adverse outcome were the same in each group. For rare events the odds ratio and the relative risk (which uses actual risks and not odds) will be very similar. See also relative risk, risk ratio.
Off-label prescribing	When a drug or device is prescribed outside its specific indication, to treat a condition or disease for which it is not specifically licensed.
Osteomyelitis	Infection of bone and bone marrow. Osteomyelitis is usually caused by bacteria. It can cause a chronic infection and disability if not treated appropriately.
Outcome	The end result of care and treatment and/or rehabilitation. In other words, the change in health, functional ability, symptoms or situation of a person, which can be used to measure the effectiveness of care/treatment/rehabilitation. Researchers should decide what outcomes to measure before a study begins; outcomes are then assessed at the end of the study.
P value	If a study is done to compare two treatments then the *P* value is the probability of obtaining the results of that study, or something more extreme, if there really was no difference between treatments. (The assumption that there really is no difference between treatments is called the 'null hypothesis'.) Suppose the *P* value was *P* = 0.03. What this means is that if there really was no difference between treatments then there would only be a 3% chance of getting the kind of results obtained. Since this chance seems quite low we should question the validity of the assumption that there really is no difference between treatments. We would conclude that there probably is a difference between treatments. By convention, where the value of *P* is below 0.05 (i.e. less than 5%) the result is seen as statistically significant. Where the value of *P* is 0.001 or less, the result is seen as highly significant. *P* values just tell us whether an effect can be regarded as statistically significant or not. In no way do they relate to how big the effect might be, for which we need the confidence interval.
Paediatric specialist	The term paediatric specialist refers to a healthcare professional who has had specific training or has recognised expertise in the management of children and their illnesses. Examples include paediatricians, or healthcare professionals working in children's emergency departments.

Peer review	Review of a study, service or recommendations by those with similar interests and expertise to the people who produced the study findings or recommendations. Peer reviewers can include professional and/or patient/carer representatives.
Performance bias	Systematic differences in care provided apart from the intervention being evaluated. For example, if study participants know they are in the control group they may be more likely to use other forms of care, people who know they are in the experimental group may experience placebo effects, and care providers may treat patients differently according to what group they are in. Masking (blinding) of both the recipients and providers of care is used to protect against performance bias.
Pilot study	A small scale 'test' of the research instrument. For example, testing out (piloting) a new questionnaire with people who are similar to the population of the study, in order to highlight any problems or areas of concern, which can then be addressed before the full-scale study begins.
Placebo	Placebos are fake or inactive treatments received by participants allocated to the control group in a clinical trial that are indistinguishable from the active treatments being given in the experimental group. They are used so that participants are ignorant of their treatment allocation in order to be able to quantify the effect of the experimental treatment over and above any placebo effect due to receiving care or attention.
Placebo effect	A beneficial (or adverse) effect produced by a placebo and not due to any property of the placebo itself.
Point estimate	A best single estimate (taken from research data) for the true value of a treatment effect or other measurement. For example, researchers in one clinical trial take their results as their best estimate of the real treatment effect – this is their estimate at their point in time. The precision or accuracy of the estimate is measured by a confidence interval. Another clinical trial of the same treatment will produce a different point estimate of treatment effect.
Positive predictive value (PPV)	The proportion of people with a positive test result who have the disease (where having the disease is indicated by the 'gold' standard test being positive).
Power	See statistical power.
Primary care	Health care delivered to patients outside hospitals. Primary care covers a range of services provided by GPs, nurses and other healthcare professionals, dentists, pharmacists and opticians.
Primary care trust (PCT)	A primary care trust is an NHS organisation responsible for improving the health of local people, developing services provided by local GPs and their teams (called primary care) and making sure that other appropriate health services are in place to meet local people's needs.
Probability	How likely an event is to occur, for example how likely a treatment or intervention will alleviate a symptom.
Procalcitonin	A precursor of the hormone calcitonin that is released into the bloodstream in response to infection or inflammation. Proclacitonin can be measured in blood samples and it is currently under development as a potential test for the detection of serious infections.
Prognostic factor	Patient or disease characteristics, for example age or co-morbidity, that influence the course of the disease under study. In a randomised trial to compare two treatments, chance imbalances in variables (prognostic factors) that influence patient outcome are possible, especially if the size of the study is fairly small. In terms of analysis these prognostic factors become confounding

	factors. See also prognostic marker.
Prognostic marker	A prognostic factor used to assign patients to categories for a specified purpose – for example for treatment, or as part of a clinical trial – according to the likely progression of the disease. For example, the purpose of randomisation in a clinical trial is to produce similar treatment groups with respect to important prognostic factors. This can often be achieved more efficiently if randomisation takes place within subgroups defined by the most important prognostic factors. Thus if age was very much related to patient outcome then separate randomisation schemes would be used for different age groups. This process is known as stratified random allocation.
Prospective study	A study in which people are entered into the research and then followed up over a period of time with future events recorded as they happen. This contrasts with studies that are retrospective.
Protocol	A plan or set of steps that defines appropriate action. A research protocol sets out, in advance of carrying out the study, what question is to be answered and how information will be collected and analysed. Guideline implementation protocols set out how guideline recommendations will be used in practice by the NHS, both at national and local levels.
Publication bias	Studies with statistically significant results are more likely to get published than those with non-significant results. Meta-analyses that are exclusively based on published literature may therefore produce biased results. This type of bias can be assessed by a funnel plot.
Qualitative research	Qualitative research is used to explore and understand people's beliefs, experiences, attitudes, behaviour and interactions. It generates non-numerical data, for example a patient's description of their pain rather than a measure of pain. In health care, qualitative techniques have been commonly used in research documenting the experience of chronic illness and in studies about the functioning of organisations. Qualitative research techniques such as focus groups and in-depth interviews have been used in one-off projects commissioned by guideline development groups to find out more about the views and experiences of patients and carers.
Quality-adjusted life years (QALYs)	A measure of health outcome that looks at both length of life and quality of life. QALYs are calculated by estimating the years of life remaining for a patient following a particular care pathway and weighting each year with a quality of life score (on a zero to one scale). One QALY is equal to 1 year of life in perfect health, or 2 years at 50% health, and so on.
Quantitative research	Research that generates numerical data or data that can be converted into numbers, for example clinical trials or the national Census that counts people and households.
Quasi-experimental study	A study designed to test whether a treatment or intervention has an effect on the course or outcome of disease. It differs from a controlled clinical trial and a randomised controlled trial in that: • the assignment of patients to treatment and comparison groups is not done randomly, or patients are not given equal probabilities of selection, or • the investigator does not have full control over the allocation and/or timing of the intervention, but nonetheless conducts the study as if it were an experiment, allocating subjects to treatment and comparison groups.
Random allocation or randomisation	A method that uses the play of chance to assign participants to comparison groups in a research study, for example, by using a random numbers table or a computer-generated random sequence. Random allocation implies that each individual (or each unit in the case of cluster randomisation) being entered into

	a study has the same chance of receiving each of the possible interventions.
Randomised controlled trial (RCT)	A study to test a specific drug or other treatment in which people are randomly assigned to two (or more) groups, with one (the experimental group) receiving the treatment that is being tested and the other (the comparison or control group) receiving an alternative treatment, a placebo (dummy treatment) or no treatment. The two groups are followed up to compare differences in outcomes to see how effective the experimental treatment was. (Through randomisation, the groups should be similar in all aspects apart from the treatment they receive during the study.)
Relative risk (RR)	A summary measure that represents the ratio of the risk of a given event or outcome (e.g. an adverse reaction to the drug being tested) in one group of subjects compared with another group. When the 'risk' of the event is the same in the two groups the relative risk is 1. In a study comparing two treatments, a relative risk of 2 would indicate that patients receiving one of the treatments had twice the risk of an undesirable outcome than those receiving the other treatment. Relative risk is sometimes used as a synonym for risk ratio.
Reliability	Reliability refers to a method of measurement that consistently gives the same results. For example, someone who has a high score on one occasion tends to have a high score if measured on another occasion very soon afterwards. With physical assessments it is possible for different clinicians to make independent assessments in quick succession – and if their assessments tend to agree then the method of assessment is said to be reliable.
Remote assessment	An assessment carried out when the patient is geographically remote from the assessor such that physical examination is not possible.
Retrospective study	A retrospective study deals with the present/past and does not involve studying future events. This contrasts with studies that are prospective.
Review	Summary of the main points and trends in the research literature on a specified topic. A review is considered non-systematic unless an extensive literature search has been carried out to ensure that all aspects of the topic are covered and an objective appraisal made of the quality of the studies.
Rigors	An episode of shaking or shivering which can occur when the child has high temperature. Unlike during a seizure episode, the child is conscious and alert. It occurs when the body increases its temperature to fight infection. Extreme shivering can be confused for febrile convulsion.
Risk ratio	Ratio of the risk of an undesirable event or outcome occurring in a group of patients receiving experimental treatment compared with a comparison (control) group. The term relative risk is sometimes used as a synonym for risk ratio.
Royal Colleges	In the UK medical/nursing world, the term Royal Colleges, as for example in 'The Royal College of …', refers to organisations that usually combine an educational standards and examination role with the promotion of professional standards.
Safety netting	The provision of support for patients in whom the clinician has some uncertainty as to whether the patient has a self-limiting illness and is concerned that their condition may deteriorate. Safety netting may take a number of forms, such as dialogue with the patient or carer about symptoms and signs to watch for, advice about when to seek further medical attention, review after a set period, and liaising with other healthcare services.
Sample	A part of the study's target population from which the subjects of the study will be recruited. If subjects are drawn in an unbiased way from a particular population, the results can be generalised from the sample to the population as a whole.

Sampling	Refers to the way participants are selected for inclusion in a study.
Sampling frame	A list or register of names that is used to recruit participants to a study.
Scottish Intercollegiate Guidelines Network (SIGN)	SIGN was established in 1993 to sponsor and support the development of evidence-based clinical guidelines for the NHS in Scotland.
Secondary care	Care provided in hospitals.
Selection bias	Selection bias has occurred if: • the characteristics of the sample differ from those of the wider population from which the sample has been drawn, or • there are systematic differences between comparison groups of patients in a study in terms of prognosis or responsiveness to treatment.
Selection criteria	Explicit standards used by guideline development groups to decide which studies should be included and excluded from consideration as potential sources of evidence.
Semi-structured interview	Structured interviews involve asking people pre-set questions. A semi- structured interview allows more flexibility than a structured interview. The interviewer asks a number of open-ended questions, following up areas of interest in response to the information given by the respondent.
Sensitivity	In diagnostic testing, sensitivity refers to the chance of having a positive test result given that you have the disease. 100% sensitivity means that all those with the disease will test positive, but this is not the same the other way around. A patient could have a positive test result but not have the disease – this is called a 'false positive'. The sensitivity of a test is also related to its negative predictive value (true negatives) – a test with a sensitivity of 100% means that all those who get a negative test result do not have the disease. To fully judge the accuracy of a test, its specificity must also be considered.
Septic	Affected by bacterial infection; hence septic shock, septic arthritis, etc.
Septicaemia	A serious medical condition in which there is rapid multiplication of bacteria in the bloodstream and in which bacterial toxins are present in the blood. Septicaemia is usually fatal unless treated promptly with parenteral antibiotics.
Shock	A pathological condition that can suddenly affect the haemodynamic equilibrium, usually manifested by failure to perfuse or oxygenate vital organs.
Sign	A finding on physical examination of a patient that provides the clinician with an objective indication of a particular diagnosis or disorder (cf. symptom).
Single-blind study	A study in which *either* the subject (patient/participant) *or* the observer (clinician/investigator) is not aware of which treatment or intervention the subject is receiving.
Social cues	A child's response to social interaction with a parent or health professional, such response to their name, smiling and/or giggling.
Specific indication	When a drug or a device has a specific remit to treat a specific condition and is not licensed for use in treating other conditions or diseases.
Specificity	In diagnostic testing, specificity refers to the chance of having a negative test result given that you do not have the disease. 100% specificity means that all those without the disease will test negative, but this is not the same the other way around. A patient could have a negative test result yet still have the disease – this is called a 'false negative'. The specificity of a test is also related to its positive predictive value (true positives) – a test with a specificity of 100% means that all those who get a positive test result definitely have the disease. To fully judge the accuracy of a test, its sensitivity must also be considered.

Standard deviation	A measure of the spread, scatter or variability of a set of measurements. Usually used with the mean (average) to describe numerical data.
Statistical power	The ability of a study to demonstrate an association or causal relationship between two variables, given that an association exists. For example, 80% power in a clinical trial means that the study has a 80% chance of ending up with a *P* value of less than 5% in a statistical test (i.e. a statistically significant treatment effect) if there really was an important difference (e.g. 10% versus 5% mortality) between treatments. If the statistical power of a study is low, the study results will be questionable (the study might have been too small to detect any differences). By convention, 80% is an acceptable level of power.
Structured interview	A research technique where the interviewer controls the interview by adhering strictly to a questionnaire or interview schedule with pre-set questions.
Study checklist	A list of questions addressing the key aspects of the research methodology that must be in place if a study is to be accepted as valid. A different checklist is required for each study type. These checklists are used to ensure a degree of consistency in the way that studies are evaluated.
Study population	People who have been identified as the subjects of a study.
Study quality	See methodological quality.
Study type	The kind of design used for a study. Randomised controlled trials, case–control studies and cohort studies are all examples of study types.
Subject	A person who takes part in an experiment or research study.
Survey	A study in which information is systematically collected from people (usually from a sample within a defined population).
Symptom	A patient's report of an abnormal feeling or sensation that provides the clinician with a subjective indication of a particular diagnosis or disorder (cf. sign).
Systematic	Methodical, according to plan; not random.
Systematic error	Refers to the various errors or biases inherent in a study. See also bias.
Systematic review	A review in which evidence from scientific studies has been identified, appraised and synthesised in a methodical way according to predetermined criteria. May or may not include a meta-analysis.
Systemic	Involving the whole body.
Tachypnoea	Abnormally rapid respiratory rate.
Target population	The people to whom guideline recommendations are intended to apply. Recommendations may be less valid if applied to a population with different characteristics from the participants in the research study, for example in terms of age, disease state or social background.
Tepid sponging	A traditional treatment for fever in which the patient is undressed and sponged with lukewarm water that is then allowed to evaporate.
Tertiary centre	A major medical centre providing complex treatments that receives referrals from both primary and secondary care. Sometimes called a tertiary referral centre. See also primary care and secondary care.
Triangulation	Use of three or more different research methods in combination; principally used as a check of validity. The more the different methods produce similar results, the more valid the findings.
Triple-blind study	A study in which the statistical analysis is carried out without knowing which treatment patients received, in addition to the patients and investigators/clinicians being unaware which treatment patients were getting.

Trust	A trust is an NHS organisation responsible for providing a group of healthcare services. An acute trust provides hospital services. A mental health trust provides most mental health services. A primary care trust buys hospital care on behalf of the local population, as well as being responsible for the provision of community health services.
Tympanic thermometer	A thermometer that is inserted into the external ear canal and measures the temperature of blood vessels in the tympanic membrane (eardrum) by detecting infrared radiation.
Validity	Assessment of how well a tool or instrument measures what it is intended to measure. See also external validity, internal validity.
Variable	A measurement that can vary within a study, for example the age of participants. Variability is present when differences can be seen between different people or within the same person over time, with respect to any characteristic or feature that can be assessed or measured.

Appendices A to L are presented as separate files