Children with fever without apparent source:
diagnosis and dilemmas

Kinderen met koorts zonder focus:
diagnose en dilemma's

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
Rector Magnificus

Prof.dr.ir. J.H. van Bemmel

en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op

woensdag 11 september 2002 om 15.45 uur

door
Sacha Elisabeth Bleeker
geboren te Apeldoorn

Promotiecommissie

Promotoren:	Prof.dr. R. de Groot
	Prof.dr. D.E. Grobbee
Overige leden:	Prof.dr. H.A. Büller
	Prof.dr.ir. J.D.F. Habbema
	Prof.dr. A.W. Hoes
Copromotoren:	Dr. H.A. Moll
	Dr. K.G.M. Moons

Cover: 'Spelende kinderen in Thailand'. Artistieke impressie van Kourosh Ajamlou naar een foto van Janet Wishnetsky uit de Novib kalender. Reprinted with kind permission of the artist.

The studies described in this thesis were supported by a grant from the Health Care Insurance Council of the Netherlands. The printing of this thesis was financially supported by the Julius Centre for General Practice and Patient Oriented Research, Utrecht and Abbott B.V.

ISBN 90-77017-95-X
Printed by Optima Grafische Communicatie, Rotterdam.

Contents

Chapter 1

Aims and outline

Aims of the studies

This thesis describes the results of diagnostic research in young children presenting with fever without apparent source at the emergency department. The study was conducted at the Sophia Children's University Hospital in Rotterdam and the Juliana Children's Hospital in The Hague, both large inner-city paediatric teaching hospitals in the Netherlands.

The specific aims of the studies are:

1. To describe trends in the management of children visiting the emergency department with fever without apparent source.
2. To develop a diagnostic prediction rule for referred patients presenting with fever without apparent source, including readily obtainable parameters from the patient's history, physical examination and laboratory tests in order to distinguish the patients with a serious bacterial infection from those without a serious bacterial infection.
3. To externally validate this developed diagnostic prediction rule for referred patients.
4. To obtain a diagnostic prediction rule for self-referred patients presenting with fever without apparent source, including the determination of the generalisability of the previously developed prediction rule for referred patients.
5. To deal with pitfalls with regard to diagnostic research on routine care data.
6. To compare results of internal and external validation of the developed diagnostic prediction rule for referred patients.
7. To develop a computer-based patient record for structured data entry for paediatric practice, in particular for recording data from the patient's history and physical examination.

Outline of this thesis

In chapter 2.1 the diagnostic and therapeutic dilemmas of young patients presenting with fever without apparent source are illustrated by two case histories. Changes during the last decade in the management of children presenting at the emergency department with fever without apparent source are presented and discussed in chapter 2.2.

Chapter 3 studies the diagnostic process in young children with fever without apparent source. A diagnostic prediction rule for the presence of a serious bacterial infection in patients referred to the emergency department by the general practitioner for the evaluation of fever without apparent source is described (3.1). External validation of this prediction rule is discussed in chapter 3.3. In chapter 3.2 the prediction of the presence of a serious bacterial infection in self-referred patients is discussed, resulting in a specific diagnostic prediction rule for this population.

In chapter 4 some aspects of the methods of performing (diagnostic) research that we encountered are deliberated and future perspectives are provided. Chapter 4.1 discusses the pros and cons of diagnostic research on routine care data. In chapter 4.2 the results of internal and external validation of a prediction rule for children with fever without apparent source are compared. Chapter 4.3 describes the development and evaluation of a computer-based patient record with structured data entry for a comprehensive paediatric patient history and physical examination.

In chapter 5 the results of the previous chapters are summarised (5.1 English; 5.2 Dutch).

Chapter 2

Introduction

2.1 The young child with fever without apparent source in the post-*Haemophilus influenzae* era

SE Bleeker, HA Moll

Nederlands Tijdschrift voor Geneeskunde 2002;146:3-5 [Het jonge kind met koorts zonder focus in het 'post-Haemophilus influenzae tijdperk']

Abstract

A 2½-year-old male and 2-month-old female patient presented with fever without apparent source. Alarming signs for the presence of a serious bacterial infection were reason to request additional laboratory tests. Pneumonia and viral meningitis were diagnosed, respectively. Due to changing prospects after eradication of invasive *Haemophilus influenzae* infections no suitable guidelines for children with fever without apparent source are available. By means of patient history and physical examination and subsequently additional laboratory tests, a first and respectively a second selection of patients (not) at risk for serious bacterial infections can be made. Furthermore, careful evaluation, clinical acumen, well-informed parents and observation are important factors.

Introduction

Children with fever (temperature ≥38°C) are a cause of concern for parents and a frequent reason for consulting a physician. Some 15% of all visits to the general practitioner and 10% to 35% of the visits to the emergency department are made in connection with children with fever (1-5). The majority of these patients are younger than 3 years of age and in 14% to 40%, no focus of fever was identified after a case history and a physical examination (3, 6). The underlying cause of the fever tends to be viral. However, depending on the population, a serious bacterial infection such as meningitis or sepsis is detected in 3% to 25% of the cases (see table) (5-7). Non-specific signs and symptoms in the young febrile child complicate the evaluation of the risk of having a serious bacterial infection even more. The following case histories illustrate that evaluating children with fever and distinguishing between those suffering and not suffering from a serious bacterial infection can be difficult.

Patient A, a boy aged 2½ years, presented with a temperature that had hovered at 41°C for three days, vomiting and a slight cough. The general practitioner was called and consulted. Additional history taking revealed symptoms of a common cold, headache, and a poor micturition and fluid intake. There was no diarrhoea. In the past, the patient had suffered febrile seizures and he had been vaccinated in accordance with the recommended immunisation schedule in the Netherlands. Clinical examination revealed no stiffness of the neck. Breathing was accelerated, no retractions were observed. The cervical lymph nodes were slightly swollen. Examination of ear-nose-throat, heart and lungs revealed no abnormalities. The general practitioner referred the patient to the paediatric emergency department for evaluation of 'fever without apparent source'. There, a moderately ill child was seen, with a good peripheral circulation and hydration status. The pulse rate was 120/minute, the body temperature 41°C, and the respiratory rate was 48/minute. There was no meningeal irritation present. No clear focus was found for the fever. The differential diagnosis included upper respiratory tract infection, pneumonia, urinary tract infection or gastritis. In addition, a blood work-up and a blood culture were performed; dipstick analysis of the urine was also carried out. No sputum could be taken.

The blood results (reference values) were: leukocytes: 20.7*10^9/L (4.0-10.0); C-reactive protein (CRP): 204 mg/L (<10); sodium: 136 mmol/L (135-145); potassium: 4.2 mmol/L (3.5-5.1); ureum: 4.4 mmol/L (2.6-5.6); creatinine: 35 μmol/L (40-70);

Table. Final diagnoses in 972 children, aged 1 month to 36 months, who presented with fever without apparent source between 1988-1998 at the emergency department of the Sophia Children's Hospital.

	Absolute number (%)
Serious bacterial infection present	178 (18,3)
Bacterial meningitis	20 (2,1)
Sepsis/ bacteraemia	14 (1,4)
Pneumonia	52 (5,3)
Bacterial gastroenteritis	6 (0,6)
Urinary tract infection	79 (8,1)
Cellulitis/ osteomyelitis/ abscess	7 (0,7)
Serious bacterial infection absent	794 (81,7)
Aseptic/ viral meningitis	12 (1,2)
Viral gastroenteritis	21 (2,2)
Acute otitis media	121 (12,4)
Viral syndrome*	556 (57,2)
Miscellaneous†	84 (8,6)

*Upper respiratory tract infection, respiratory syncytial virus bronchiolitis, (non-)specified viral infection.
†Immunisation reaction, malaria, skin infections, no diagnosis.

acid base status: pH: 7.46 (7.34-7.45); P_{CO2}: 4,2 kPa (4.7-5.9); bicarbonate: 22.0 mmol/L (21-27); base excess: –0.7 mmol/L (–2.3-+2.3). Dipstick analysis of the urine was negative. The chest X-ray showed an infiltration of the upper lobe of the right lung. As oral treatment was ruled out because of the vomiting, the boy was hospitalised and put on intravenous antibiotics.

Within 24 hours, the temperature had normalised and the patient was discharged in good condition with antibiotics per os. Blood culture and virus serologies were negative. In view of the temperature, mild tachypnoea, leukocytosis, elevated serum CRP and the abnormalities seen on the chest X-ray, the probability diagnosis was 'bacterial pneumonia'. At follow-up at the outpatient clinic the patient was found to have recovered completely.

Patient B, a 2-month old baby girl was seen by her general practitioner because of one day of inconsolable crying, groaning, temperature reaching 39°C, vomiting and a slight cough. The child did not feed as well as she normally did; the micturition was normal.

The patient's history revealed irritability and crying when her legs were manipulated (flexing the hips with extended legs). In view of her age, the patient had not yet received any vaccinations. The general practitioner saw a crying infant with a temperature of 39°C. Examination of the heart, lungs, ear, nose and throat area and abdomen revealed no abnormalities. The patient was referred by the general practitioner to the paediatric emergency department in order to discover the cause of the fever.

Physical examination showed an inconsolable infant with a pulse of 188/minute and temperature of 39.6°C. The peripheral circulation and hydration status were good. The child was hypertonic and meningeally irritable, but alert. There were no petechiae; the fontanelle was not bulging. The differential diagnosis was meningitis, sepsis, pneumonia or urinary tract infection. Laboratory blood results (reference values) were as follows: CRP: 3 mg/L (<10); thrombocytes: $602*10^9$/L (200-473); leukocytes: $14.1*10^9$/L (5-21); glucose: 7.4 mmol/L (2.6-6.0). The urinary dipstick analysis was negative. Cerebral spinal fluid test results were: 316 (of which 224 polymorphs) cells per $3*10^{-6}$ L (0-10); glucose: 2.6 mmol/L (2.1-4.2); negative gram stain. Bacteriological and viral cerebrospinal fluid cultures were obtained, in addition to a blood culture. Radiological examination (X-ray) of the chest revealed no abnormalities. The working diagnosis at this time was 'meningitis (bacterial or viral)' and the infant was admitted and placed on intravenous antibiotics. The temperature dropped to normal within 24 hours. However, after 36 hours another fever spike occurred up to 38.6°C. Within 48 hours of admission the child's condition had improved considerably and the fever had dropped to below 38°C. Bacteriological blood and cerebral spinal fluid cultures remained sterile, after which the antibiotic therapy was stopped after 72 hours. The viral cerebral spinal fluid culture remained sterile, although a Coxsackie virus type B5 was isolated from the throat and stool. The final diagnosis, based on the clinical symptoms and the elevated cerebral spinal fluid cell count, was 'viral meningitis'. The child was discharged in good condition.

Discussion

Both patients were referred by the general practitioner on the basis of alarm signals (8, 9). In patient A these included the duration and height of the fever, a poor micturition and fluid intake, and tachypnoea; in patient B these were inconsolable crying and groaning in combination with the patient's young age. These symptoms, in

combination with the additional alarm signals found by the paediatrician (patient A: moderately ill; patient B: meningeal irritation and hypertonia), were grounds for additional testing. In patient A, the blood results indicated a bacterial infection. The chest X-ray revealed the presence of pneumonia. In patient B the diagnosis was based on cerebral spinal fluid tests. Both patients were hospitalised. Patient A, however, would not have been admitted to hospital had oral administration of antibiotics been possible. Patient B was hospitalised in order to receive intravenous antibiotic therapy, which in retrospect appeared to have been unnecessary.

Fever without apparent source in young children frequently poses a diagnostic and therapeutic dilemma. To date evidence-based guidelines for specific patient groups are uncommon, and are an issue of debate. In 1993, a guideline for children from 0 to 36 months old with fever was published in the American literature; a policy strategy for general practitioner assessment and referral to the paediatrician that was derived from this American guideline was published in this journal in 1999 (8, 9).

The question is, however, whether this policy in children with fever without apparent source still suffices in today's practice, where invasive *Haemophilus influenzae* type b (Hib) infections have become rare (a decrease in the Netherlands from 700 Hib-infections per annum to 8 in 1997), or that a more expectative policy is justified (10, 11). Hib-vaccinations are given to children born later than April 1 1993; the vaccine is administered in the 2nd, 3rd, 4th and 11th month of life (before January 1 1999 the schedule was in the 3rd, 4th, 5th and 11th month). As of January 1997, a vaccination level of 95.5% has been achieved and the expected vaccine effectiveness is 99.4% (12, 13).

In the following, the policy strategy formulated in 1999 is briefly summarised per age group and discussed in the light of current insights.

0-28 days. According to the described policy strategy, for this age group applies (9): paediatric care, hospitalisation for diagnostics; treatment dependent on the risk profile. With respect to the current insights and in the light of the non-specific presentation at this age, the immature immune system and absence of protection by vaccination, this strategy still holds.

1 month -3 months. The described policy strategy formulated (9): in the absence of alarm symptoms seen by the general practitioner (ill clinical appearance, body temperature <36°C or ≥39°C, bulging fontanelle, <4 wet nappies/ 24 hours, diarrhoea): no referral; obtain urinary culture. In the case of ≥1 alarm symptom(s) at

the general practitioner: referral; depending on the laboratory results at paediatric care (leukocytes, band count and thrombocytes; urine sediment) admission for diagnostics and parenteral antibiotic treatment. With the current insights and in view of the negative or incomplete level of immunisation against Hib-infections, this strategy still holds.

3-36 months. The policy strategy described (9): in the absence of alarm symptoms (ill clinical appearance, temperature ≥39.5°C, bulging fontanelle, tachypnoea, diarrhoea) seen by the general practitioner: no referral; obtain urinary culture. In the case of ≥1 alarm symptom(s) at the general practitioner: referral; dependent on results of laboratory tests carried out at paediatric care (leukocyte and neutrophile count, CRP, urine sediment) hospital-bed admission for diagnostics and parenteral antibiotic treatment. Under current insights this strategy must be adjusted.

As invasive Hib-infections accompanied by severe morbidity, mortality and late sequelae virtually no longer occur, this policy would appear too stringent, in particular regarding the referral from the general practitioner to the paediatrician.

This shift in the spectrum of organisms causing serious bacterial infections has led to *Streptococcus pneumoniae* now being cultured in >90% of patients with bacterial infections. Some 6% of all cases of pneumococcal bacteraemia develop into a serious bacterial infection, and in children aged 3 to 36 months with fever ≥39°C, the risk of a pneumococcal meningitis is 0.1% (14, 15). Blood cultures only very rarely reveal *Neisseria meningitidis* (9, 16).

One option would be a more expectative policy strategy for these older children, including less hasty referrals to paediatric care, more reticence in deciding on treatment with antibiotics and a more important role for observation as a major diagnostic tool (17).

Detailed guidelines for a new policy strategy for children between 3 to 36 months of age with fever without apparent source require well-substantiated diagnostic studies, wholly centred on patients presenting with fever without apparent source.

Until studies of this kind have been carried out, careful evaluation (at first by the general practitioner), by which is meant the conduct of a directed history taking and physical examination, will continue to remain of first importance in sifting out initially the children with and those without a serious bacterial infection.

Readily laboratory testing will help with the further task of winnowing afflicted children out. Time, as a diagnostic instrument, is another key factor.

An expectative policy at the general practitioner must be conditioned on: (a) a hitherto healthy child, (b) parents capable of receiving instruction and with whom clear agreements can be made (careful observation of the child, take action if a change in the child's condition should occur) and (c) opportunity for re-evaluation in time if necessary. At this, one should bear in mind that it is impossible to eliminate all risks and that guidelines are drafted to support clinicians when making decisions, not to replace them. A guideline should not be applied as a rigid instrument in practice, but should be tailored to fit the needs of each individual.

Nowadays, in this "post-*Haemophilus influenzae* era", a more expectative policy is advised towards children aged 3 to 36 months, with less hasty referrals to paediatric care. Careful evaluation, clinical acumen, well-informed parents and time as a diagnostic instrument are all important factors in this respect.

Acknowledgements

Professor Dr. R. de Groot critically reviewed the manuscript.

References

1. Hart HE, Bruijnzeels MA, Wouden JC van der, Suijlekom-Smit LWA van, Does E van der, Velden J van der. [The child with fever in general practice]. Huisarts Wet 1992;35:246-248.
2. McCarthy PL. Fever. Pediatr in Rev 1998;19:401-8.
3. Baraff LJ, Lee SI. Fever without source: management of children 3 to 36 months of age. Pediatr Infect Dis J 1992;11:146-51.
4. Steensel-Moll HA van, Jongkind CJ, Aarsen RSR, et al. [A problem-oriented patient classification system for general paediatrics II]. Tijdschr Kindergeneesk 1996;64:99-104.
5. Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteraemia in young febrile children. Ann Emerg Med 1998;31:679-87.
6. Baker MD, Bell LM, Avner JR. Outpatient management without antibiotics of fever in selected infants. N Engl J Med 1993;329:1437-41.
7. Baraff LJ, Oslund SA, Schriger DL, Stephen ML. Probability of bacterial infections in febrile infants less than three months of age: a meta-analysis. Pediatr Infect Dis J 1992;11:257-64.
8. Baraff LJ, Bass JW, Fleisher GR, et al. Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Paediatrics 1993;92:1-12.
9. Oostenbrink R, de Groot R, Moll HA. [Fever of unknown origin in a young child:

diagnosis and treatment]. Ned Tijdschr Geneeskd 1999;143:185-90.
10. Bol P. [Epidemiology of *Haemophilus influenzae* type b infections in the Netherlands and elsewhere]. Ned Tijdschr Geneeskd 1991;135:7-9.
11. Conyn-van Spaendonck MA, Veldhuijzen IK, Suijkerbuijk AW, Hirasing RA. [Significant decline of the number of invasive *Haemophilus influenzae* infections in the first 4 years after introduction of vaccination against H. influenzae type B in children]. Ned Tijdschr Geneeskd 2000;144:1069-73.
12. Inspectie voor de Gezondheidszorg. Vaccinatietoestand Nederland per 1 januari 1997. Den Haag: Inspectie voor de Gezondheidszorg; 1998.
13. Barbour ML, Mayon-White RT, Coles C, Crook DW, Moxon ER. The impact of conjugate vaccine on carriage of *Haemophilus influenzae* type b. J Infect Dis 1995;171:93-8.
14. Alario AJ, Nelson EW, Shapiro ED. Blood cultures in the management of febrile outpatients later found to have bacteraemia. J Pediatr 1989;115:195-9.
15. Baraff LJ. Management of fever without source in infants and children. Ann Emerg Med 2000;36:602-14.
16. Furth AM van, Zaaijer HL. [Meningococcal disease in the Netherlands: media hype, but not an epidemic]. Ned Tijdschr Geneeskd 2001;145:1716-8.
17. Long SS. Antibiotic therapy in febrile children: "best-laid schemes". J Pediatr 1994;124: 585-8.

2.2 Trends in antibiotic prescription and hospitalisation in children with fever without apparent source at the emergency department

SE Bleeker, G Derksen-Lubsen, WK Redekop, KGM Moons, JB Oostenbrink, DE Grobbee, R de Groot, HA Moll

Submitted

Abstract

During the last decade, the management of children with fever without apparent source was characterised by the introduction of the *Haemophilus influenzae* type b vaccine, the realisation of the growing emergence of antibiotic resistance in major paediatric pathogens, and the establishment of guidelines for diagnosis and therapy.

The aim of this study was to compare the management of children presenting at the emergency department with fever without apparent source before and after the aforementioned developments. To this purpose, retrospective data from patients referred by their general practitioner with fever without apparent source from two periods (1988-1992, n=379 and 1996-1998, n=231; Sophia Children's University Hospital in Rotterdam and Juliana Children's Hospital in The Hague, the Netherlands) were compared. Collected data included general characteristics, diagnosis, treatment, hospitalisation and follow-up. In data analysis, χ^2-tests (parametric data), Wilcoxon tests (non-parametric data) and t-tests (comparing means) were used. P-values of less than 0.05 were considered significant.

The frequency of serious bacterial infections did not change over time (21% versus 25%). Measurements of serum C-reactive protein increased from 33% to 87% ($p<0.001$). The percentage of children with cultures of cerebrospinal fluid, blood or faeces and the proportion of children treated with antibiotics reduced significantly. Admission rates remained stable (26% of all patients with fever without apparent source). The mean length of hospitalisation lessened considerably. The mean

associated costs of the management per patient reduced, mounting to 46% in the children without a serious bacterial infection.

These findings indicate that the management in children with fever without apparent source has become less invasive and more selective, reducing the costs without showing a worsening effect on the clinical outcome.

Introduction

Fever evokes anxiety in parents and is encountered in 10% to 35% of all children evaluated at an emergency department (ED) (1-4). Careful evaluation is necessary because young children, who present with non-specific signs and symptoms, occasionally develop a serious bacterial infection. In 14% to 40% of these febrile children no focus can be identified after history taking and physical examination (2, 5, 6). The physician should balance the risks for a serious bacterial infection and several subsequent decisions must be made. The first is whether additional diagnostic tests should be performed, and subsequently, whether antibiotic treatment is indicated. Confirmation of serious bacterial infections, such as meningitis or sepsis, generally involves positive cultures. These culture results will not be available within 12 to 24 hours. Therefore, a preliminarily decision on treatment often has to be made, using an empirical choice of antibiotics. Also, the clinician must determine whether the child should be hospitalised or treated as an outpatient. Several guidelines have been issued to help clinicians navigate through these complex issues (4, 5, 7-18).

In the last decade important changes occurred which possibly influenced the diagnostic and therapeutic management of children with fever without apparent source. First, the *Haemophilus influenzae* type b (Hib) conjugate vaccine was introduced (1993), resulting in a near elimination of invasive Hib-infections. Second, the awareness of ominous antibiotic resistance is growing, in particular penicillin resistance in *Streptococcus pneumoniae* (19-21). Finally, the aforementioned guidelines for diagnosis and therapy were introduced.

The aim of this study was to compare the diagnostic and therapeutic management in children presenting at the ED with fever without apparent source before (1988-1992) and after (1996-1998) the above-mentioned developments. We hypothesised that the proportion children with antibiotic treatments and the number of admissions had decreased during time. Additionally, we estimated the costs of medical practice in both time periods.

Patients and methods

Patients

This study was conducted as part of a large ongoing study on paediatric diagnostic management (18, 22-24) and was approved by the Institutional Review Boards of both participating hospitals. Patients between 1 month and 36 months of age who attended the ED of the Sophia Children's University Hospital Rotterdam (1988-1992 and 1996-1998) and the Juliana Children's Hospital in The Hague (1998) were enrolled. Both are large inner-city teaching hospitals in the Netherlands. Patients with acute fever without apparent source, including suspected sepsis, were enrolled. Fever without apparent source was defined as a body temperature of at least thirty-eight degrees Celsius for which no apparent source was found after evaluation by the general practitioner or history taking by the paediatrician. Patient data were retrieved by means of a problem-oriented patient classification system, in which the main reason for encounter is classified (3, 25). Patients who were not referred by a general practitioner or who suffered from immune deficiencies were excluded.

Measurements and data retrieval

Data were collected by reviewing the standardised medical records. These included information on patient history, physical examination, diagnostic tests (retrieved from the computer-documented hospital information system), treatment, hospitalisation and follow-up.

For each patient the final diagnosis was determined either by a reference standard (cultures of blood, cerebrospinal fluid, urine, stools) or based on a consensus diagnosis. Outcome diagnosis was the presence or absence of a serious bacterial infection. Serious bacterial infection was defined as the presence of bacterial meningitis, sepsis or bacteraemia, pneumonia, urinary tract infection, bacterial gastroenteritis, osteomyelitis or ethmoiditis. A follow-up period of two weeks was the standard for ruling out the possibility of a missed diagnosis of serious bacterial infection. Detailed definitions of specific diagnoses have been described previously (18, 23).

Outcome on treatment was dichotomised into antibiotic treatment (oral and parenteral antibiotic treatment) and no antibiotic treatment. Outcome on follow-up was dichotomised into hospitalisation and no hospitalisation (outpatient treatment and discharge without follow-up).

Data analysis

The patient group from the time period 1988-1992 was compared to the group with patients who presented at the ED between 1996 and 1998, with respect to general characteristics, diagnosis, treatment, hospitalisation and follow-up. In order to analyse frequency distributions, χ^2-tests were used for parametric data and Wilcoxon tests for non-parametric data. T-tests were used for comparison of means. P-values were considered significant at 0.05 or less.

The costs of diagnosis and treatment were estimated by multiplying the resource use of each patient (using the empirical data) with the estimated average unit prices of diagnostic tests and therapeutic intervention (24). Resource use included the ED visit, diagnostic tests at presentation (laboratory and radiographic tests) or during clinical course, and in-hospital or ambulatory treatment (nursing days, attendance by paediatrician, and prescription of medication use) (26). Unit cost calculations in Euro (1 Euro = $0.946, January 2001) were based on the financial accounts of the Sophia Children's University Hospital Rotterdam and the Juliana Children's Hospital of 1996.

Results

Six hundred and ten children with fever without apparent source were included in this study. Of those, 379 attended the ED between 1988 and 1992 and 231 between 1996 and 1998. The demographic, clinical characteristics and final diagnoses for the two patient groups are presented in table 1. Demographic and clinical characteristics were similar. The mean age of the patients was 1.2 years and the mean duration of fever at presentation was 2.8 days. No material differences were found in the final diagnoses between both patient groups, with the exception that in 1996-1998 less patients had a viral syndrome, while more had a miscellaneous diagnosis ($p<0.05$). A serious bacterial infection was diagnosed in 21% of the children in 1988-1992 and in 25% in the recent patient group.

The frequencies of diagnostic tests that were performed within the first 24 hours of the evaluation of the febrile patients at the ED are shown in table 2. The serum C-reactive protein (CRP) level was measured more frequently in the second period. In contrast, cultures of cerebrospinal fluid, blood or faeces were performed significantly less often.

Table 1. Demographic and clinical characteristics of patients presenting with fever without apparent source in 1988-1992 and 1996-1998*.

	1988-1992 (n†=379)	1996-1998 (n=231)
Male gender	216 (57)	122 (53)
Age (years)‡	1.2 (0.7)	1.1 (0.8)
Weeks of gestation‡	39.4 (1.8)	39.3 (2.1)
Duration of fever (days)‡	2.8 (2.4)	2.8 (2.4)
Body temperature at physical examination (°C)‡	39.4 (0.9)	39.4 (1.0)
Final diagnoses		
Serious bacterial infection present	*78 (20.6)*	*58 (25.1)*
Bacterial meningitis	9 (2.4)	3 (1.3)
Sepsis/ bacteraemia	7 (1.8)	3 (1.3)
Pneumonia	26 (6.9)	28 (12.1)
Bacterial gastroenteritis	3 (0.8)	2 (0.9)
Urinary tract infection	31 (8.2)	22 (9.5)
Osteomyelitis/ ethmoiditis	2 (0.5)	0 (0)
Serious bacterial infection absent	*301 (79.4)*	*173 (74.9)*
Aseptic/ viral meningitis	9 (2.4)	3 (1.3)
Viral gastroenteritis	6 (1.6)	9 (3.9)
Acute otitis media	15 (4.0)	7 (3.0)
Viral syndrome§	221 (58.3)	104 (45.0)
Miscellaneous¶	50 (13.2)	50 (21.6)

*Values represent absolute numbers (percentages) unless stated otherwise.

†Number of patients.

‡Mean (standard deviation).

§Upper respiratory tract infection, respiratory syncytial virus bronchiolitis, non-specified viral infection.

¶Drugfever, immunisation reaction, Kawasaki disease, exanthema subitum, vasculitis, malaria, no diagnosis.

The initial management at the ED of the children presenting with fever without apparent source is presented in table 3. The ratios of admitted patients, patients treated on ambulatory basis and of patients discharged without follow-up remained stable over time. The frequency of ambulatory care with empirical prescription of antibiotics decreased, whereas that of the outpatient management without any antibiotic treatment increased markedly.

Table 2. Diagnostic tests in patients presenting with fever without apparent source over time*.

Diagnostic test	1988-1992 (n†=379)	1996-1998 (n=231)	p-value
Serum white blood cell count (*10^9/L)	352 (93)	205 (89)	0.10
Serum C-reactive protein (mg/L)	126 (33)	202 (87)	0.01
Cultures			
Cerebrospinal fluid	83 (22)	32 (14)	0.01
Blood	131 (35)	60 (26)	0.03
Urine	123 (33)	60 (26)	0.10
Faeces	85 (22)	29 (13)	0.01
Chest X-ray	137 (36)	73 (32)	0.26

*Values represent absolute numbers (percentages).
†Number of patients.

Table 3. Initial management in children with fever without apparent source at the emergency department over time*.

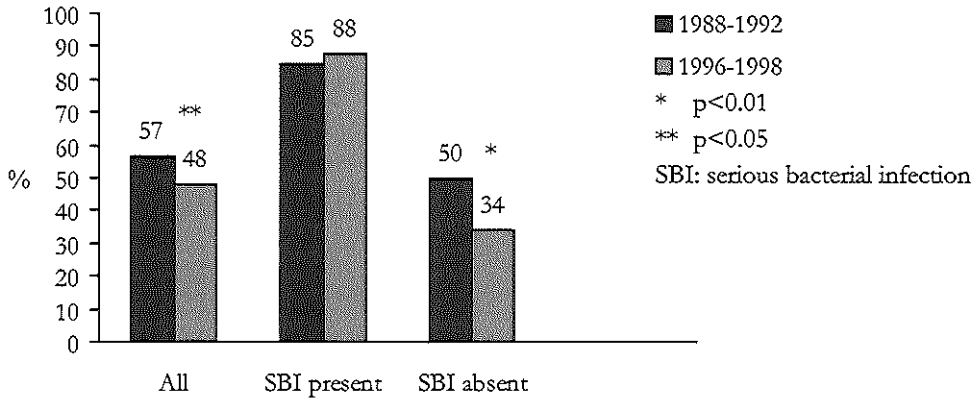
Management	1988-1992 (n†=379)	1996-1998 (n=231)	p-value
Hospitalisation with antibiotic treatment	81 (21)	50 (22)	0.9
Hospitalisation without antibiotic treatment	16 (4)	12 (5)	0.6
Ambulatory care with antibiotic treatment	95 (25)	38 (17)	0.01
Ambulatory care without antibiotic treatment	78 (21)	73 (32)	0.01
Discharge with antibiotic treatment	38 (10)	18 (8)	0.4
Discharge without antibiotic treatment	71 (19)	40 (17)	0.7

*Values represent absolute patient numbers (percentages).
†Number of patients.

The frequencies of overall antibiotic treatments (i.e. in-hospital or as outpatient) in both patient groups are shown in figure 1. In the period 1996-1998 the prescription of antibiotics in patients with fever without apparent source decreased significantly compared to the earlier time period.

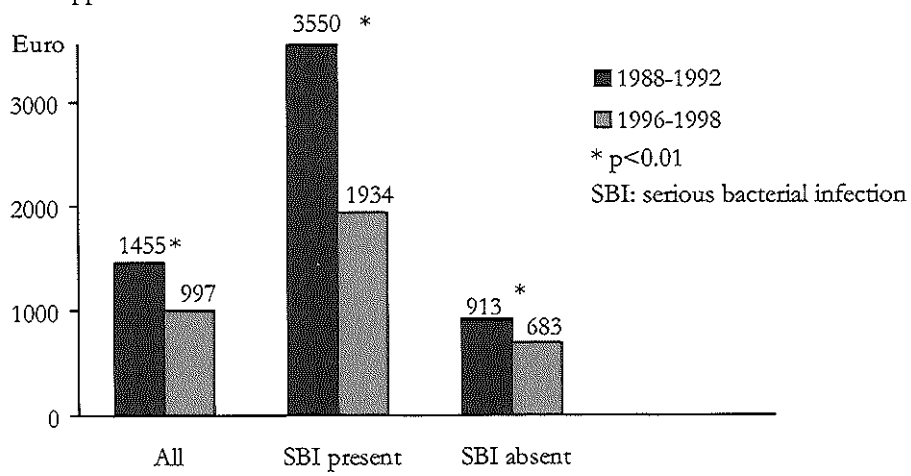
Approximately a quarter of all patients in both time periods were admitted to the hospital at the initial evaluation. No differences were found between the groups in the frequencies of patients hospitalised only after a re-evaluation (3% versus 2%). The mean duration of hospitalisation declined from 9 to 6 days over time ($p<0.01$).

Figure 1. Frequencies of patients with fever without apparent source treated with antibiotics in 1988-1992 and in 1996-1998.



In the patients in whom eventually a serious bacterial infection was diagnosed, antibiotic treatments remained unchanged (figure 1), the hospitalisation rate remained approximately 50%, and the mean duration of the hospital stay decreased with more than 4 days (12 versus 7.6 days). In those patients eventually without serious bacterial infections, prescription of antibiotics reduced by 16% over time (figure 1), the hospitalisation rate remained stable (20%), and the hospital stay shortened from 6.3 to 4.9 days.

Figure 2: Mean costs (Euro) per patient presenting at the emergency department with fever without apparent source over time.



Costs of the management per patient presenting at the ED with fever without apparent source are depicted in figure 2. The costs in both the patients without and with serious bacterial infections decreased importantly during time, mounting to a 46% reduction in the latter group.

Discussion

This study describes the changes over time in diagnostic and therapeutic management of children presenting with fever without apparent source and being at risk for serious bacterial infections. Data from two historical periods (1988-1992 versus 1996-1998) were compared, overall and per outcome diagnosis (presence or absence of a serious bacterial infection). The results showed a trend towards a less invasive management in young children with fever without apparent source.

Such an approach agrees with issued recommendations. During the late eighties and the nineties, the management in febrile children shifted from a broad approach including obtaining cultures, with easy admission and treatment, to a more selective and less invasive (outpatient) management after judicious evaluation (8, 16, 27-30). In addition, the introduction of the Hib-vaccine, the availability of professional recommendations on the management of febrile children, and the emergence of antibiotic resistance had a clear impact (4, 6, 10, 17, 19-21, 31-34).

The stable frequency of serious bacterial infections over time might at first seem contradictory to the decline in the occurrence of invasive Hib-infections due to the Hib-vaccine. However, the proportion of Hib-invasive diseases (meningitis) among serious bacterial infections was relatively small in our first patient population (3/379) and an elimination of this agent was unnoticed.

Besides the reduced fear for the invasive Hib-infections, the decrease in obtaining cultures might be influenced by the enabling of rapid quantification of levels of CRP, as expressed in an increase in CRP measurement (35).

Although the length of hospitalisation decreased significantly, the percentage of hospitalised patients did not. This is in accordance with the observation that the frequency of serious bacterial infections changed neither. In both time periods, one in the five not-afflicted children was hospitalised. No information was available on the reasons for admission in these children without a serious bacterial infection, but could have been suspicion of serious bacterial infection or need for in-hospital treatment other than because of a serious bacterial infection. The reduction in the costs in

children with fever without apparent source is mainly a reflection of a shortened duration of hospitalisation.

In conclusion, when patients presenting with fever without apparent source between 1988-1992 are compared to those presenting between 1996-1998, the frequency of serious bacterial infection appears constant over time. Empirical antibiotic treatment is withheld more frequently in patients without serious bacterial infections in ambulatory care. The length of hospitalisation decreased significantly. Overall costs reduced materially over time. These results confirm a trend towards a less invasive and more selective approach in young children with fever without apparent source, reducing the costs without worsening the clinical outcome.

Acknowledgements

We gratefully acknowledge Wilfried de Jong, Femke Mineur and Haley C. Kang, medical students, for support in data collection. The Health Care Insurance Council of the Netherlands financially supported this project.

References

1. McCarthy PL. Fever. Pediatr in Rev 1998;19:401-8.
2. Baraff LJ, Lee SI. Fever without source: management of children 3 to 36 months of age. Pediatr Infect Dis J 1992;11(2):146-51.
3. Steensel-Moll HA van, Jongkind CJ, Aarsen RSR, et al. [A problem-oriented patient classification system for general paediatrics II]. Tijdschr Kindergeneesk 1996;64:99-104.
4. Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteraemia in young febrile children. Ann Emerg Med 1998;31(6):679-87.
5. Baraff LJ. Management of fever without source in infants and children. Ann Emerg Med 2000;36(6):602-14.
6. Baker MD, Bell LM, Avner JR. Outpatient management without antibiotics of fever in selected infants. N Engl J Med 1993;329(20):1437-41.
7. McCarthy PL, Lembo RM, Fink HD, Baron MA, Cicchetti DV. Observation, history, and physical examination in diagnosis of serious illnesses in febrile children less than or equal to 24 months. J Pediatr 1987;110(1):26-30.
8. Lieu TA, Schwartz JS, Jaffe DM, Fleisher GR. Strategies for diagnosis and treatment of children at risk for occult bacteraemia: clinical effectiveness and cost-effectiveness. J Pediatr 1991;118(1):21-9.

9. Downs SM, McNutt RA, Margolis PA. Management of infants at risk for occult bacteraemia: a decision analysis. Pediatr 1991;118(1):11-20.
10. Baraff LJ, Bass JW, Fleisher GR, et al. Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Paediatrics 1993;92:1-12.
11. Isaacman DJ, Shults J, Gross TK, Davis PH, Harper M. Predictors of bacteraemia in febrile children 3 to 36 months of age. Paediatrics 2000;106(5):977-82.
12. Lee GM, Harper MB. Risk of bacteraemia for febrile young children in the post-*Haemophilus influenzae* type b era. Arch Pediatr Adolesc Med 1998;152(July):624-8.
13. Baker MD, Avner JR, Bell LM. Failure of infant observation scales in detecting serious illness in febrile, 4- to 8-week-old infants. Paediatrics 1990;85(6):1040-3.
14. Bonadio WA. Assessing patient clinical appearance in the evaluation of the febrile child. Am J Emerg Med 1995;13(3):321-6.
15. Dagan R, Powell K, Hall C, Menegus M. Identification of infants unlikely to have serious bacterial infection although hospitalised for suspected sepsis. J Pediatr 1985;107:855-860.
16. Kramer MS, Lane DA, Mills EL. Should blood cultures be obtained in the evaluation of young febrile children without evident focus of bacterial infection? A decision analysis of diagnostic management strategies. Paediatrics 1989;84(1):18-27.
17. Kramer MS, Shapiro ED. Management of the young febrile child: a commentary on recent practice guidelines. Paediatrics 1997;100:128-34.
18. Bleeker SE, Moons KGM, Derksen-Lubsen G, Grobbee DE, Moll HA. Predicting serious bacterial infection in young children with fever without apparent source. Acta Paediatrica 2001;90(11):1226-1232.
19. Neu HC. The crisis in antibiotic resistance. Science 1992;257(5073):1064-73.
20. Friedland IR, McCracken GH, Jr. Management of infections caused by antibiotic-resistant Streptococcus pneumoniae. N Engl J Med 1994;331(6):377-82.
21. Kaplan SL, Mason EO, Jr., Barson WJ, et al. Three-year multicenter surveillance of systemic pneumococcal infections in children. Paediatrics 1998;102(3 Pt 1):538-45.
22. Oostenbrink R, Moons KGM, Donders ART, Grobbee DE, Moll HA. Prediction of bacterial meningitis in children with meningeal signs: reduction of lumbar punctures. Acta Paediatr 2001;90(6):611-7.
23. Oostenbrink R, Moons KGM, Theunissen CC, Derksen-Lubsen G, Grobbee DE, Moll HA. Signs of meningeal irritation at the emergency department: how often bacterial meningitis? Pediatr Emerg Care 2001;17(3):161-4.
24. Oostenbrink R, Oostenbrink JB, Moons KGM, et al. Cost-utility analysis of patient care in children with meningeal signs. Int J Technol Assess 2002; 18(3):485-96.

25. Derksen-Lubsen G, Jongkind CJ, Kraayenoord S, et al. [A problem-oriented patient classification system for general paediatrics I]. Tijdschr Kindergeneeskd 1996;64(3):93-8.
26. Drummond MF, O'Brien B, Stoddart GL, Torrance GW. Methods for the Economic Evaluation of Health Care Programmes. 2nd ed. Oxford: Oxford University Press; 1997.
27. Klein JO, Schlesinger PC, Karasic RB. Management of the febrile infant three months of age or younger. Pediatr Infect Dis 1984;3(1):75-9.
28. McLellan D, Giebink GS. Perspectives on occult bacteraemia in children. J Pediatr 1986;109(1):1-8.
29. Gehlbach SH. Fever in children younger than three months of age. A pooled analysis. J Fam Pract 1988;27(3):305-12.
30. Lieu TA, Baskin MN, Schwartz JS, Fleisher GR. Clinical and cost-effectiveness of outpatient strategies for management of febrile infants. Paediatrics 1992;89(6 Pt 2):1135-44.
31. Bass JW, Steele RW, Wittler RR, et al. Antimicrobial treatment of occult bacteraemia: a multicenter cooperative study. Pediatr Infect Dis J 1993;12(6):466-73.
32. Long SS. Antibiotic therapy in febrile children: "best-laid schemes". J Pediatr 1994;124(4):585-8.
33. Browne GJ, Ryan JM, McIntyeare P. Evaluation of a protocol for selective empiric treatment of fever without localising signs. Arch Dis Child 1997;76(2):129-33.
34. Rothrock SG, Harper MB, Green SM, et al. Do oral antibiotics prevent meningitis and serious bacterial infections in children with Streptococcus pneumoniae occult bacteraemia? A meta-analysis. Paediatrics 1997;99(3):438-44.
35. Jaye DL, Waites KB. Clinical applications of C-reactive protein in paediatrics. Pediatr Infect Dis J 1997;16:735-46; quiz 746-7.

3.1 Predicting serious bacterial infection in young children with fever without apparent source

SE Bleeker, KGM Moons, G Derksen-Lubsen, DE Grobbee, HA Moll

Acta Paediatrica 2001; 90: 1226-1232

Abstract

The aim of this study was to design a clinical rule to predict the presence of a serious bacterial infection in children with fever without apparent source. Information was collected from the records of children, aged 1 month to 36 months, who attended the paediatric emergency department with fever without source (temperature ≥38°C and no apparent source found after evaluation by a general practitioner or history taking by a paediatrician). Serious bacterial infection included bacterial meningitis, sepsis, bacteraemia, pneumonia, urinary tract infection, bacterial gastroenteritis, osteomyelitis and ethmoiditis. Using multivariable logistic regression and the area under the receiver operating characteristic curve (ROC area), the diagnostic value of predictors for serious bacterial infection was judged, resulting in a risk stratification.

Twenty-five percent of the 231 patients enrolled in the study (mean age 1.1 years) had a serious bacterial infection. Independent predictors from history and examination were duration of fever, poor micturition, vomiting, age, temperature <36.7 °C or ≥40 °C at examination, chest-wall retractions and poor peripheral circulation (ROC area (95% confidence interval): 0.75 (0.68-0.83)). Independent predictors from laboratory tests were white blood cell count, serum C-reactive protein and the presence of ≥70 white blood cells in urinalysis (ROC area (95% confidence interval): 0.83 (0.77-0.89)). The risk stratification for serious bacterial infection ranged from 6% to 92%.

Conclusion: The probability of a serious bacterial infection in the individual patient with fever without source can be estimated more precisely by using a limited number of symptoms, signs and laboratory tests.

Introduction

Fever without apparent source constitutes a common diagnostic and therapeutic dilemma for paediatricians. Approximately 10% to 35% of all visits to paediatric emergency departments concern febrile children (1-4) and in 14% to 40% no apparent source is found after a case history and physical examination are undertaken (2, 5). The underlying cause of fever varies from a mild viral to a serious bacterial infection, such as sepsis or meningitis (2). Bacterial infections are reported in 3% to 15% of febrile children (4-6).

Early discrimination of patients with and without a serious bacterial infection would enhance more appropriate management. Recommendations for the evaluation and treatment of febrile children vary greatly (4, 7-16). However, validation of these recommendations appeared to be either poor, or they were based on studies issued before the introduction of the *Haemophilus influenzae* type b vaccines and as a consequence the current value may have been diminished (17, 18). Moreover, it has been argued that some of the earlier recommendations may induce overtesting, overtreatment and even increased penicillin resistance among isolates of *Streptococcus pneumoniae* (19). Furthermore, most of the diagnostic studies in this field did not select patients on presenting symptoms (i.e. children with fever either with or without apparent source) but on the presence or absence of bacteraemia or other serious bacterial infections potentially leading to seriously biased results (20). Finally, univariable rather than the necessary multivariable analysis was commonly applied in most studies (21).

As a consequence the evaluation of febrile children remains a challenging problem for clinicians. The aim of the present study was to develop a diagnostic prediction rule for patients presenting with fever without apparent source, including readily obtainable parameters from the patient history, physical examination and laboratory tests in order to distinguish the patients with a serious bacterial infection from those without a serious bacterial infection.

Patients and methods

Patients

This study was conducted as part of a large ongoing study on paediatric diagnostic management (22, 23) and was approved by The Institutional Review Boards of both participating hospitals. Patients between 1 month and 36 months of age who attended

the emergency department of the Sophia Children's University Hospital Rotterdam (1996-1998) and the Juliana Children's Hospital in The Hague (1998) for the evaluation of acute fever without apparent source (including suspected sepsis) were enrolled. Both hospitals are large inner-city teaching hospitals in the Netherlands. Patient data were retrieved by means of a problem-oriented patient classification system, in which the main reason for encounter after evaluation of the general practitioner or history taking by the paediatrician is classified (3, 24). In this classification the category 'infectious diseases' comprised 1) fever with meningeal signs, 2) fever with cough, 3) fever with micturition problems, 4) fever with vomiting and/or diarrhoea, 5) fever with at least two obvious signs of an upper respiratory tract infection, 6) fever with signs or symptoms of conjunctivitis, 7) fever without apparent source. The last category applied to patients with a body temperature of thirty-eight degrees Celsius or higher and for whom classifications 1 to 6, as described above, were not applicable. Patients not referred by a general practitioner, referred from other hospitals or with immune deficiencies were excluded.

Potential diagnostic determinants

Data were collected by reviewing the standardised medical records. Documented data from patient history and physical examination included information on age, gender, weeks of gestation, body weight, body temperature, duration of fever (body temperature ≥38.0°C), coughing, vomiting, diarrhoea, micturition, intake, crying pattern, vital signs, clinical appearance, fontanelle and information on ear-nose-throat, skin, and the respiratory-, circulatory- and abdominal tract. Data from laboratory tests included haematology, blood chemistry and dipstick urinalysis, which were retrieved from the computer-documented hospital information system.

Reference standard

For each patient the final diagnosis was determined either by a reference standard (cultures of blood, spinal fluid, urine, stool positive for a pathogen) or based on a consensus diagnosis. Outcome diagnosis was the presence or absence of a serious bacterial infection. A serious bacterial infection was defined as the presence of bacterial meningitis, sepsis or bacteraemia, pneumonia, urinary tract infection, bacterial gastroenteritis, osteomyelitis or ethmoiditis. A follow-up period of two weeks was the standard for ruling out the possibility of a missed diagnosis of serious bacterial

infection. Detailed definitions of specific diagnoses are described in a previous report (22).

Statistical analysis

The association between each finding from patient history, physical examination and laboratory tests and the presence of a serious bacterial infection was assessed using univariable logistic regression analyses. Continuous variables were analysed both on a linear and on a transformed scale, i.e. logarithmic or quadratic, to determine which scale was the better predictor of outcome (25). Variables with a univariable p-value of 0.15 or less were subsequently entered into a stepwise multivariable logistic regression procedure. Variables with a multivariable p-value of less than 0.10 were considered to be independent predictors of a serious bacterial infection. First, a subset of independent predictors from the patient's history and physical examination was defined ('clinical model'). Then each univariably significant laboratory test was consecutively added in various orders to estimate its incremental value to the clinical model. The model including the independent predictors from patient history, physical examination and laboratory tests is referred to as the 'clinical + lab model'.

For each model, the ability to discriminate between patients with and without a serious bacterial infection was quantified using the area under the receiver operating characteristic curve (ROC area) (25). The ROC area can range from 0.5 (no discrimination, like flipping a coin) to 1.0 (perfect discrimination). Differences in discriminative value between models were estimated by differences in ROC area with 95% confidence intervals (95%CI), taking into account the correlation between models as they were based on the same cases (26). The reliability (goodness of fit) of each model was quantified using the Hosmer & Lemeshow test (27). Subsequently, we used bootstrapping techniques to validate the defined models, i.e. to adjust the models for optimism (25).

To reduce bias and increase statistical efficiency, missing values in the data were completed by imputation using SOLAS (version 1.2) (28). This method uses all available data to impute the missing values based on the correlation between each variable with missing values and all other variables. Taking into consideration uncertainties in imputed data, multiple imputations were performed by repeating the imputation five times (28).

Table 1. General characteristics and final diagnoses of the 231 patients with fever without apparent source.

Male gender*	122 (53)
Age (years)†	1.1 (0.8)
Weeks of gestation†	39.2 (2.2)
Duration of fever (days)†	2.8 (2.4)
Body temperature at physical examination (°C)†	39.4 (1.0)
Final diagnoses*	
Serious bacterial infection present	*58 (25.1)*
Bacterial meningitis	3 (1.3)
Sepsis/ bacteraemia	3 (1.3)
Pneumonia	28 (12.1)
Bacterial gastroenteritis	2 (0.9)
Urinary tract infection	22 (9.5)
Serious bacterial infection absent	*173 (74.9)*
Aseptic/ viral meningitis	3 (1.3)
Viral gastroenteritis	9 (3.9)
Acute otitis media	7 (3.0)
Viral syndrome‡	104 (45.0)
Miscellaneous§	50 (21.6)

*Absolute number (percentage).

†Mean (standard deviation).

‡Upper respiratory tract infection, respiratory syncytial virus bronchiolitis, non-specified viral infection.

§Drugfever, immunisation reaction, Kawasaki disease, exanthema subitum, vasculitis, malaria, no diagnosis.

Results

In total 231 patients with fever without apparent source were included in the study. The demographic and clinical characteristics and final diagnoses are presented in table 1. The mean age was 1.1 year and 15% of the children were 1 month to 3 months old, 16% 3 to 6 months, 26% 6 to 12 months and 43% were older than 1 year. A serious bacterial infection was present in 25% of the patients. Table 2 shows the variables from patient history, physical examination and laboratory tests that were significantly (p-value <0.15) associated with the presence or absence of a serious bacterial infection in a univariable analysis.

Table 2. Variables from patient history, physical examination and laboratory tests univariably associated ($p<0.15$) with presence or absence of serious bacterial infection (SBI).

Characteristics*	SBI absent 173 (74.9)	SBI present 58 (25.1)
Patient history		
Vomiting	64 (37.0)	33 (56.9)
Poor micturition	57 (32.9)	12 (20.7)
Weeks of gestation†	39.1 (2.2)	39.5 (2.4)
Poor intake	63 (36.4)	15 (25.9)
Duration of fever (days)†	2.6 (2.2)	3.2 (2.8)
Age >1 year	70 (40.5)	30 (51.7)
Physical examination		
Purulent nasal discharge in history or at examination	35 (20.2)	27 (46.6)
Temperature <36.7 or ≥40 at examination (°C)	53 (30.6)	28 (48.3)
Decreased consciousness	6 (3.5)	5 (8.6)
Bulging fontanelle	9 (5.2)	6 (10.3)
Chest-wall retractions ± tachypnoea	9 (5.2)	16 (27.6)
Poor peripheral circulation	19 (11.0)	13 (22.4)
Crepitations	4 (2.3)	4 (6.9)
Bulging abdomen	10 (5.8)	7 (12.1)
Laboratory tests		
White blood cell count (*10^9/L)†	12.9 (6.5)	18.8 (9.0)
Absolute band count†	2.7 (3.6)	5.6 (5.7)
Serum C-reactive protein (mg/L)†	33.2 (44.1)	86.0 (68.7)
Serum haemoglobin†	7.3 (0.9)	7.0 (0.8)
≥70 white blood cells/μl in dipstick urinalysis	39 (22.5)	28 (48.3)
≥25 erythrocytes/μl in dipstick urinalysis	43 (24.9)	40 (69.0)

*Values represent absolute patient numbers (percentages) unless stated otherwise.
†Means (standard deviations).

Independent predictors from patient history and physical examination (i.e. the 'clinical model') included duration of fever at presentation, a history of poor micturition, vomiting, age above 1 year, a body temperature below 36.7 °C or equal to or above 40°C at examination, chest-wall retractions either with or without tachypnoea and poor peripheral circulation (table 3, 'clinical model'). Although of borderline significance, duration of fever and poor peripheral circulation were retained

in the model in view of their clinical relevance. After adding the univariably significant laboratory variables to the 'clinical model', white blood cell (WBC) count, C-reactive protein (CRP) and the presence of 70 or more white blood cells in dipstick urinalysis retained significant and contained additional diagnostic information (table 3, 'clinical + lab model'). The ROC area of the 'clinical model' increased significantly from 0.75 to 0.83.

Table 3. Results from multivariable analyses, corrected for optimism: independent predictors of presence or absence of a serious bacterial infection.

	Clinical model*		**Clinical + Lab model†**	
Characteristics	ß§	Odds ratio (90%CI)	ß	Odds ratio (90%CI)
Duration of fever (days)‡	0.91	2.5 (0.8-7.5)	0.31	1.4 (0.4-5.1)
History of poor micturition	-0.66	0.5 (0.3-1.0)	-0.84	0.4 (0.2-1.0)
History of vomiting	0.82	2.3 (1.2-4.3)	0.76	2.1 (1.0-4.7)
Age >1 year	0.42	1.5 (0.9-2.7)	0.12	1.1 (0.6-2.1)
Temperature <36.7 or ≥40 at examination (°C)	0.52	1.7 (0.9-3.0)	0.54	1.7 (0.8-3.5)
Chest-wall retractions ± tachypnoea	1.59	4.9 (2.3-10.7)	1.55	4.7 (2.0-11.1)
Poor peripheral circulation	0.47	1.6 (0.7-3.6)	0.44	1.6 (0.7-3.5)
White blood cell count (*10⁹/L)			0.04	1.04 (1.0-1.1)
Serum C-reactive protein (mg/L)			0.01	1.01 (1.0-1.01)
≥70 white blood cells/µl in dipstick urinalysis			0.54	1.7 (0.8-3.8)
ROC area (95%CI)		0.75 (0.68-0.83)		0.83 (0.77-0.89)

*Intercept of the model was -2.49.

†Intercept of the model was -3.26.

‡Odds ratio and 90% confidence interval estimated after logarithmic transformation of the characteristic.

§β = Regression coefficient.

These two final models were then transformed to obtain two readily applicable diagnostic scoring rules (see appendix). For each individual patient a 'clinical score' and a 'lab score' were estimated by assigning the corresponding points for each variable present. The 'clinical score' ranged from 0-26, the 'lab score' from 0-40. As

Table 4. Risk of a serious bacterial infection across clinical score categories with and without laboratory tests (n†=231).

			Lab score		
			≤5	5-15	>15
Clinical score		**Without lab***	(n=86)	(n=84)	(n=61)
≤7	(n=112)	10%	6%	7%	28%
7-12	(n=86)	29%	11%	21%	52%
>12	(n=33)	67%	38%	62%	92%

†Number of patients.

*Without lab means using the 'clinical score' only.

the ROC area reflects only the overall discriminative value of a rule and not directly its clinical value in terms of absolute patient numbers (21, 27, 29), we estimated the absolute number of patients across various categories of the estimated scores from the diagnostic rule. Risk stratification based on the 'clinical score' and after addition of laboratory tests ('lab score') is shown in table 4. For example, a patient of 2 years of age, with fever lasting two days, normal micturition, without vomiting, a body temperature of 40.2 °C, without retractions and with normal circulation has a 'clinical score' of 9 (see appendix), corresponding to a 29% risk of contracting a serious bacterial infection at the moment of evaluation (table 4). If the clinician decides on laboratory tests (WBC count, CRP, urine dipstick analysis) for this patient, the probability can change to an 11% to 52% risk, depending on the test results. For instance, if the WBC count was 19 and the CRP level was also 19 and if less than 70 white blood cells were present in the urine dipstick analysis, the 'lab score' would be 5 (see appendix), changing the patient's risk from 29 % to 11% (table 4). However, the risk would be 52% if the WBC count and CRP were, for example, 25 and 84, respectively. Table 4 shows that in patients with fever without apparent source, additional laboratory tests are useful, as the risk of an individual patient having a serious bacterial infection can be estimated precisely.

Discussion

This study of patients with fever without apparent source has identified predictors that can be helpful in distinguishing between patients with and without serious bacterial infections based on patient history, physical examination and additional

laboratory tests. The predictors include duration of fever, a history of poor micturition, vomiting, age above 1 year, a body temperature below 36.7 °C or equal to or above 40 °C at examination, chest-wall retractions either with or without tachypnoea, poor peripheral circulation, WBC count, CRP and the presence of 70 or more white blood cells in dipstick urinalysis.

Our study has followed the diagnostic work-up as generally applied in medical practice as far as possible in both design (selection of patients on problem of referral) and analysis (applying chronological multivariable modelling). This has resulted in a validated evidence-based and readily applicable diagnostic rule including a limited number of symptoms, signs and laboratory tests. This rule allows a more precise estimate of the probability of a serious bacterial infection in the individual patient compared to the prior probability (prevalence).

To appreciate the results several aspects need to be addressed. First, inclusion criteria and definition of fever without apparent source and outcome differ widely across studies (5-7, 10, 12-14, 30). We have defined fever without apparent source as fever without source after evaluation by the general practitioner or after history taking by the paediatrician, i.e. before physical examination by the paediatrician. Our motivation for this was the intention to discriminate patients with and without a serious bacterial infection as early as possible in the diagnostic work-up, preferably with history (other than focal symptoms) and physical examination data only. Our definition of the outcome serious bacterial infection was based on the most consistent literature (5, 6, 12, 13). Second, in our study the prevalence of serious bacterial infection is high (25%) compared with that reported in other studies (4-6). This is probably due to different inclusion criteria (e.g. febrile children versus children with fever without focus) and other outcome definitions (e.g. serious bacterial infection including versus excluding the final diagnosis of pneumonia). These discrepancies in inclusion criteria and outcome variables hinder comparisons of values across studies. Third, it has been argued whether febrile infants presenting with focal signs of infection pose a diagnostic or a therapeutic dilemma (12, 31). Our study shows that they do pose such dilemmas. This may be further illustrated by an example based on our data: crepitations at physical examination have been found in only 4 patients out of 28 with pneumonia and also in 4 patients out of 203 subjects without pneumonia (positive predictive value: 50%, sensitivity: 14%), which shows that focal signs can be non-specific. Fourth, the outcome in children with fever without apparent source is heterogeneous, as reflected in the diversity of final diagnoses (table 1). This

complicates both the clinical approach and the design of scientific studies. Heterogeneity in outcome makes it difficult to make a complete distinction between patients with and without the outcome. Nonetheless, the discriminative value of our scoring rule was 0.83.

Our results partly agree and partly disagree with the literature. The predictors selected in the present study have been discussed extensively in other published works (1, 2, 4, 7, 8, 12, 15, 16, 30-35). Toxic appearance, components of the Yale Observation Scale (quality of cry, state variation, hydration), poor intake and absolute neutrophil counts are other frequently described characteristics associated with bacteraemia or serious infections (4, 7, 12, 15, 16, 30, 31). We have considered all these factors, but have not been able to confirm their independent diagnostic value. Nonetheless, our predictors have never ever been combined into one diagnostic rule as derived and validated by a multivariable approach.

In this study the *a priori* risk for a serious bacterial infection in an individual patient presenting with fever without apparent source is 25%. Based on patient history and physical examination only, the risk varies between 10% and 67% ('clinical score') and the value of clinical screening has been stated already (7, 8). By adding laboratory results further modification is possible ranging from 6% to 92% ('clinical + lab score'). This shows that laboratory evaluation is of substantial additional value in these patients. Nevertheless, table 4 indicates that if the 'clinical score' exceeds twelve, the management with respect to treatment and hospital admission is evident anyhow and probably would not change even with knowledge of additional laboratory tests. Of the patients with a 'clinical score' ≤ 7 without laboratory evaluation, 10% were identified as having serious bacterial infections (table 4). Of the fifty-one patients with a 'clinical score' ≤ 7 and a 'lab score' ≤ 5, three patients (6%) eventually were diagnosed as having a serious bacterial infection (all urinary tract infections, ages 2, 3 and 6 months). In forty-three patients with a 'clinical score' ≤ 7 and a 'lab score' 5-15 as well three patients were determined as having a serious bacterial infection (1 case of bacterial meningitis, at age 2 months; 1 case of bacterial gastroenteritis, at age 3 months; 1 case of pneumonia, at age 7 months). This stresses the difficulty in selecting patients with a serious bacterial infection, particularly among 'low risk' patients (i.e. 'clinical score' ≤ 7).

In conclusion, diagnostic and therapeutic management in children with fever without apparent source remains complex. In this paper we describe a diagnostic rule by which the risk for the presence of a serious bacterial infection in an individual

patient can be estimated. This estimate is meant to augment but not to replace clinical acumen, and the clinician should make a judgement in combination with experience and knowledge.

Acknowledgements

We express our gratitude to R. Oostenbrink, MD, PhD for support in research procedures and data collection, M.J. Twijnstra, MSc for data management, A.R.T. Donders, PhD for assisting in the imputation analyses and Professor R. de Groot, MD, PhD for critically reviewing the manuscript. The Health Care Insurance Council of the Netherlands financially supported this project.

References

1. McCarthy PL. Fever. Pediatr in Rev 1998;19:401-8.
2. Baraff LJ, Lee SI. Fever without source: management of children 3 to 36 months of age. Pediatr Infect Dis J 1992;11:146-51.
3. Steensel-Moll HA van, Jongkind CJ, Aarsen RSR, et al. [A problem-oriented patient classification system for general paediatrics II]. Tijdschr Kindergeneesk 1996;64:99-104.
4. Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteraemia in young febrile children. Ann Emerg Med 1998;31:679-87.
5. Baker MD, Bell LM, Avner JR. Outpatient management without antibiotics of fever in selected infants. N Engl J Med 1993;329:1437-41.
6. Baraff LJ, Oslund SA, Schriger DL, Stephen ML. Probability of bacterial infections in febrile infants less than three months of age: a meta-analysis. Pediatr Infect Dis J 1992;11:257-64.
7. McCarthy PL, Lembo RM, Baron MA, Fink HD, Cicchetti DV. Predictive value of abnormal physical examination findings in ill-appearing and well-appearing febrile children. Paediatrics 1985;76:167-71.
8. McCarthy PL, Lembo RM, Fink HD, Baron MA, Cicchetti DV. Observation, history, and physical examination in diagnosis of serious illnesses in febrile children less than or equal to 24 months. J Pediatr 1987;110:26-30.
9. Kramer MS, Lane DA, Mills EL. Should blood cultures be obtained in the evaluation of young febrile children without evident focus of bacterial infection? A decision analysis of diagnostic management strategies. Paediatrics 1989;84:18-27.
10. Lieu TA, Schwartz JS, Jaffe DM, Fleisher GR. Strategies for diagnosis and treatment of children at risk for occult bacteraemia: clinical effectiveness and cost-effectiveness. J Pediatr 1991;118:21-9.

11. Downs SM, McNutt RA, Margolis PA. Management of infants at risk for occult bacteraemia: a decision analysis. Pediatr 1991;118:11-20.
12. Baraff LJ, Bass JW, Fleisher GR, et al. Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Paediatrics 1993;92:1-12.
13. Dagan R, Powell K, Hall C, Menegus M. Identification of infants unlikely to have serious bacterial infection although hospitalised for suspected sepsis. J Pediatr 1985;107:855-860.
14. Bonadio WA. Assessing patient clinical appearance in the evaluation of the febrile child. Am J Emerg Med 1995;13:321-6.
15. Isaacman DJ, Shults J, Gross TK, Davis PH, Harper M. Predictors of bacteraemia in febrile children 3 to 36 months of age. Paediatrics 2000;106:977-82.
16. Lee GM, Harper MB. Risk of bacteraemia for febrile young children in the post-*Haemophilus influenzae* type b era. Arch Pediatr Adolesc Med 1998;152:624-8.
17. Teach SJ, Fleisher GR. Efficacy of an observation scale in detecting bacteraemia in febrile children three to thirty-six months of age, treated as outpatients. Occult Bacteraemia Study Group. J Pediatr 1995;126:877-81.
18. Baker MD, Avner JR, Bell LM. Failure of infant observation scales in detecting serious illness in febrile, 4- to 8-week-old infants. Paediatrics 1990;85:1040-3.
19. Kramer MS, Shapiro ED. Management of the young febrile child: a commentary on recent practice guidelines. Paediatrics 1997;100:128-34.
20. Knottnerus JA, Leffers JP. The influence of referral patterns on the characteristics of diagnostic tests. J Clin Epidemiol 1992;45:1143-1154.
21. Moons KG, Es GA van, Michel BC, Büller HR, Habbema JD, Grobbee DE. Redundancy of single diagnostic test evaluation. Epidemiology 1999;10:276-81.
22. Oostenbrink R, Theunissen CCW, Moons KGM, Derksen-Lubsen G, Grobbee DE, Moll HA. Signs of meningeal irritation at the emergency department; how often bacterial meningitis? Ped Emerg Care, 2001;17:161-4.
23. Oostenbrink R, Moons KGM, Donders ART, Grobbee DE, Moll HA. Prediction of bacterial meningitis in children with meningeal signs: reduction of lumbar punctures. Acta Paediatr, 2001;90:611-7.
24. Derksen-Lubsen G, Jongkind CJ, Kraayenoord S, et al. [A problem-oriented patient classification system for general paediatrics I]. Tijdschr Kindergeneeskd 1996;64:93-8.
25. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-87.
26. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839-43.
27. Hosmer DW, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons, Inc; 1989.

28. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. Am J Epidemiol 1995;142:1255-64.
29. Moons KGM, Stijnen T, Es GA van, et al. Treatment thresholds in diagnostic test evaluation: an alternative to the comparison of areas under receiver operating characteristic curves. Med Decis Making 1997;17:447-54.
30. Hewson PH, Humphries SM, Roberton DM, McNamara JM, Robinson MJ. Markers of serious illness in infants under 6 months old presenting to a children's hospital. Arch Dis Child 1990;65:750-6.
31. Berger RM, Berger MY, Steensel-Moll HA van, Dzoljic-Danilovic G, Derksen-Lubsen G. A predictive model to estimate the risk of serious bacterial infections in febrile infants. Eur J Pediatr 1996;155:468-73.
32. Teach SJ, Fleisher GR. Duration of fever and its relationship to bacteraemia in febrile outpatients three to 36 months old. The Occult Bacteraemia Study Group. Pediatr Emerg Care 1997;13:317-9.
33. Jaffe DM, Fleisher GR. Temperature and total white blood cell count as indicators of bacteraemia. Paediatrics 1991;87:670-4.
34. Jaye DL, Waites KB. Clinical applications of C-reactive protein in paediatrics. Pediatr Infect Dis J 1997;16:735-46; quiz 746-7.
35. Kramer MS, Tange SM, Mills EL, Ciampi A, Bernstein ML, Drummond KN. Role of the complete blood count in detecting occult focal bacterial infection in the young febrile child. J Clin Epidemiol 1993;46:349-57.

Appendix

'Clinical score'.

Variable	Points to assign		Score
Duration of fever	Days	Points	
	½-1	1	
	1½-2	2	
	2½-4½	3	
	5-8	4	
	≥8½	5	
History of poor micturition	no = 0/ yes = -3		
History of vomiting	no = 0/ yes = 4		
Age	≤1 year = 0		
	>1 year = 2		
Temperature <36.7 or ≥40 at examination (°C)*	no = 0/ yes = 3		
Chest-wall retractions ± tachypnoea	no = 0/ yes = 8		
Poor peripheral circulation	no = 0/ yes = 2		
			Total +2 =

*If temperature at examination is <38°C, then temperature in history must be ≥38°C.

'Lab score'.

Variable	Points to assign		Score
White blood cell count	Count	Points	
(*10^9/L)	<10	0	
	10-19	4	
	20-29	8	
	30-39	12	
	≥40	16	
Serum C-reactive protein	0-99:	1st integer, e.g.	(max 16)
(mg/L)		CRP=*84*: *8* points	
	≥100	1st and 2nd integers, e.g.	
		CRP=*132*: *13* points,	
		maximum=16 points	
≥70 white blood cells/µl			
in dipstick urinalysis	no = 0/ yes = 8		
			Total =

3.2 Self-referred young patients with fever without apparent source: predicting serious bacterial infections

SE Bleeker, KGM Moons, G Derksen-Lubsen, DE Grobbee, HA Moll

Submitted

Abstract

Objective: To predict the presence of serious bacterial infections (SBIs) in self-referred patients presenting with fever without apparent source. *Design:* The generalisability of a previously developed prediction model for the presence of SBIs in referred patients with fever without apparent source to self-referred patients was determined. Subsequently, an improved prediction model for the presence of SBIs was derived on self-referred patients with fever without apparent source using multivariable logistic regression, resulting in a risk stratification. *Setting:* Paediatric emergency department, urban. *Patients:* 1 month to 36 months old, presenting with fever without apparent source between 1997-1998. Patients were classified according to the presence of a SBI (bacterial meningitis, sepsis/ bacteraemia, pneumonia, urinary tract infection, bacterial gastroenteritis, osteomyelitis/ ethmoiditis). *Interventions:* None. *Outcome Measures:* The discriminating performance (area under the receiver operating characteristic curve (95% confidence interval): ROC area (95%CI)) of a prediction model. *Results:* The ROC area (95%CI) of the prediction model for referred patients was 0.60 (0.47-0.73) in the self-referred patients. Independent predictors from history and examination for the presence of SBIs in self-referred patients were duration of fever at presentation at the emergency department, age and degree of body temperature at examination (ROC area (95%CI): 0.70 (0.58-0.82)). Independent additional predictors were white blood cell count and serum C-reactive protein (ROC area (95%CI): 0.81 (0.71-0.91)). The risk stratification for SBIs ranged from 4% to 83%. *Conclusions:* A limited number of symptoms and signs are useful as a first screening in estimating the risk having a SBI, in an individual self-referred patient with fever without apparent source. Additional laboratory evaluation can specify this risk further.

Introduction

Fever frequently occurs in children. It is one of the most common problems for which parents seek medical care (1-4). The underlying cause varies widely from a mild viral infection to a serious bacterial infection, such as sepsis or meningitis (3% to 15%) (2, 4-6).

Patients visiting the emergency department (ED) by self-referral are growing in number, especially among young children (7, 8). Depending on the domain (e.g. age of patients, in- or excluding trauma cases), 35% to 75% of the ED attendees are self-referred (7-11). It has been suggested that self-referred patients more often visit the ED for non-urgent use than referred patients. In particular, infectious diseases appear a reason for inappropriate use of EDs, which may amount from 17% to 83% (7, 12). Nonetheless, part of the self-referred patients does present with emergent problems requiring paediatric care. So, early discrimination of non-urgent and emergent problems in self-referred patients, presenting with fever without apparent source, would optimise management.

The aim of the present study was to obtain a diagnostic prediction rule for self-referred patients presenting with fever without apparent source, in order to distinguish the patients with a serious bacterial infection from those without a serious bacterial infection. This included the determination of the generalisability of a previously developed prediction model for the presence of serious bacterial infections in patients referred by a general practitioner with fever without apparent source to self-referred patients as well (13).

Patients and methods

Patients

Self-referred patients between 1 month and 36 months of age who attended the ED of the Sophia Children's University Hospital Rotterdam (1997-1998) and the Juliana Children's Hospital in The Hague (1998) were enrolled. Both hospitals are large inner-city teaching hospitals in the Netherlands. Reason for attending had to be evaluation of acute fever without apparent source, including suspected sepsis. Patient data were retrieved by means of a problem-oriented patient classification system, in which the main reason for encounter after evaluation of the general practitioner or history taking by the paediatrician is classified (3, 14). In this classification the category 'infectious diseases' comprised 1) fever with meningeal signs, 2) fever with cough, 3) fever with

micturition problems, 4) fever with vomiting and/or diarrhoea, 5) fever with at least two obvious signs of an upper respiratory tract infection, 6) fever with signs or symptoms of conjunctivitis, 7) fever without apparent source. The last category applied to patients with a body temperature of at least thirty-eight degrees Celsius and for whom classifications 1 to 6, as described above, were not applicable. Patients with immune deficiencies were excluded.

This study formed part of an ongoing study on diagnostic management in acute paediatric patients and was approved by The Institutional Review Boards of both participating hospitals (13, 15, 16).

Outcome

Referral status of patients was obtained from the records and was noted as either referred by a general practitioner or as self-referred. Self-referred patients were defined as patients who bypassed primary care and independently visited the ED for paediatric care (10).

For each patient, the final diagnosis was determined either by a reference standard (cultures of blood, spinal fluid, urine, stool positive for a pathogen) or based on a consensus diagnosis. Based on the final diagnosis patients were classified according to the presence or absence of a serious bacterial infection. A serious bacterial infection was defined as the presence of bacterial meningitis, sepsis or bacteraemia, pneumonia, urinary tract infection, bacterial gastroenteritis, osteomyelitis or ethmoiditis (5, 6, 17, 18). A follow-up period of two weeks was the standard for ruling out the possibility of a missed diagnosis of serious bacterial infection. Detailed definitions of specific diagnoses have been described in previous reports (13, 16).

Validation of a 'referred patients model' on self-referred patients with fever without source

In a previous study, a model to predict the presence of serious bacterial infections in patients presenting with fever without apparent source was developed (13). The population on which this model was developed comprised 231 patients who were referred by the general practitioner for the evaluation of fever without apparent source to the EDs of the Sophia Children's University Hospital Rotterdam (1997-1998) and the Juliana Children's Hospital in The Hague (1998). The prediction model comprised seven variables from patient history and physical examination: duration of fever at the moment of presentation at the ED, a history of poor micturition, vomiting, age above 1 year, a body temperature below 36.7 °C or equal to or above 40 °C at examination,

chest-wall retractions either with or without tachypnoea and poor peripheral circulation ('referred patients model'). This model was internally validated using bootstrap methods (13).

To determine whether this 'referred patients model' could be extended to self-referred patients, the 'referred patients model' was applied to the group of self-referred patients. Its discriminating performance as reflected by the area under the receiver operating characteristic curve (ROC area) of this 'referred patients model' was estimated in the group of self-referred patients and compared to that found in the group of referred patients.

Derivation of a new prediction model on self-referred patients with fever without source

Data of the self-referred patients were collected by reviewing the standardised medical records. Documented data included information from patient history, physical examination and laboratory tests (13). The potential diagnostic determinants that were evaluated were the same seven predictors as included in the 'referred patients model' and additionally, based on published works, changed crying pattern (<1 year), diarrhoea, poor intake, ill clinical appearance, bulging fontanelle (<1 year), serum white blood cell count, serum C-reactive protein and white blood cells in dipstick urinalysis (1, 2, 4, 17-26).

The association between each finding from patient history, physical examination and laboratory tests and the presence of a serious bacterial infection was assessed using univariable logistic regression analyses. Variables with a p-value of 0.15 or less were subsequently entered into a stepwise multivariable logistic regression procedure. Variables with a multivariable p-value of less than 0.10 were considered to be independent predictors of a serious bacterial infection. First, a subset of independent predictors from the patient's history and physical examination was defined ('clinical model'). Then each univariably selected laboratory test was consecutively added in various orders to estimate its incremental value to the clinical model. The model including the independent predictors from patient history, physical examination and laboratory tests is referred to as the 'clinical + lab model'.

For each model, the ability to discriminate between patients with and without a serious bacterial infection was quantified using the ROC area (27). The ROC area can range from 0.5 (no discrimination, like flipping a coin) to 1.0 (perfect discrimination). Differences in discriminative value between models were estimated by differences in ROC area with 95% confidence intervals (95%CI), taking into account the correlation

between models as they were based on the same cases (28). Subsequently, we used bootstrapping techniques to validate the defined models, i.e. to adjust the models for optimism (27).

Some variables had missing values. As simple exclusion of records with missing values commonly causes biased results and decreases statistical efficiency, missing values in the data were completed by imputation using SOLAS (version 2.0) (29, 30). This method uses all available data to impute the missing values based on the correlation between each variable with missing values and all other variables (30). Taking into consideration uncertainties in imputed data, multiple imputations were performed by repeating the imputation five times (29).

Results

In total 109 self-referred patients with fever without apparent source were included in the study, which was 38% of all patients presenting with that problem. The demographic and clinical characteristics and the final diagnoses are presented in table 1. The mean age was 1.1 year. A serious bacterial infection was present in 19% of the patients.

When applying the 'referred patients model' to the self-referred patients with fever without apparent source, the ROC area (95%CI) was 0.60 (0.47-0.73), whereas this was 0.75 (0.68-0.83) in the group of referred patients. Because of the poor discriminative ability, we rated this 'referred patients model' to be inadequate to classify future self-referred patients presenting with fever without apparent source. The model could not be generalised to the self-referred patients.

Table 1. General characteristics and final diagnoses of the 109 self-referred patients with fever without apparent source*.

Male gender	59 (54)
Age (years) †	1.1 (0.7)
Weeks of gestation†	39.2 (2.5)
Duration of fever at presentation (days) †	2.2 (1.9)
Body temperature at physical examination (°C) †	39.7 (0.9)
Serious bacterial infection present	21 (19)
Hospitalisation	24 (22)

*Values represent absolute patient numbers (percentages) unless stated otherwise; †Mean (standard deviation).

Hence, a new prediction model for the presence of serious bacterial infections was developed on the data from 109 self-referred patients with fever without apparent source. Table 2 shows the variables from patient history, physical examination and laboratory tests that were associated ($p<0.15$) with the presence or absence of a serious bacterial infection in a univariable analysis. Independent predictors from patient history and physical examination (i.e. the 'clinical model') were duration of fever at the moment of presentation at the ED, age and degree of body temperature at examination (table 3, 'clinical model'). After adding the univariably significant laboratory variables to the 'clinical model', serum white blood cell count and C-reactive protein retained significant and contained additional diagnostic information (table 3, 'clinical + lab model'). The ROC area of the 'clinical model' increased significantly from 0.70 to 0.81.

Table 2. Variables from patient history, physical examination and laboratory tests univariably associated ($p<0.15$) with presence or absence of serious bacterial infection (SBI) in self-referred patients with fever without apparent source*.

Characteristics	SBI absent 88 (81)	SBI present 21 (19)
Patient history		
Duration of fever at presentation (days)†	2.0 (1.8)	2.8 (2.2)
Changed crying pattern (only <1 year)	31 (35)	11 (52)
Age†	1.1 (0.7)	0.8 (0.7)
Poor intake	23 (26)	10 (48)
Physical examination		
Temperature at examination (°C)†	39.6 (0.9)	39.9 (0.8)
Ill clinical appearance	44 (50)	16 (76)
Chest-wall retractions ± tachypnoea	10 (11)	4 (19)
Laboratory tests		
White blood cell count (*10^9/L)†	13.1 (7.0)	20.3 (7.9)
Serum C-reactive protein (mg/L)†	26.5 (31.3)	77.2 (43.3)
≥70 white blood cells/μl in dipstick urinalysis	23 (26)	10 (48)

*Values represent absolute patient numbers (percentages) unless stated otherwise.
†Means (standard deviations).

Table 3. Independent predictors of presence or absence of a serious bacterial infection in self-referred patients with fever without apparent source. Results from multivariable analyses, corrected for optimism.

	Clinical model*		**Clinical + Lab model†**	
Characteristics	β‡	Odds ratio (90%CI)	β	Odds ratio (90%CI)
Duration of fever at presentation (days)	0.21	1.2 (1.0-1.6)	0.10	1.1 (0.8-1.6)
Age >1 year	-0.90	0.4 (0.2-1.0)	-1.28	0.3 (0.1-1.0)
Temperature at examination (°C)	0.50	1.7 (0.9-3.0)	0.59	1.8 (0.8-3.9)
Serum C-reactive protein (mg/L)			0.03	1.03 (1.01-1.05)
White blood cell count (*10^9/L)			0.07	1.08 (0.98-1.18)
ROC area (95%CI)		0.70 (0.58-0.82)		0.81 (0.71-0.91)

*Intercept of the model –32.1.

†Intercept of the model –34.7.

‡β = Regression coefficient.

These two final models were then transformed to obtain two readily applicable diagnostic scoring rules (see appendix). For each individual patient a 'clinical score' and a 'lab score' were estimated by assigning the corresponding points for each variable present. The 'clinical score' ranged from 0-22, the 'lab score' from 0-66. As the ROC area reflects only the overall discriminative value of a rule and not directly its clinical value in terms of absolute patient numbers, we estimated the absolute number of patients across various categories of the estimated scores from the diagnostic rule (31-33). Risk stratification based on the 'clinical score' and after addition of laboratory tests ('lab score') is shown in table 4. For example, a patient of 9 months of age, with fever lasting two and a half days and a body temperature of 40.2°C has a 'clinical score' of 11 (see appendix: 0+3+5+3). This corresponds to a 32% risk of contracting a serious bacterial infection at the moment of evaluation (table 4). If the clinician decides on laboratory tests (WBC count, CRP) for this patient, the probability can change to a 5% or 83% risk, depending on the test results. For instance, if the WBC count was 19 and the CRP level was also 19, the 'lab score' would be 10 (see appendix: 7+3), changing the patient's risk from 32% to 5% (table 4). However, the

Table 4. Risk of a serious bacterial infection across clinical score categories with and without laboratory tests (n*=109).

			Lab score	
			0-24	25-66
Clinical score		Without lab†	(n=79)	(n=30)
≤6	(n=36)	6%	4%	14%
7-8	(n=39)	20%	7%	55%
≥9	(n=34)	32%	5%	83%

*Number of patients.

†Without lab means using the 'clinical score' only.

risk would be 83% if the WBC count and CRP were, for example, 25 and 80, respectively ('lab score' of 14+24=38).

Discussion

In this study, predictors for the presence of serious bacterial infections in self-referred patients presenting with fever without apparent source are described. These include duration of fever, age, degree of body temperature at examination, serum white blood cell count and C-reactive protein.

Our study has followed the diagnostic work-up as generally applied in medical practice as far as possible in both design (selection of patients on problem of referral) and analysis (applying chronological multivariable modelling). This has resulted in a readily applicable diagnostic rule including a limited number of symptoms, signs and laboratory tests. This rule provides a more precise estimate of the probability of a serious bacterial infection in the individual patient compared to the prior probability (prevalence). The *a priori* risk for a serious bacterial infection in an individual patient presenting with fever without apparent source is 19%. Based on patient history and physical examination the risk varies between 6% and 32% ('clinical score'), supporting the importance of clinical screening (19, 20). By adding laboratory results, further discrimination is possible, with probabilities ranging from 4% to 83% ('clinical + lab score'). Of the patients with a 'clinical score' ≤6 without laboratory evaluation, two patients (6%) were identified as having serious bacterial infections (both pneumoniae, ages 1.4 and 2.8 years) (table 4).

The proportion of self-referral in this study (38%) is similar to other studies (7-11). The predictors identified in this study have been discussed in other published

works too (1, 2, 4, 13, 17, 19-21, 23-26). However, these predictors have never ever been combined into one prediction model as derived by a multivariable approach.

Prospective methods, such as nurse triage, for identifying patients attending EDs with problems that could be managed appropriately in general practice, have been developed (12). In this study, however, we have identified predictors for self-referred patients presenting with the problem fever without apparent source at the ED, in particular. To increase efficiency and reduce costs, it might be of interest to prospectively evaluate the value of the proposed 'clinical model' as a first screening instrument applied by nurse practitioners (nurse triage) in self-referred patients. For instance, based on such nurse triage, the self-referred patient could be sent back to the general practitioner or, optionally after laboratory evaluation requested by the nurse practitioner, evaluated by the paediatrician. All three clinical variables of the screening instrument would be readily obtainable by nurse practitioners (duration of fever at the moment of presentation, age and degree of body temperature). Of course, an approach based on nurse triage has to be validated first, before implementation in practice.

Applying an earlier developed prediction model for referred patients to the self-referred patients yielded a poor discriminative power. This was not a complete surprise, as referral pattern influences the composition of a patient population (34). When a patient is referred by a general practitioner, the patient already underwent a diagnostic work-up (patient history, physical examination and optionally laboratory evaluation) before presentation at the ED. Hence, the general practitioner already filters out the non-urgent patients, and potentially urgent patients will be sent to the hospital. In the self-referred group this filtering by the general practitioner is lacking, resulting in a more heterogeneous group. Therefore, referred patients and self-referred patients should be considered as different patient populations. Duration of fever at presentation, age and degree of body temperature have been identified as predictors in both the referred and the self-referred patients. Probably, this is due to the heterogeneity of the outcome in children with fever without apparent source, resulting in more general than specific predictors. However, the relative weight of the predictors is different across both groups. This is reflected by different odds ratios.

To appreciate the results, several aspects need to be addressed. We have defined fever without apparent source as fever without source after history taking by the paediatrician, i.e. before physical examination by the paediatrician. Our motivation for this was the intention to discriminate patients with and without serious bacterial

infections as early as possible in the diagnostic work-up, preferably with history (other than focal symptoms) and physical examination data only. Our definition of the outcome serious bacterial infection agrees with most literature (5, 6, 17, 18). The prevalence of serious bacterial infection is higher (19%) than in some previous reports (4-6). Different inclusion criteria (e.g. febrile children versus children with fever without focus) and other outcome definition (e.g. serious bacterial infection including versus excluding the final diagnosis of pneumonia) probably cause this. Because of the small group of self-referred patients, this study has limited power. As a consequence, an unstable selection of variables, using stepwise selection methods, may arise (35-38). Nonetheless, the results agree with literature, and bootstrapping has further documented robustness of the estimates. However, although internally validated by bootstrap methods, external validation of the prediction model is necessary before implementation in paediatric practice (39-41).

Bypassing primary care by self-referral of patients to paediatric care is on the increase. It may endanger continuity of health care and increase costs (8, 11, 42, 43). Hence, discrimination of non-urgent patients and urgent patients early in presentation is important.

In conclusion, risk stratification shows that readily obtainable variables from the patient's history and physical examination are useful as a first screening. The risk having a serious bacterial infection can be estimated in each individual self-referred patient with fever without apparent source. Guided by such screening the clinician can decide for each patient whether additional laboratory evaluation would be useful for further specification of this risk. This estimate is meant to augment but not to replace clinical acumen, and the clinician should judge it in combination with experience and knowledge.

Acknowledgements

We gratefully acknowledge A.R.T. Donders, PhD for assisting in the imputation analyses, and Wilfried de Jong, Femke Mineur and Haley C. Kang, medical students, for support in data collection. The Health Care Insurance Council of the Netherlands financially supported this project.

References

1. McCarthy PL. Fever. Pediatr in Rev 1998;19:401-8.
2. Baraff LJ, Lee SI. Fever without source: management of children 3 to 36 months of age. Pediatr Infect Dis J 1992;11(2):146-51.
3. Steensel-Moll HA van, Jongkind CJ, Aarsen RSR, et al. [A problem-oriented patient classification system for general paediatrics II]. Tijdschr Kindergeneesk 1996;64(3):99-104.
4. Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteraemia in young febrile children. Ann Emerg Med 1998;31(6):679-87.
5. Baker MD, Bell LM, Avner JR. Outpatient management without antibiotics of fever in selected infants. N Engl J Med 1993;329(20):1437-41.
6. Baraff LJ, Oslund SA, Schriger DL, Stephen ML. Probability of bacterial infections in febrile infants less than three months of age: a meta-analysis. Pediatr Infect Dis J 1992;11(4):257-64.
7. Sharma V, Simon SD, Bakewell JM, Ellerbeck EF, Fox MH, Wallace DD. Factors influencing infant visits to emergency departments. Paediatrics 2000;106(5):1031-9.
8. Bedford HE, Jenkins SM, Shore C, Kenny PA. Use of an east end children's accident and emergency department for infants: a failure of primary health care? Qual Health Care 1992;1(1):29-33.
9. Rothschild M, Gilboa S, Sagi H, Berger I, Wolach B. Referral, admission, and discharge patterns in a paediatric emergency department in Israel. Pediatr Emerg Care 1993;9(2):72-6.
10. Fieselmann J, Hendryx M. Characteristics of self-reffered patients. Academic Medicine 1994;69(5):376-81.
11. Beebe SA, Casey R, Magnusson MR, Pasquariello PS, Jr. Comparison of self-referred and physician-referred patients to a paediatric diagnostic center. Clin Pediatr (Phila) 1993;32(7):412-6.
12. Dale J, Green J, Reid F, Glucksman E. Primary care in the accident and emergency department: I. Prospective identification of patients. BMJ 1995;311:423-6.
13. Bleeker SE, Moons KGM, Derksen-Lubsen G, Grobbee DE, Moll HA. Predicting serious bacterial infection in young children with fever without apparent source. Acta Paediatrica 2001;90(11):1226-1232.
14. Derksen-Lubsen G, Jongkind CJ, Kraayenoord S, et al. [A problem-oriented patient classification system for general paediatrics I]. Tijdschr Kindergeneeskd 1996;64(3):93-8.
15. Oostenbrink R, Moons KGM, Donders ART, Grobbee DE, Moll HA. Prediction of bacterial meningitis in children with meningeal signs: reduction of lumbar punctures. Acta Paediatr 2001;90(6):611-7.
16. Oostenbrink R, Moons KGM, Theunissen CC, Derksen-Lubsen G, Grobbee DE, Moll HA. Signs of meningeal irritation at the emergency department: how often bacterial meningitis? Pediatr Emerg Care 2001;17(3):161-4.

17. Baraff LJ, Bass JW, Fleisher GR, et al. Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Paediatrics 1993;92:1-12.
18. Dagan R, Powell K, Hall C, Menegus M. Identification of infants unlikely to have serious bacterial infection although hospitalised for suspected sepsis. J Pediatr 1985;107:855-860.
19. McCarthy PL, Lembo RM, Baron MA, Fink HD, Cicchetti DV. Predictive value of abnormal physical examination findings in ill-appearing and well-appearing febrile children. Paediatrics 1985;76(2):167-71.
20. McCarthy PL, Lembo RM, Fink HD, Baron MA, Cicchetti DV. Observation, history, and physical examination in diagnosis of serious illnesses in febrile children less than or equal to 24 months. J Pediatr 1987;110(1):26-30.
21. Lee GM, Harper MB. Risk of bacteraemia for febrile young children in the post-*Haemophilus influenzae* type b era. Arch Pediatr Adolesc Med 1998;152(July):624-8.
22. Hewson PH, Humphries SM, Roberton DM, McNamara JM, Robinson MJ. Markers of serious illness in infants under 6 months old presenting to a children's hospital. Arch Dis Child 1990;65(7):750-6.
23. Berger RM, Berger MY, Steensel-Moll HA van, Dzoljic-Danilovic G, Derksen-Lubsen G. A predictive model to estimate the risk of serious bacterial infections in febrile infants. Eur J Pediatr 1996;155(6):468-73.
24. Jaffe DM, Fleisher GR. Temperature and total white blood cell count as indicators of bacteraemia. Paediatrics 1991;87(5):670-4.
25. Jaye DL, Waites KB. Clinical applications of C-reactive protein in paediatrics. Pediatr Infect Dis J 1997;16:735-46; quiz 746-7.
26. Teach SJ, Fleisher GR. Duration of fever and its relationship to bacteraemia in febrile outpatients three to 36 months old. The Occult Bacteraemia Study Group. Pediatr Emerg Care 1997;13(5):317-9.
27. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine 1996;15(4):361-87.
28. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148(3):839-43.
29. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. Am J Epidemiol 1995;142(12):1255-64.
30. Little RA. Regression with missing X's: A review. J Am Stat Assoc 1992;87:1227-1237.
31. Moons KG, Es GA van, Michel BC, Buller HR, Habbema JD, Grobbee DE. Redundancy of single diagnostic test evaluation. Epidemiology 1999;10(3):276-81.
32. Hosmer DW, Lemeshow S. Applied logistic regression. New York: John Wiley & Sons, Inc; 1989.

33. Moons KGM, Stijnen T, Es GA van, et al. Treatment thresholds in diagnostic test evaluation: an alternative to the comparison of areas under receiver operating characteristic curves. Med Decis Making 1997;17:447-54.
34. Knottnerus JA, Leffers JP. The influence of referral patterns on the characteristics of diagnostic tests. J Clin Epidemiol 1992;45:1143-1154.
35. Harrell FE, Lee KL, Matchar DB, Reichert TA. Regression models for prognostic prediction: Advantages, problems, and suggested solutions. Cancer Treat Rep 1985;69:1071-7.
36. Steyerberg EW, Eijkemans MJ, Habbema JD. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. J Clin Epidemiol 1999;52(10):935-42.
37. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med 2000;19(8):1059-79.
38. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models. Efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54(8):774-81.
39. Justice AC, Covinsky KE, Berlin JA. Assessing the generalisability of prognostic information. Ann Intern Med 1999;130(6):515-24.
40. Knottnerus JA. Prediction rules: statistical reproducibility and clinical similarity. Med Decis making 1992;12:286-287.
41. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. Jama 2000;284(1):79-84.
42. Halfon N, Newacheck PW, Wood DL, St Peter RF. Routine emergency department use for sick care by children in the United States. Paediatrics 1996;98(1):28-34.
43. Smith RD, McNamara JJ. Why not your paediatrician's office? A study of weekday paediatric emergency department use for minor illness care in a community hospital. Pediatr Emerg Care 1988;4(2):107-11.

Appendix

'Clinical score'.

Variable	Points to assign				Score
Duration of fever at presentation	For each day of fever 1 point (½ days are rounded up), maximum 10 points.				
Age	≤1 year = 0 >1 year = -4				
Body temperature at examination	°C	Points	°C	Points	
	<38*	0	40.0-40.4	5	
	38.0-38.4	1	40.5-40.9	6	
	38.5-38.9	2	41.0-41.4	7	
	39.0-39.4	3	41.5-41.9	8	
	39.5-39.9	4	≥42.0	9	
					Total + 3 =

*If temperature at examination is <38°C, then temperature in history must be ≥38°C.

'Lab score'.

Variable	Points to assign				Score
White blood cell count ($*10^9$/L)	Count		Points		
	<10		0		
	10-19		7		
	20-29		14		
	≥30		21		
	Count	Points	Count	Points	
Serum C-reactive protein (mg/L)	<10	0	80-89	24	
	10-19	3	90-99	27	
	20-29	6	100-109	30	
	30-39	9	110-119	33	
	40-49	12	120-129	36	
	50-59	15	130-139	39	
	60-69	18	140-149	42	
	70-79	21	≥150	45	
					Total =

3.3 External validation and update of a prediction rule for serious bacterial infection in young children referred with fever without apparent source

SE Bleeker, G Derksen-Lubsen, DE Grobbee, ART Donders, KGM Moons, HA Moll

Submitted

Abstract

Objective: to externally validate a previously developed rule for prediction of the presence of serious bacterial infections in patients presenting with fever without apparent source, and to improve the predictive ability of the rule. *Methods:* patients, aged 1 month to 36 months, presenting at the emergency department with fever without source between 2000 and 2001, were prospectively enrolled. Serious bacterial infection included bacterial meningitis, sepsis, bacteraemia, pneumonia, urinary tract infection, bacterial gastroenteritis, osteomyelitis and ethmoiditis. The generalisability of the original rule was determined. Subsequently, the predictive ability of the rule was improved on all available data of the patients with fever without source (1996-1998 and 2000-2001, n=381) using multivariable logistic regression, resulting in a risk stratification. *Results:* the generalisability of the original prediction rule appeared insufficient in the new patients (n=150). In the updated rule, independent predictors from history and examination were duration of fever at presentation, vomiting, ill clinical appearance, chest-wall retractions and poor peripheral circulation (area under the receiver operating characteristic curve (95% confidence interval): ROC area (95%CI): 0.69 (0.63-0.75)). Additional independent predictors from laboratory tests were white blood cell count, serum C-reactive protein and the presence of ≥70 white blood cells in urinalysis (ROC area (95%CI): 0.83 (0.78-0.88). The risk stratification for serious bacterial infection ranged from 4% to 54%. *Conclusions:* in patients with fever without apparent source and at risk for serious bacterial infection information on patient history and physical examination can be used as a first screening tool. Additional laboratory testing may refine the individual risk estimate (range: 4%-54%).

Data analysis

In spite of the prospective data collection, some data were missing (<3% in data from patient history and physical examination; laboratory testing was done at the discretion of the individual treating physician: WBC count and CRP and respectively urinalysis were not performed in 17 and 20 children). Missing values were completed by imputation (17). In this, we mimicked the application of the model in clinical practice, in which information on a particular predictor may be missing (not measured) as well. The expected value, based on the mean of each predictor with missing values, was used for imputation.

To determine the generalisability of the prediction rule in the validation set, first the calibration of the model was evaluated by comparing the predictive probability and the observed proportion of serious bacterial infections for deciles of the patient population (Hosmer & Lemeshow test). Subsequently, the discriminating performances as reflected by the areas under the receiver operating characteristic curves (ROC areas) of the 'clinical model' and the 'clinical + lab model' were estimated in the validation set and compared to the ROC areas found in the derivation set.

II. Optimisation of the prediction model

Data analysis

In order to improve the predictive ability of the prediction model, the derivation set and validation set were merged ('total set'). The potential diagnostic determinants that were evaluated included the same ten predictors as in the previously derived prediction rule. Based on literature and clinical practice, additional relevant variables, on which information had been collected prospectively, were evaluated as well. These included gestational age, quality of cry (<1 year), poor intake, ill clinical appearance judged by the (trainee) paediatrician, state of consciousness, body weight and pale skin (1, 9-12, 18-26). The total number of candidate variables was 17.

To optimise the prediction model, variables with 50% or more missing values were excluded from the analyses. Missing values in the total data set (derivation and validation set) were completed by imputation using S-plus library MICE (version 1.0) (27). This method uses all available data to impute the missing values based on the correlation between each variable with missing values and all other variables (17). Taking into consideration uncertainties in imputed data, multiple imputations were performed by repeating the imputation five times (28).

The association between potential diagnostic determinants and the presence of a serious bacterial infection was assessed using logistic regression analyses (SPSS version 9.0). Continuous variables were analysed both on a linear and a transformed scale, e.g. logarithmic or quadratic, to determine which scale was the better predictor of outcome. First, the predictive ability of the full model, i.e. including all fourteen clinical variables, was estimated using multivariable logistic regression. Subsequently, backward selection from the full model was performed using the likelihood ratio (LR) test in which the model chi-square of each reduced model was compared to that of the full model. If the LR test became significant (p-value of less than 0.10) no further reduction was performed. This yielded the 'reduced clinical model'. Then, all three candidate laboratory variables (WBC blood count, serum CRP and the presence of 70 or more WBC in dipstick urinalysis) were added to this 'reduced clinical model' to estimate their incremental value. Similarly, lab variables with a p-value $\geq$0.10, based on the LR test, were excluded from the model. The final model including the selected predictors from patient history, physical examination and laboratory tests is referred to as the 'reduced clinical + lab model'.

For each model, its calibration and discriminating performance were evaluated as described in the external validation methods section. Subsequently, we used bootstrapping techniques applied to the full model to validate the two reduced models, i.e. to adjust the models for optimism (4). The entire process of stepwise variable selection was also included in the bootstrapping techniques. The bootstrapping technique was executed using the Hmisc and Design libraries in S-plus.

Results

I. External validation of the prediction rule

In total 150 patients with fever without apparent source were included in the validation study. In table 1 the distribution of several patient characteristics, including final diagnoses and the predictors of the diagnostic rule, of both the validation and the derivation set are presented. The two populations were largely comparable, although the patients in the validation set were slightly younger and had less often a history of vomiting. A serious bacterial infection was present in 27% of the patients in the validation set, compared to 25% in the derivation set.

The 'clinical' model and the 'clinical + lab' model both were applied to the validation set to assess their generalisability. Both models were well calibrated

Table 1. General characteristics and final diagnoses of the patients with fever without apparent source in the derivation and validation set*.

	Validation set (n†=150)	Derivation set (n=231)
Male gender	85 (57)	122 (53)
Age (years)	0.9 (0.8)	1.1 (0.8)
Weeks of gestation‡	39.4 (1.5)	39.2 (2.2)
Duration of fever (days)‡	2.8 (2.4)	2.8 (2.4)
History of poor micturition	37 (25)	63 (27)
History of vomiting	43 (29)	95 (41)
Body temperature at physical examination (°C) ‡	39.2 (1.0)	39.4 (1.0)
Chest-wall retractions ± tachypnoea	18 (12)	25 (11)
Poor peripheral circulation	27 (18)	45 (20)
White blood cell count (*10^9/L) ‡	15.4 (8.6)	14.8 (8.0)
Serum C-reactive protein (mg/L) ‡	57.6 (60.7)	48.3 (58.5)
≥70 white blood cells/μl in dipstick urinalysis	34 (23)	70 (30)
Serious bacterial infection present	41 (27)	58 (25)

*Values represent absolute numbers (percentages) unless stated otherwise.

†Number of patients.

‡Mean (standard deviation).

(p-values Hosmer & Lemeshow tests: 0.7 and 0.6, respectively). But, the ROC area (95%CI) of the 'clinical model' was 0.60 (0.49-0.70), which was substantially lower compared to that in the derivation set (0.75 (0.68-0.83)), and discriminated poorly. The ROC area (95%CI) of the 'clinical + lab model' was 0.78 (0.69-0.86), which was closer to the ROC area (95%CI) in the derivation set (0.83 (0.77-0.89)) (3). The poor discriminative ability of the 'clinical model' forced us to rate the rule inadequate to classify future patients presenting with fever without apparent source and thus useless for paediatric care. We believed that better discrimination based on patient history and physical examination could be achieved. The dissatisfying results of the 'clinical model' in particular, prompted us to optimise the prediction of a serious bacterial infection by deriving a new prediction model using all data, i.e. all patients with fever without apparent source ('total set', n=381).

II. Optimisation of the prediction model

The distribution of the seventeen candidate variables in the total set is presented in table 2. Variable selection yielded five strong predictors from patient history and physical examination. These were duration of fever at presentation, history of vomiting, ill clinical appearance judged by the (trainee) paediatrician, chest-wall retractions and poor peripheral circulation (table 3, 'reduced clinical model'). The ROC area (95%CI) of this model after adjustment for overfitting (bootstrapping) was 0.69 (0.63-0.75). After adding the laboratory variables to this 'reduced clinical model',

Table 2. Distribution of candidate variables from patient history, physical examination and laboratory tests in the total set (n*=381).

Characteristics†	SBI‡ absent (n=282)	SBI present (n=99)
Patient history		
Age (years) §	1.0 (0.8)	1.1 (0.8)
Weeks of gestation§	39.3 (1.9)	39.5 (1.6)
Duration of fever (days) §	2.6 (2.3)	3.2 (2.6)
Changed crying pattern (only <1 year)	151 (54)	39 (39)
Poor intake	107 (38)	36 (36)
History of vomiting	87 (31)	49 (50)
History of poor micturition	72 (26)	27 (27)
Physical examination		
Ill clinical appearance	128 (49)	63 (64)
Decreased consciousness	12 (4)	11 (11)
Body temperature at physical examination (°C) §	39.3 (0.9)	39.4 (1.1)
Body weight (kilograms) §	8.6 (3.1)	9.2 (3.1)
Poor peripheral circulation	31 (11)	26 (26)
Pale skin	41 (15)	28 (28)
Chest-wall retractions	19 (7)	24 (24)
Laboratory		
White blood cell count (*10^9/L) §	13.5 (7.3)	19.1 (9.3)
Serum C-reactive protein (mg/L) §	35.7 (45.3)	94.0 (70.0)
≥70 white blood cells/μl in dipstick urinalysis	58 (21)	46 (47)

*Number of patients.

†Values represent absolute numbers (percentages) unless stated otherwise.

‡serious bacterial infection.

§Mean (standard deviation).

WBC count, CRP and the presence of 70 or more WBCs in dipstick urinalysis contained additional diagnostic information (table 3, 'reduced clinical + lab model'). The ROC area (95%CI) of the 'reduced clinical model' increased significantly to 0.86 (0.82-0.90). The goodness-of-fit tests in both models were far from significant (p-values: 0.7 and 0.3, respectively), demonstrating good fit.

Table 3. Results from multivariable analyses on the total set, corrected for optimism: predictors of presence of a serious bacterial infection.

	Reduced clinical model*		**Reduced clinical + lab model†**	
Characteristics	ß§	Odds ratio (90%CI)	ß	Odds ratio (90%CI)
Duration of fever (days)‡	0.79	2.2 (1.2-4.1)	0.12	1.1 (0.5-2.4)
History of vomiting	0.52	1.7 (1.1-2.6)	0.53	1.7 (1.0-2.8)
Ill clinical appearance	0.45	1.6 (1.0-2.4)	0.46	1.6 (0.9-2.7)
Chest-wall retractions ± tachypnoea	1.18	3.3 (1.8-6.0)	1.33	3.8 (1.9-7.7)
Poor peripheral circulation	0.70	2.0 (1.1-3.6)	1.09	3.0 (1.4-6.3)
White blood cell count (*10^9/L)			0.05	1.05 (1.01-1.08)
Serum C-reactive protein (mg/L)			0.01	1.01 (1.01-1.02)
≥70 white blood cells/µl in dipstick urinalysis			1.07	2.9 (1.4-6.3)
ROC area (95%CI)		0.69 (0.63-0.75)		0.86 (0.82-0.90)

*Intercept of the model was –2.04.

†Intercept of the model was –3.74.

‡Odds ratio and 90% confidence interval estimated after logarithmic transformation of the characteristic.

§ß = Regression coefficient.

Two readily applicable diagnostic scoring rules were constructed by transforming the two final reduced models (see appendix). By assigning the corresponding points for each variable present, a 'clinical score' and a 'lab score' were obtained for each individual patient. The 'clinical score' and the 'lab score' ranged from 0-38 and 0-33, respectively; the ROC area (95%CI) of those transformed scores were 0.75 (0.69-0.81) and 0.85 (0.81-0.89), respectively. The absolute numbers of patients across various

Table 4. Risk of a serious bacterial infection across clinical score categories with and without laboratory tests (n*=381).

			Lab score	
			≤8	>8
Clinical score		**Without lab†**	(n=201)	(n=180)
≤10	(n=203)	12%	4%	31%
>10	(n=178)	42%	15%	54%

*Number of patients.

†Without lab means using the 'clinical score' only.

categories of the estimated risk scores from the diagnostic rules were estimated (table 4). Using patient history and physical examination only, the *a priori* risk of having a serious bacterial infection of 26% (i.e. the prevalence: 99/381) could be changed to 12% or 42% ('clinical score'). For instance, in the case of a vomiting patient, with fever since a few hours and an ill clinical appearance judged by the physician, but no further signs or symptoms, the 'clinical score' is 9 (0+5+4+0+0=9; see appendix). This 'clinical score' corresponds to a 12% risk of contracting a serious bacterial infection (table 4). Additional laboratory tests, if requested, changes this probability to a 4% or a 31% risk, depending on the results. For example, a WBC count of 15, a CRP level of 40 and a negative (<70 WBCs) urine dipstick analysis, would result in a 'lab score' of 6 (2+4+0; see appendix), changing the patient's risk from 12 % to 4% (table 4).

Discussion

Commonly, the discriminating performance of a prediction rule degrades from the patient data on which the rule has been constructed to new patient data (4, 8, 29-31). Therefore, we have externally validated a previously constructed diagnostic prediction rule, to assess its generalisability (3). The purpose of the rule is to guide in sifting out the children with and those without a severe bacterial infection in patients who initially present with fever without apparent source. The prediction rule is based on two subsequent models: the 'clinical model' and the 'clinical + lab model'. The 'clinical + lab' model proves similar discriminating performance in the new patients as compared to the derivation set. Unfortunately, the 'clinical model' shows substantial

loss of overall discriminating performance in the validation set. We have concluded that the previously developed 'clinical model' is not satisfactory to classify future patients presenting with fever without apparent source and that the best calibrated and best discriminating model was not defined yet.

We have further optimised the previous prediction model for a serious bacterial infection by performing a new multivariable logistic regression analysis on all available data in which we have merged the derivation and validation set. This increased the number of patients with 39% as compared to the derivation set. The analyses show that the most important predictors are: duration of fever at the moment of presentation, vomiting, ill clinical appearance, chest-wall retractions, poor peripheral circulation ('reduced clinical model'), and additionally WBC count, CRP and the presence of ≥70 WBCs in dipstick urinalysis ('reduced clinical + lab model'). Both models have a reasonable to good discriminating performance. Afflicted and non-afflicted patients cannot be discriminated perfectly by the scoring rules. Four patients of the 110 (4%) patients with a 'clinical score' ≤10 and a 'lab score' ≤8 after all suffer from a serious bacterial infection. In three of those cases, urinary tract infections were diagnosed (ages 2, 3 and 10 months), the fourth case concerned a bacterial gastroenteritis (age 3 months). This once again illustrates that the management of children with fever without apparent source and selection of patients with a serious bacterial infection continues to be a delicate task, particularly among 'low risk' patients (i.e. 'clinical score' ≤10) (3). Nonetheless, in spite of the imperfect separation by the prediction rule, its resulting post-test probabilities may affect the management and support the clinicians (clinical validity) (30, 32, 33). The results of the risk stratification suggest that on individual levels, the information on patient history and physical examination can be used as a first screening tool. Additional laboratory testing refines the individual risks. We emphasise to use the diagnostic prediction rule to support decision making in which the clinical experience and knowledge of the clinician should be involved, rather than an independent tool.

Predictors of serious infections in children with fever have been debated extensively in literature. Inherently, our results are partly in agreement and partly in disagreement (1, 9, 12, 18-21, 23-26, 34-37). All predictors identified in this study have been discussed in other published works, sometimes confirming our finding, sometimes not.

Except for 'ill clinical appearance', all of the predictors of the 'optimised rule' overlap with those in the previously developed prediction rule. The recurrence of the

predictors suggests that they are robust, although the magnitude of the associations (regression coefficients) has changed in the overall analyses as compared to the analyses based on the derivation set only (3). The predictors that are included in the initial rule but excluded in the new rule (poor micturition, age above 1 year and body temperature) might indicate that these are unstable predictors and that the initial model has been overfitted. This may be due to limited power and to the use of stepwise selection methods (38-41). By merging the derivation set and validation set the power of the study increased. In this total set, only a limited set of candidate variables has been predefined based on the literature, resulting in a better events-per-variable-ratio (4, 39, 40).

This validation study has certain limitations. The original validation set is a relative small sample (n=150). To our knowledge, it is yet unknown how large (i.e. number of events) a validation study should be (41). With hindsight, it would have been better to derive and validate the prediction rule on a larger sample. The unsatisfying validation results of the 'clinical model' have urged us to update the previously defined prediction rule (42). To improve precision, we have merged the derivation and validation sets. This undermines the principle of an external validation study, in which the performance of a model should be tested on independent data. Although we have adjusted our newly defined rule for optimism using bootstrapping techniques, before applying the updated rule in clinical practice, its generalisability has to be evaluated again in another external validation procedure. Third, inherent to the non-specific presentation in young children, the outcome in children presenting with fever without apparent source is diverse. The endpoint serious bacterial infection is a cluster of several final diagnoses (9-12). When a child presents at the ED with fever without apparent source, firstly the risk for development of any serious bacterial infection has to be estimated. The assessment of a specified diagnosis might follow in a later stage. Such heterogeneity in the outcome makes both the clinical approach and the design of prediction research more complicated, compared to a more straightforward outcome, such as bacterial meningitis (14). It makes it difficult to make a complete distinction between patients with and without the outcome, as shown by our findings.

To conclude, the management of young children with fever without apparent source remains associated with uncertainties. Risks can not be reduced to zero. In this validation study, a previously obtained prediction rule was updated. This prediction rule enables to classify patients with fever without apparent source in different risk

categories of having a serious bacterial infection. Patients can be screened based on information on history and physical examination (duration of fever, vomiting, ill clinical appearance, chest wall retractions, and poor peripheral circulation). If necessary, subsequent laboratory testing (WBC count, CRP, urinalysis) may refine this risk (range from 4% to 54%). The diagnostic prediction rule can be used as a support in decision making, in combination with the clinical expertise of the clinician.

Acknowledgements

We express our gratitude to all (trainee) paediatricians and nurses of the emergency departments of the participating hospitals, and to R. Oostenbrink, MD, PhD, and Wilfried de Jong, Femke Mineur and Haley C. Kang, medical students, for support in data collection, and to M.J. Twijnstra, MSc for data management. The Health Care Insurance Council of the Netherlands financially supported this project.

References

1. Baraff LJ, Lee SI. Fever without source: management of children 3 to 36 months of age. Pediatr Infect Dis J 1992;11:146-51.
2. Isaacman DJ, Kaminer K, Veligeti H, Jones M, Davis P, Mason JD. Comparative practice patterns of emergency medicine physicians and paediatric emergency medicine physicians managing fever in young children. Paediatrics 2001;108:354-8.
3. Bleeker SE, Moons KGM, Derksen-Lubsen G, Grobbee DE, Moll HA. Predicting serious bacterial infection in young children with fever without apparent source. Acta Paediatrica 2001;90:1226-1232.
4. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine 1996;15:361-87.
5. Justice AC, Covinsky KE, Berlin JA. Assessing the generalisability of prognostic information. Ann Intern Med 1999;130:515-24.
6. Laupacis A, Wells G, Richardson WS, Tugwell P. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. Jama 1994;272:234-7.
7. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. Jama 2000;284:79-84.

8. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. New England Journal of Medicine 1985;313:793-9.
9. Baraff LJ, Bass JW, Fleisher GR, et al. Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Paediatrics 1993;92:1-12.
10. Baraff LJ, Oslund SA, Schriger DL, Stephen ML. Probability of bacterial infections in febrile infants less than three months of age: a meta-analysis. Pediatr Infect Dis J 1992;11:257-64.
11. Baker MD, Bell LM, Avner JR. Outpatient management without antibiotics of fever in selected infants. N Engl J Med 1993;329:1437-41.
12. Dagan R, Powell K, Hall C, Menegus M. Identification of infants unlikely to have serious bacterial infection although hospitalised for suspected sepsis. J Pediatr 1985;107:855-860.
13. Oostenbrink R, Moons KGM, Theunissen CC, Derksen-Lubsen G, Grobbee DE, Moll HA. Signs of meningeal irritation at the emergency department: how often bacterial meningitis? Pediatr Emerg Care 2001;17:161-4.
14. Oostenbrink R, Moons KGM, Donders ART, Grobbee DE, Moll HA. Prediction of bacterial meningitis in children with meningeal signs: reduction of lumbar punctures. Acta Paediatr 2001;90:611-7.
15. Oostenbrink R, Oostenbrink JB, Moons KGM, et al. Cost-utility analysis of patient care in children with meningeal signs. Int J Technol Assess 2002;18(3):485-96.
16. Oostenbrink R, Maas M, Derksen-Lubsen G, Moll HA. Sequelae after bacterial meningitis in childhood. Scan J Inf Dis 2002;34:379-82.
17. Little RA. Regression with missing X's: A review. J Am Stat Assoc 1992;87:1227-1237.
18. McCarthy PL. Fever. Pediatr in Rev 1998;19:401-8.
19. Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteraemia in young febrile children. Ann Emerg Med 1998;31:679-87.
20. McCarthy PL, Lembo RM, Baron MA, Fink HD, Cicchetti DV. Predictive value of abnormal physical examination findings in ill-appearing and well-appearing febrile children. Paediatrics 1985;76:167-71.
21. McCarthy PL, Lembo RM, Fink HD, Baron MA, Cicchetti DV. Observation, history, and physical examination in diagnosis of serious illnesses in febrile children less than or equal to 24 months. J Pediatr 1987;110:26-30.
22. Kramer MS, Lane DA, Mills EL. Should blood cultures be obtained in the evaluation of young febrile children without evident focus of bacterial infection? A decision analysis of diagnostic management strategies. Paediatrics 1989;84:18-27.
23. Isaacman DJ, Shults J, Gross TK, Davis PH, Harper M. Predictors of bacteraemia in febrile children 3 to 36 months of age. Paediatrics 2000;106:977-82.
24. Hewson PH, Humphries SM, Roberton DM, McNamara JM, Robinson MJ. Markers of serious illness in infants under 6 months old presenting to a children's hospital. Arch Dis Child 1990;65:750-6.

25. Berger RM, Berger MY, Steensel-Moll HA van, Dzoljic-Danilovic G, Derksen-Lubsen G. A predictive model to estimate the risk of serious bacterial infections in febrile infants. Eur J Pediatr 1996;155:468-73.
26. Teach SJ, Fleisher GR. Duration of fever and its relationship to bacteraemia in febrile outpatients three to 36 months old. The Occult Bacteraemia Study Group. Pediatr Emerg Care 1997;13:317-9.
27. Buuren S van, Oudshoorn CGM. Multiple imputation online [TNO Prevention and Health: Multiple imputation web site]. Available at: http://www.multiple-imputation.com. Updated: November 20, 2001.
28. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. Am J Epidemiol 1995;142:1255-64.
29. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Ass 1983;78:316-31.
30. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. Jama 1997;277:488-94.
31. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19:453-73.
32. 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 1994;271:389-91.
33. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. Jama 1994;271:703-7.
34. Jaffe DM, Fleisher GR. Temperature and total white blood cell count as indicators of bacteraemia. Paediatrics 1991;87:670-4.
35. Kramer MS, Tange SM, Mills EL, Ciampi A, Bernstein ML, Drummond KN. Role of the complete blood count in detecting occult focal bacterial infection in the young febrile child. J Clin Epidemiol 1993;46:349-57.
36. Lee GM, Harper MB. Risk of bacteraemia for febrile young children in the post-*Haemophilus influenzae* type b era. Arch Pediatr Adolesc Med 1998;152:624-8.
37. Jaye DL, Waites KB. Clinical applications of C-reactive protein in paediatrics. Pediatr Infect Dis J 1997;16:735-46; quiz 746-7.
38. Harrell FE, Lee KL, Matchar DB, Reichert TA. Regression models for prognostic prediction: Advantages, problems, and suggested solutions. Cancer Treat Rep 1985;69:1071-7.
39. Steyerberg EW, Eijkemans MJ, Habbema JD. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. J Clin Epidemiol 1999;52:935-42.

40. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med 2000;19:1059-79.
41. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models. Efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54:774-81.
42. Houwelingen HC van, Thorogood J. Construction, validation and updating of a prognostic model for kidney graft survival. Stat Med 1995;14:1999-2008.

Appendix

'Clinical score'.

Variable	Points to assign		Score
Duration of fever	Days	Points	
	½	0	
	1	2	
	1½	4	
	2-2½	5	
	3-3½	6	
	4-4½	7	
	5-6	8	
	6½-8½	9	
	≥9	10	
History of vomiting	no = 0/ yes = 5		
Ill clinical appearance	no = 0/ yes = 4		
Chest-wall retractions ± tachypnoea	no = 0/ yes = 12		
Poor peripheral circulation	no = 0/ yes = 7		
			Total =

'Lab score'.

Variable	Points to assign		Score
White blood cell count	Count	Points	
(*10^9/L)	<10	0	
	10-19	2	
	20-29	4	
	30-39	6	
	≥40	8	
Serum C-reactive protein	0-99:	1st integer,	(max 16)
(mg/L)		e.g. CRP=*72*: *7* points	
	≥100	1st and 2nd integers, e.g. CRP=*143*: *14* points, maximum=16 points	
≥70 white blood cells/μl in dipstick urinalysis	no = 0/ yes = 9		
			Total =

Chapter 4

Methods and perspectives

4.1 Diagnostic research on routine care data: prospects and problems

R Oostenbrink, KGM Moons, SE Bleeker, HA Moll, DE Grobbee
In press, Journal of Clinical Epidemiology

Abstract

Diagnostic practice is a sequential process starting with a patient with a particular set of signs and symptoms. To serve practice diagnostic research should aim to quantify the added value of a test to clinical information that is commonly available before the test would be applied. Routine care databases commonly include all documented patient information, and seem therefore be suitable to quantify a test's added value to prior information. It is well known, however, that retrospective use of routine care data in diagnostic research may cause various methodological problems. But, given the increased attention of computer-based patient records including data from routine patient care, we believe it is time to reconsider these problems. We discuss four problems related to routine care databases. First, most databases do not label patients by their symptoms or signs but by their final diagnosis. Second, in routine care the diagnostic work-up of a patient is by definition determined by previous diagnostic (test) results. Therefore, routinely documented data are subject to so-called work-up bias. Third, the reference test is in practice always interpreted with knowledge of the preceding test information, such that in scientific studies the diagnostic value of a test under evaluation is commonly overestimated. Fourth, routinely documented databases are likely to contain missing data. Per problem we discuss methods, which are presently available and may (partly) overcome each problem. All this could contribute to more frequent and appropriate use of routine care data in diagnostic research. The discussed methods to overcome the above problems may as well be useful to prospective diagnostic studies.

Introduction

Diagnostic practice is a sequential, stepwise process starting with a patient with a particular set of signs and symptoms. In order to ascertain or rule out a diagnosis, the physician decides upon additional tests based on his findings in previous steps, to increase or decrease the probability of a particular disease (target disease). Hence, to serve practice, diagnostic research should select patients conform practice, follow the sequential process of making a diagnosis in practice and should aim to quantify the added value of a test to clinical information that available before the test would be applied. Although this has been recognised for years (1-15), at present the majority of diagnostic studies still include single test studies, aiming to estimate a test's sensitivity and specificity, without considering other (previous) patient information and quantifying the test's added value (15-17). Routine care databases or computer-based patient records commonly include all information that is related to patient care. Hence, they include all patient information that is considered relevant to ascertain a diagnosis in routine practice. This makes routinely documented data very well suitable for quantifying the value of a particular test additional to other (previous) information. It has widely been discussed and illustrated (e.g. in studies on the value of exercise stress testing in diagnosis of coronary artery disease), that studies using retrospective patient data as obtained from routine care, provide invalid results. This particularly includes the problem of selection (or also called referral, work-up or verification) bias (2-4, 6, 12, 16, 18-30). However, the use of large electronic databases or computer-based patient records in medical practice is still increasing, notably in general practice (31, 32). We therefore believe it is time to reconsider the methodological disadvantages of diagnostic research based on routine care data and potential solutions to overcome them.

In this paper, we first put forward the nature of diagnosis in practice and the preferred design for quantification of (added) value of diagnostic tests. Subsequently, we summarise the well-known problems that can be encountered in diagnostic research using routine care data. Per problem, we discuss the currently available and sometimes novel methods that may (partly) solve the problem. We believe that these methods could contribute to more frequent and appropriate use of routine care data in diagnostic research. For illustration purposes, we will use throughout an example study that aimed to quantify the (added) value of various diagnostic tests in children suspected of having bacterial meningitis (33, 34).

Example study

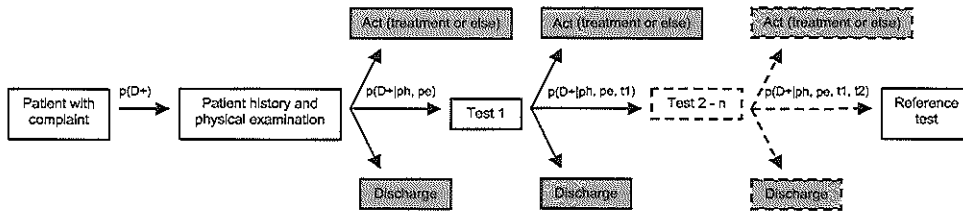
At the emergency department of the Sophia Children's Hospital, Rotterdam, The Netherlands, we performed a study on children visiting the emergency department because of meningeal signs (33, 34). These children pose a diagnostic dilemma for the physician, because they are at risk of bacterial meningitis (target disease (22)), though they may have self-limiting diseases in 50% to 60% as well (34-36). The question is in which of these children a lumbar puncture should be performed and empirical antibiotic treatment should be started such that not a single case of bacterial meningitis will be missed. We performed a study aiming to quantify which diagnostic tests, including findings from patient history, physical examination, and additional tests, have independent (added) value in the discrimination between presence or absence of bacterial meningitis in children presenting with meningeal signs at the emergency department. We developed and validated a simple prediction rule for bacterial meningitis in this patient domain. For further details, we refer to previous publications (33, 34).

Diagnostic practice and diagnostic research

In practice, a diagnosis starts with a patient with a clinical problem (symptoms or signs) suspected of having a particular disease, the so-called target disease (2, 5, 8, 15, 22). As shown in figure 1, the physician commonly applies a phased work-up, starting with patient history and physical examination. Subsequent steps in general may be additional laboratory tests, imaging and finally so-called gold standard or reference tests such as arthroscopy, angiography or, as in our example, lumbar puncture. After each phase the physician will consider the available diagnostic information to implicitly estimate the diagnostic probability of the presence of the target disease ($p(D+|ph, pe)$ or $p(D+|ph, pe, t1)$ in figure 1). In patients with a very low probability, the physician may refrain from further testing for that disease and, if necessary, will search (or test) for the presence of alternative diseases or may even discharge the patient. In patients with a high target disease probability treatment will be initiated. As long as uncertainty remains, further diagnostic tests are applied until a treatment decision (including no treatment) can safely be made. To set a diagnosis is to know that more information would not change the decision to act as if the patient has the disease. To rule out a disease is to know that more information would not change the decision to act as if the patient did not have the disease (37, 38).

The motive of diagnostic research is efficiency: decreasing the patient burden and measurement costs of the diagnostic work-up in practice considering the consequences of false (missed) diagnoses. To serve diagnostic practice, diagnostic research should reflect the sequential process of probability estimation in practice. In this view, diagnostic research is a typical form of pragmatic, prediction research. Hence, in diagnostic research, study patients should also be selected on their suspicion of having a particular disease and the aim should be to quantify which tests (truly) contribute to the estimation of probability of presence (absence) of that disease (2, 7-9, 15, 22, 39, 40). Accordingly, every test (including findings from history and physical examination, which are diagnostic tests as well) that may be relevant to or used in the probability estimation, should be considered in the study. For this purpose, databases including routinely documented data may be very useful, since in principle they comprise all patient data that may be relevant to ascertain a diagnosis in practice. Furthermore, as routinely documented data include the entire diagnostic and therapeutic process, they allow for stepwise analysis of tests in the sequence as they occurred in routine practice and allow to estimate the added value of tests to previous findings. However, there are some problems related to the use of retrospective or routine care data in scientific diagnostic research. These, as well as their possible solutions are discussed below.

Figure 1. Illustration of a typical diagnostic process in clinical practice.
ph: patient history; pe: physical examination; t1: result of test 1; t2 - n: result of test 2 up to test n; Dashed elements not necessarily occur in all patients.
p(D+): probability of disease; p(D+ | ph, pe): probability of disease given the information from ph and pe.



Problems of using routine care data in diagnostic research and methods to solve them

Selection of the proper patient population

In scientific diagnostic studies or test evaluations, patients are often selected on the established presence or absence of a particular disease, i.e. patients that had undergone the so-called gold standard or reference test in routine care are selected for the study. Similarly, studies aiming to quantify the diagnostic accuracy of a particular test often include only those patients that had undergone that test in routine care. Both types of patient selection for diagnostic studies, however, leads to the inclusion of only selected patient groups because in routine care physicians selectively refer patients for additional tests (including the reference test, figure 1) (3, 4, 6, 9, 12, 18, 23, 24, 27, 40). In particular, patients with a low probability of the target disease based on previous information are not referred for the usual more burdening and costly tests. But, as discussed above, to enhance applicability of study results to clinical practice, in diagnostic studies patients should be selected on the problem, i.e. symptoms or signs, with which they visit the physician (5, 6, 9, 10, 12, 15, 39-42). In our example study children with meningeal signs, suspected of having bacterial meningitis were included. This selection includes the complete spectrum of differential diagnoses of patients with meningeal signs conform practice (34). Of course, this population does not reflect all patients with meningitis, since meningitis might be present without meningeal signs (43, 44).

The necessary selection of patients on presenting problem (i.e. symptoms or signs) can be done using existing or routine care databases. However, it requires a problem oriented patient classification database in which patients are classified to their presenting problem. Using such classification system, for example, all children referred with 'meningeal signs' to the outpatient or emergency department as the main reason for encounter, can easily be selected (45, 46). In contrast to general practice research, where symptom and problem oriented databases (ICPC (47)) have already been considered of key importance (31, 32), most routine care patient classification systems in hospitals are based on the final diagnosis only. But problem oriented coding in routine care databases of hospitals may be recommended to those interested in diagnostic research.

Verification bias

Although patients may be properly selected for a scientific study on their presenting symptoms or signs by using routine care data with a problem oriented classification system, there may still be the problem of so-called verification, referral or work-up bias when using routine care data for scientific diagnostic studies (1-4, 6, 12, 16, 18-30, 48). For scientific purposes, ideally, all patients undergo the same and entire diagnostic work-up including the reference test to determine the final diagnosis. In practice, however, as said, not all patients (routinely) undergo the reference test (figure 1), but commonly a selected sample only. Due to this selective verification, work-up or referral, it has extensively been shown that use of routine care data for research purposes leads to biased estimates of the accuracy of the test(s) under study (1-4, 6, 12, 16, 18-30, 48).

However, we believe that this problem with routine care data may be overcome if the target disease is fatal or rapidly progressing without treatment. In such instances, it seems reasonable to consider, in absence of a reference test, the target disease as absent if the patient recovers without treatment (good clinical course over time), and as present if such patient shortly returns with evident symptoms of the target disease. Hence, provided that information from patient follow-up is indeed available (which is common), such a pragmatic approach using data from patient follow-up may be valid in studies based on routine care data. The problem of verification bias when using routine care data particularly applies if the (target) disease has a subclinical or self-limiting course such that the actual presence of the disease under study may be missed. In our example study, the outcome diagnosis was the presence or absence of bacterial meningitis, which is a rapidly progressing disease without treatment (49). In patients who did not routinely undergo a lumbar puncture, bacterial meningitis was ruled out based on an uneventful clinical course without treatment.

The pragmatic approach of using routine care data from clinical follow-up to determine the final diagnosis in each patient is not uncommon in diagnostic studies where it is unethical to apply the reference test to all patients suspected of having the disease. Examples are studies on diagnosis of pulmonary embolism and breast cancer (50, 51). Finally, it must be noted that some (target) diseases lack a single reference standard, such as heart failure. Also, in these instances, a more pragmatic approach of consensus diagnosis using data from patient follow-up is used for research purposes (52, 53).

Blinding

Another problem of using routine care data is the problem of absence of blinding. For research purposes, the reference test is interpreted preferably without knowledge of (or blinded for) preceding test results, notably of the test(s) under study which accuracy is to be quantified. If this blinding is not guaranteed, the information provided by the test(s) under study may partly be used and included ('incorporated') by the observer of the reference test. Consequently the estimated accuracy of the test(s) under study will be biased, so-called 'incorporation bias', 'test review bias', or 'diagnostic review bias' (1, 2, 9, 16, 22, 40, 54, 55). Theoretically this bias may result in an under- or overestimation of the test's accuracy. Though, it often results in an overestimation since the results of the test under evaluation and the results of the reference test become more alike (matched), incorrectly decreasing the number of false positive and false negative test results (55).

In routine diagnostic practice, the result of each test (including the reference test) is always considered with knowledge of preceding test information. Hence, in scientific studies using routinely documented data blinding is commonly not guaranteed. We believe, however, that this does not pose major problems for reference tests yielding merely objective results, i.e. not requiring subjective interpretation. In our example study, the reference test included an increased cell count in cerebrospinal fluid (CSF) and a positive culture from CSF of blood, which is rather objective and not influenced by subjective interpretation. Incorporation bias particularly occurs when the interpretation of the reference test is subject to intra- and inter-observer variation, as applies to most imaging tests.

Missing data and analysis

Finally, missing data are more likely in routinely documented data than in data documented for a specific prospective study. Standard data analytical techniques exclude subjects with a missing value on a variable (56, 57). Consequently, conventional analysis of diagnostic research data includes only the cases without missing data (so-called complete cases). In case of a high frequency of missing data or if many variables are considered, this may grossly reduce the number of analysable patients and therefore the power of the study. Furthermore, complete case analysis may yield biased results, since the cases excluded because of missing data often differ systematically from the complete cases (56-59). A patient with very severe symptoms and signs, for example a patient with meningeal signs and convulsions, may already be

referred for more sophisticated diagnostic tests (e.g. lumbar puncture) before full completion of patient history and physical examination. On the other hand, as illustrated in figure 1, in a patient presenting with very mild or no symptoms, additional test information may be incomplete, as the physician already ruled out a serious disease early in the diagnostic process and did not consider additional tests to be necessary (37). Hence, missing data in routine care tend not to occur at random, but for a reason ('on indication').

Imputation techniques may serve to overcome the bias due to missing data in diagnostic research using routine patient care data and to increase the efficiency of the analysis. Frequently applied methods, e.g. imputation of mean values of available data or the so-called 'indicator method', however, may also lead to biased estimates (57-59). More recent methods, i.e. the maximum-likelihood method and multiple imputation, have been proposed which seem to yield more valid estimates. These imputation methods use all information available from the actually documented data to impute the missing values (57-59). By multiple imputation each missing data is imputed more than once, e.g. using a Bayesian algorithm, such that multiple databases with complete data are obtained. Repeating the analyses, commonly five times, on each of the imputed data sets and averaging the results according to standard statistical techniques (57) will result in estimates of the regression coefficients and their standard errors, incorporating the uncertainty of the imputed values (53, 55). The multiple imputation method is an accepted and valid statistical method for the problem of missing data, as long as the probability of 'missingness' depends on observed variable values only. Whether missingness depends on unobserved variables, however, can obviously not be assessed from the observed data but can only be judged (57-59). Therefore, also analyses based on (multiple) imputed data should still be interpreted with care although it must be noted that there are no straightforward methods to handle missing data that depend on unobserved data. The number of repeated imputations depends on the proportion of missing data (per variable). The allowable proportion of missing data to validly apply (multiple) imputation techniques is yet unknown. We believe, however, when using routine care data, variables with a very high proportion of missing data (e.g. >50%) should be excluded from the analysis anyway, since they are presumably difficult to obtain in practice or frequently unavailable or not considered clinically important.

Given a complete data set after proper (multiple) imputation of missing data, routinely documented data including the entire stepwise diagnostic work-up and the

clinical course facilitate proper analysis of diagnostic tests in the sequence as they commonly occur in routine practice. The (overall) diagnostic probability is the central source of information, instead of the (single) test results and corresponding sensitivity and specificity (4, 7, 8, 10, 14, 15, 38, 39). Multivariable logistic regression modeling allows for probability estimation of the presence of a disease as a combined function of available characteristics from patient history, examination and additional laboratory tests (7, 14, 15, 38, 39, 41, 56, 60). This probability estimation based on all diagnostic information together directly coheres with the diagnostic process in practice. The discriminative value of a diagnostic model can be expressed by the area under the Receiver Operator Characteristic curve (ROC area) (61, 62). From such a logistic model a prediction or decision rule can be derived, that is easy applicable in practice (7, 38, 41, 60). Of course, before introducing a diagnostic rule in clinical practice, prospective application of the rule in a new group of similar patients is necessary to get insight in the rule's usefulness and its impact on the diagnostic process in practice (5, 16, 38, 56, 60). For specific details on methods of analysis we refer to the (statistical) literature (7, 38, 39, 56, 61-67).

Conclusions

It should be noted that the described problems encountered in diagnostic research using routine care data and the methods to handle them are not limited to retrospective diagnostic studies only. They may as well apply to diagnostic research based on prospectively documented data. As diagnostic research is a type of pragmatic, prediction research, it should always reflect the sequential diagnostic work-up of medical practice to allow for clinically meaningful inferences. It should start with patients selected on their disease suspicion, who underwent all clinically relevant steps in the diagnostic work-up. Preferably, the reference test is interpreted independently from the preceding information or its result does not require subjective interpretation. Commonly, the added value of tests is always of interest, with reference to preceding information documented anyway. Hence, after some adjustments, such as imputation of missing data if necessary, diagnostic research may be based on routinely documented data including the entire diagnostic, therapeutic and prognostic phase as recorded in practice.

References

1. Ransohoff DF, Feinstein AR. Problems of spectrum and bias in evaluating the efficacy of diagnostic tests. N Engl J Med 1978;299:926-930.
2. Feinstein AR. Clinical Epidemiology: the architecture of clinical research. Philadelphia: WB Saunders Company; 1985.
3. Diamond GA. Selection bias and the evaluation of diagnostic tests: a metadissent. J Chron Dis 1986;39:359-360.
4. Begg CB. Biases in the assessment of diagnostic tests. Stat Med 1987;6:411-423.
5. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Making 1991;11:88-94.
6. Knottnerus JA, Leffers JP. The influence of referral patterns on the characteristics of diagnostic tests. J Clin Epidemiol 1992;45:1143-1154.
7. Knottnerus JA. Application of logistic regression to the analysis of diagnostic data: exact modeling of a probability tree of multiple binary variables. Medical Decision Making 1992;12:93-108.
8. Sackett DL. A primer on the precision and accuracy of the clinical examination. JAMA 1992;267:2638-2644.
9. Jaeschke R, Guyatt G, Sackett DL. Users'guides to the medical literature. III: How to use an article about a diagnostic tests. A. Are the results ofthe study valid? JAMA 1994;271:389-391.
10. Grobbee DE, Miettinen OS. Clinical epidemiology introduction to the discipline. Neth J Med 1995;47:2-5.
11. Mackenzie R, Dixon AK. Measuring the effects of imaging: an evaluative framework. Clin Radiol 1995;50:513-518.
12. Schouw YT van der, Dijk R van, Verbeek ALM. Problems in selecting the adequate patient population from existing data files for assessment studies of new diagnostic tests. J Clin Epidemiol 1995;48:417-422.
13. Dalla-Palma L, Dixon AK, Durand-Zaleski I, Reiser M, Soimakallio S. An overview of cost-effective radiology. Eur Radiol 1997;7:147-150.
14. Moons KGM, Es GA van, Deckers JW, Habbema JDF, Grobbee DE. Limitations of sensitivity, specificity, likelihood ratio, and Bayes' theorem in assessing diagnostic probabilities: a clinical example. Epidemiology 1997;8:12-17.
15. Moons KGM, Es GA van, Michel BW, Büller HR, Habbema JDF, Grobbee DE. Reduncancy of single diagnostic test evaluation. Epidemiology 1999;10:276-281.
16. Reid MC, Lachs MS, Feinstein AR. Use of methodological standards in diagnostic test research. Getting better but still not good. JAMA 1995;274:645-651.
17. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. Jama 1999;282:1061-1066.

18. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. Biometrics 1983;39:206-215.
19. Greenes RA, Begg CB, Cain KC, Swets JA, Feehrer CE, McNeil BJ. Patient-oriented performance measures of diagnostic tests. 2. Assignment potential and assignment strength. Med Decis Making 1984;4:17-31.
20. Greenes RA, Cain KC, Begg CB. Patient-oriented performance measures of diagnostic tests. 1. Tools for prospective evaluation of test order decisions. Med Decis Making 1984;4:7-15.
21. Hlatky MA, Pryor DB, Harrell FE, Jr., Califf RM, Mark DB, Rosati RA. Factors affecting sensitivity and specificity of exercise electrocardiography. Multivariable analysis. Am J Med 1984;77:64-71.
22. Sackett DL, Haynes RB, Tugwell P. Clinical epidemiology; a basic science for clinical medicine. Boston: Little, Brown & Co; 1985.
23. Begg CB. Statistical methods in medical diagnosis. [Review]. Crit Rev Med Inform 1986;1:1-22.
24. Diamond GA, Rozanski A, Forrester JS, et al. A model for assessing the sensitivity and specificity of tests subject to selection bias. Application to exercise radionuclide ventriculography for diagnosis of coronary artery disease. J Chron Dis 1986;39:343-355.
25. Fletcher RH. Carcinoembryonic antigen. [Review]. Ann Intern Med 1986;104:66-73.
26. Knottnerus JA. The effect of disease verification and referral on the relationship between symptoms and diseases. Med Decis Making 1987;7:139-48.
27. Panzer RJ, Suchman AL, Griner PF. Workup bias in prediction research. Med Decis Making 1987;7:115-119.
28. Detrano R, Janosi A, Lyons KP, Marcondes G, Abbassi N, Froelicher VF. Factors affecting sensitivity and specificity of a diagnostic test: the exercise thallium scintigram. Am J Med 1988;84:699-710.
29. Coughlin SS, Trock B, Criqui MH, Pickle LW, Browner D, Tefft MC. The logistic modelling of sensitivity, specificity and predictive value of a diagnostic test. J Clin Epidemiol 1992;45:9-13.
30. Diamond GA. What is the effect of sampling error on ROC analysis in the face of verification bias? [letter; comment]. Med Decis Making 1992;12:155-156.
31. Metsemakers JF, Hoppener P, Knottnerus JA, Kocken RJ, Limonard CB. Computerized health information in the Netherlands: a registration network of family practices. Br J Gen Pract 1992;42:102-106.
32. Dombal FT de. Transporting databanks of medical information from one location to another. Eff Health Care 1983;1:155-163.
33. Oostenbrink R, Moons KGM, Donders ART, Grobbee DE, Moll HA. Prediction of bacterial meningitis in children with meningeal signs: reduction of lumbar punctures. Acta Paediatr 2001;90:611-617.

34. Oostenbrink R, Theunissen CCW, Moons KGM, Derksen-Lubsen G, Grobbee DE, Moll HA. Signs of meningeal irritation at the emergency department; how often bacterial meningitis? Ped Emerg Care 2001;17:161-164.
35. Barnett ED, Bauchner H, Teele DW, Klein JO. Serious bacterial infections in febrile infants and children selected for lumbar puncture. Pediatr Infect Dis J 1994;13:950-953.
36. Levy M, Wong E, Fried D. Diseases that mimic meningitis. Analysis of 650 lumbar punctures. Clin Pediatr 1990;29:254-255, 258-261.
37. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. N Engl J Med 1980;302:1109-1117.
38. Sox HC, Jr. Probability theory in the use of diagnostic tests. An introduction to critical study of the literature. Ann Int Med 1986;104:60-66.
39. Miettinen OS, Henschke CI, Yankelevitz DF. Evaluation of diagnostic imaging tests: diagnostic probability estimation. J Clin Epidemiol 1998;51:1293-1298.
40. Begg CB, McNeil BJ. Assessment of radiologic tests: control of bias and other design considerations. Radiology 1988;167:565-569.
41. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. N Engl J Med 1985;313:793-799.
42. Braitman LE, Davidoff F. Predicting clinical states in individual patients. Annals of Internal Medicine 1996;125:406-412.
43. Riordan FA, Thomson AP, Sills JA, Hart CA. Bacterial meningitis in the first three months of life. Postgrad Med J 1995;71:36-38.
44. Valmari P, Peltola H, Ruuskanen O, Korvenranta H. Childhood bacterial meningitis: initial symptoms and signs related to age, and reasons for consulting a physician. Eur J Pediatr 1987;146:515-518.
45. Steensel-Moll HA van, Jongkind CJ, Aarsen RSR, et al. [A problem oriented patient classification system for general paediatrics II]. Tijdschr Kindergeneesk 1996;64:99-104.
46. Derksen-Lubsen G, Jongkind CJ, Kraayenoord S, et al. [A problem oriented patient classification system for general paediatrics I]. Tijdschr Kindergeneesk 1996;64:93-98.
47. Lamberts H, Wood M, Hofmans-Okkes I, editors. The international Classification for Primary Care in the European Community; with a multip-language layer. Oxford: Oxford University Press; 1993.
48. Bates AS, Margolis PA, Evans AT. Verification bias in paediatric studies evaluating diagnostic tests. J Paediatrics 1993;122:585-590.
49. Feigin RD, McCracken GH, Jr., Klein JO. Diagnosis and management of meningitis. Pediatr Infect Dis J 1992;11:785-814.
50. PIOPED TI. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the prospective investigation of pulmonary embolism diagnosis (PIOPED). JAMA 1990;263:2753-2759.

51. Elmore JG, Wells CK, Lee CH, Howard DH, Feinstein AR. Variability in radiologists' interpretations of mammograms. N Engl J Med 1994;331:1493-1499.
52. Weller SC, Mann NC. Assessing rater performance without a "gold standard" using consensus theory. Med Decis Making 1997;17:71-79.
53. Hoes AW, Mosterd A, Grobbee DE. An epidemic of heart failure. Recent evidence from Europe. Eur Heart J 1998;19:L2-L9.
54. Bergus GR, Chapman GB, Levy BT, Ely JW, Oppliger RA. Clinical diagnosis and the order of information. Med Decis Making 1998;18:412-417.
55. Swets JA. Measuring the accuracy of diagnostic systems. Science 1988;240:1285-1293.
56. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996;15:361-87.
57. Little RA. Regression with missing X's: A review. J Am Stat Assoc 1992;87:1227-1237.
58. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. Am J Epidemiol 1995;142:1255-1264.
59. Schafer JL. Multiple imputation: a primer. Stat Methods Med Res 1999;8:3-15.
60. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. Jama 1997;277(6):488-494.
61. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 1982;143:29-36.
62. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839-843.
63. Feinstein AR. Clinical biostatistics. XLI. Hard science, soft data, and the challenges of choosing clinical variables in research. Clin Pharmacol Therap 1977;22:485-498.
64. Centor RM, Schwartz JS. An evaluation of methods for estimating the area under the receiver operating characteristic (ROC) curve. Med Decis Making 1985;5:149-156.
65. Hosmer DW, Lemeshow S, editors. Applied logistic regression. New York: John Wiley & Sons, Inc; 1989.
66. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol 1996;49:1373-1379.
67. Moons KGM, Stijnen T, Michel BC, Buller HR, Grobbee DE, Habbema JDF. Treatment thresholds in diagnostic test evaluation: an alternative approach to the comparison of areas under the receiver operating characteristic curve. Med Decis Making 1997;17:447-454.

4.2 External validation is necessary in prediction research: a clinical example

SE Bleeker, HA Moll, EW Steyerberg, ART Donders, G Derksen-Lubsen, DE Grobbee, KGM Moons

Submitted

Abstract

Prediction models tend to perform better on data on which the model was constructed than on new data. This difference in performance is an indication of the optimism in the apparent performance in the derivation set. For internal model validation, bootstrapping methods are recommended to provide bias corrected estimates of model performance. Results are often accepted without sufficient regard to the importance of external validation.

This report illustrates the limitations of internal validation to determine generalisability of a diagnostic prediction model to future settings. A prediction model for the presence of serious bacterial infections in children with fever without source was derived and validated internally using bootstrap resampling techniques. Subsequently, the model was validated externally.

In the derivation set (n=376), nine predictors were identified. The apparent area under the receiver operating characteristic curve (95% confidence interval) of the model was 0.83 (0.78-0.87) and 0.76 (0.67-0.85) after bootstrap correction. In the validation set (n=179) the performance was 0.57 (0.47-0.67).

Thus, for relatively small data sets, internal validation of prediction models by bootstrap techniques may not be sufficient and indicative for the model's performance in future patients. External validation is essential before implementing prediction models in clinical practice.

pattern, vital signs, clinical appearance and information on ear-nose-throat, skin, and the respiratory-, circulatory- and abdominal tract.

Reference standard. For each patient, the final diagnosis was determined either by a reference standard (cultures of blood, spinal fluid, urine, stool positive for a pathogen) or based on a consensus diagnosis (27). Outcome diagnosis was the presence or absence of a serious bacterial infection. A serious bacterial infection was defined as the presence of bacterial meningitis, sepsis or bacteraemia, pneumonia, urinary tract infection, bacterial gastroenteritis, osteomyelitis or ethmoiditis (24, 26). As obtaining cultures of blood, spinal fluid, urine or stool was dependent on the clinical evaluation and not required by protocol for each patient, a follow-up period of two weeks was used as the standard for ruling out the possibility of a missed diagnosis of serious bacterial infection.

Data analysis. The association between potential diagnostic determinants and the presence of a serious bacterial infection was assessed using logistic regression analyses (SPSS version 9.0). Continuous variables were analysed both on a linear and a transformed scale, e.g. logarithmic or quadratic, to determine which scale was the better predictor of outcome (2). Variables with 50% or more missing values were excluded from the analyses. Then, based on the literature and clinical practice (18, 19, 21-23, 28-36), 57 variables were considered as candidate predictors for the analyses. Of these, variables with a univariable p-value of 0.15 or less were subsequently entered into a forward stepwise multivariable logistic regression procedure. Variables with a multivariable p-value of less than 0.10 and clinically relevant were selected as predictors of a serious bacterial infection.

Some of the 57 variables had missing values. Simple exclusion of patients with missing values on one or more of the variables commonly causes biased results and decreases statistical efficiency (37, 38). Therefore, missing values in the data were completed by single imputation using SOLAS (version 2.0). This method uses all available data to impute the missing values based on the correlation between each variable with missing values and all other variables (38).

The ability of a prediction model to discriminate between children with and without a serious bacterial infection was quantified using the area under the receiver operating characteristic curve (ROC area) (2). The ROC area indicates the likelihood that a patient with a serious infection has a higher predicted probability among a

randomly chosen pair of patients of whom only one has a serious infection. Although encountering such a pair of patients is a rather artificial situation, the ROC area is often used as the primary criterion to quantify model performance. As a measure of overall performance we considered the explained variation (R^2) by the model using the definition proposed by Nagelkerke (39). R^2 quantifies the explained variation on the loglikelihood scale, which is the natural scale to study performance of logistic regression models.

Internal validity of the model was determined by using bootstrap techniques (2, 12). Random bootstrap samples were drawn with replacement from the derivation set consisting of all patients (200 replications). Using S-plus (version 2000) both the univariable selection of variables with a p-value less than 0.15 and the multivariable selection of variables with a p-value less than 0.10 were repeated within each bootstrap sample. The model as estimated in the bootstrap sample was evaluated in the bootstrap sample and in the derivation set. The difference between the performance in the bootstrap sample and the performance in the derivation set was considered as an estimate of the optimism in the apparent performance in the derivation set. This difference was estimated for each bootstrap sample (200 times). The 200 differences were averaged to obtain a stable estimate of the optimism. The optimism was subtracted from the apparent performance in the derivation set to estimate the internally validated performance (2, 3, 16). Further, a shrinkage factor was derived from the bootstrap samples by calculating the slope of the linear predictor in the derivation set, where the linear predictor was calculated with the regression coefficients as estimated in the bootstrap sample. The shrinkage factor was used as a multiplier for the logistic regression coefficients in order to re-calibrate the predictive model (2-4, 40, 41). To estimate the 95% confidence interval (95%CI) around the performance measures we used the empirical distributions in the 200 bootstrap samples (12).

External validation study

Patients, diagnostic determinants and reference standard. To determine generalisability of the derived prediction rule to new patients, the rule was applied to a new data set (validation set, n=179). This validation set included children from a different time period and from an additional Children's Hospital from a different city. It comprised children with fever without apparent source who visited the Sophia Children's University Hospital Rotterdam between 1997 and 1998 and the Juliana Children's

Hospital in The Hague in 1998. Both hospitals are large inner city teaching hospitals in the Netherlands. Other inclusion and exclusion criteria were identical to those used for the derivation set. Data collection, definitions of diagnostic determinants and the reference standard were identical to the derivation study.

Data analysis. The prediction model was applied to the children in the validation set. The performance (ROC area and R^2) of the model as well as the calibration were assessed. A graphical impression of the calibration of model predictions in the validation set was obtained by plotting the observed proportions versus predicted probabilities (42). In addition, subgroup analyses per hospital (Sophia Children's University Hospital Rotterdam and Juliana Children's Hospital) were performed. Subsequently, we hypothesised that the multivariable associations of the predictors with the outcome in the validation set would not differ from those in the derivation set. As an overall test of this hypothesis we compared the re-estimated regression coefficients in the validation set with the regression coefficients from the derivation set before bootstrapping. Hereto, a logistic regression analysis was performed in the validation set including a linear predictor variable based on the coefficients from the derivation set as an offset variable. This analysis assumes the regression coefficients of the derivation set to be fixed. It is a one-sample test for the coefficients in the validation set.

Results

The derivation set comprised 376 children with fever without apparent source and the validation set consisted of 179 children who had been referred for the same reason (3 respectively 0 patients were excluded because of isolation of Hib). Except for the variable pale skin, no material differences were found in the distribution of the general characteristics and the predictors between the two sets (table 1). A serious bacterial infection was present in 20% of the children in the derivation set and in 25% of the validation set. Of the 57 considered variables in the univariable analyses, 34 had a p-value of 0.15 or less and 24 variables had a p-value of 0.05 or less. Table 2 shows the variables with a univariable p-value of 0.01 or less and the results of the multivariable analysis. Strong predictors of serious bacterial infection included age above 1 year, duration of fever, changed crying pattern, nasal discharge or earache in history, ill clinical appearance, pale skin, chest-wall retractions, crepitations and signs of

pharyngitis or tonsillitis. The ROC area (95%CI) of this model was 0.825 (0.78-0.87) and the R^2 (95%CI) 32.3% (15.1%-49.4%). The estimated optimism by bootstrapping was 0.068 and 14.1%, reducing the ROC area (95%CI) and R^2 (95%CI) to 0.756 (0.66-0.86) and 18.0% (5.7%-30.0%) respectively. The shrinkage factor (95%CI) for correction of the regression coefficients was 0.66 (0.38-0.93).

Table 1. Distribution of patient characteristics in the derivation and validation set*.

	Derivation set SCH† (1988-1992; n‡=376)	Validation set SCH, JCH† (1997-1998; n=179)
Male gender	214 (57)	91 (51)
Age (years) §	1.1 (0.7)	1.0 (0.8)
Weeks of gestation§	39.3 (2.1)	39.1 (2.3)
Duration of fever (days) §	2.8 (2.4)	2.8 (2.5)
Changed crying pattern (only <1 year)	139 (37)	79 (44)
Nasal discharge or earache	246 (65)	129 (72)
Ill clinical appearance	219 (58)	103 (58)
Body temperature at physical examination (°C)§	39.7 (1.0)	39.8 (0.9)
Pale skin	66 (18)	15 (8)
Chest-wall retractions	18 (5)	16 (9)
Crepitations	15 (4)	7 (4)
Signs of pharyngitis or tonsillitis	159 (42)	82 (46)
Serious bacterial infection present	75 (20)	45 (25)

*Values represent absolute patient numbers (percentages) unless stated otherwise.

†SCH: Sophia Children's University Hospital; JCH: Juliana Children's Hospital.

‡Number of patients.

§Mean (standard deviation).

Subsequently, the model was applied to the validation set to test its predictive performance. The ROC area (95%CI) dropped to 0.57 (0.47-0.67) and the R^2 to 2.0%. Figure 1 shows a poor calibration of the model in the validation set. The predicted probabilities of the presence of a serious bacterial infection ranged from 0.02 to 0.73 (mean 0.19). In particular, in the clinically important lower categories the predicted probabilities corresponded poorly with the observed proportions. Similar differences were found when the model was tested in the children of the Sophia Children's

University Hospital Rotterdam or to the children of the Juliana Children's Hospital separately.

The poor results in the external validation were confirmed by refitting the multivariable model (i.e. re-estimating the regression coefficients) on the data of the validation set. Overall, the regression coefficients in the validation set were significantly different from the derivation set, in particular, the regression coefficients of four of the nine predictors (changed crying pattern, nasal discharge or earache, chest-wall retractions and crepitations). The ROC area (95%CI) of this refitted model was 0.70 (0.61-0.79), and the R^2 16.9%.

Table 2. Derivation set: variables with univariable p-value ≤0.01 and results of multivariable analysis.

	Percentage		Odds Ratio (95%CI)	
Characteristic	SBI* absent (n†=301)	SBI present (n=75)	Univariable	Multivariable‡
Patient history				
Age >1 year	51	57	1.3 (0.8-2.2)	0.4 (0.1-1.1)
Duration of fever (days) §	2.5 (2.2)	3.8 (2.8)	1.22 (1.10-1.34)	1.26 (1.12-1.41)
Changed crying pattern (only <1 year)	34	51	2.0 (1.2-3.4)	5.0 (1.7-15.0)
Nasal discharge or earache	69	51	0.5 (0.3-0.8)	0.4 (0.2-0.7)
Physical examination				
Body weight (kilograms) §	9.8 (2.9)	8.9 (3.4)	0.89 (0.82-0.98)	-
Ill clinical appearance	55	73	2.3 (1.3-4.0)	2.6 (1.4-5.0)
Poor peripheral circulation	10	25	3.2 (1.7-6.1)	-
Pale skin	14	33	3.2 (1.7-5.7)	2.2 (1.1-4.4)
Chest-wall retractions¶	4	9	2.7 (1.0-7.3)	3.4 (1.1-10.1)
Crepitations	2	12	6.7 (2.3-19.5)	12.8 (3.8-42.9)
Signs of pharyngitis or tonsillitis	47	24	0.4 (0.2-0.6)	0.3 (0.2-0.6)

*serious bacterial infection.

†Number of patients.

‡Intercept of the model was –2.29.

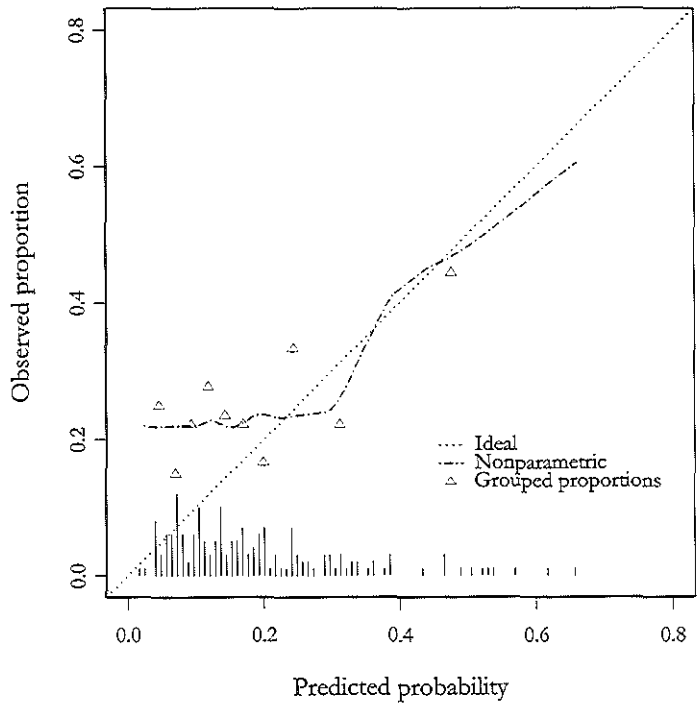
§Mean (standard deviation).

¶Univariable p-value: 0.06.

Discussion

The aim of this study was to construct and validate a diagnostic prediction model to distinguish children with and without serious bacterial infections in children referred with fever without apparent source. Selected predictors in the derivation set were age above 1 year, duration of fever, changed crying pattern, nasal discharge or earache in history, ill clinical appearance, pale skin, chest-wall retractions, crepitations and signs of pharyngitis or tonsillitis. These results agree with other published studies (19, 28, 29, 31, 32, 34-36). The model seemed to perform well, according to common performance criteria (ROC area and R^2). Bootstrapping suggested a substantial optimism. Nonetheless, applying the model to the validation set showed a much larger decrease in predictive performance. This decrease was of such a degree that the model

Figure 1. Comparison of the model's predicted probabilities and observed proportions. Triangles: deciles of the predicted probabilities in the validation set (n=179). Diagonal line: reflection of ideal situation (predicted probability = observed proportion). Dashed line: reflection of the relation non-parametrically. Lower part of the figure: histogram of the predicted probabilities.



appeared useless for paediatric care. Hence, our study illustrates and confirms that internal validation per se is no guarantee for generalisability and thus no substitute for external validation (5).

Although external validation is needed, results of external validations are not always unambiguous and trustworthy. In new settings, a novel external validation may be required with a substantial sample size to provide sufficient power to detect possible differences in performance (5).

It is generally recommended that the overall characteristics of the derivation and validation set should be compared before an internally validated prediction model is applied to a new patient population(6, 8). These characteristics include predictors in the regression model, but also general aspects such as the selection of patients (e.g. referral pattern) and the definition of predictors. If no important differences are found, the validation set is usually considered to provide a comparable population. Subsequently, a largely similar performance as estimated from the derivation set should be found in the validation set. Table 1 showed no major differences between the two data sets. Apparently, this does not assure good performance of the model in a validation set.

Some aspects of our study need to be addressed to appreciate the results. Firstly, we considered a relatively large number of candidate predictors. Secondly, by using stepwise selection methods, an unstable selection of variables may arise by limited power to select diagnostically important variables in small data sets. Multiple testing and underestimation of standard errors and p-values are issues of concern (1, 10, 11, 16). Sample sizes were small in both the derivation (n = 376) and validation set (n = 179). Finally, we only applied single imputation instead of a more appropriate multiple imputation (37). Moreover, the single imputation was only performed in the derivation set and was not repeated in the bootstrap procedure. Anyhow, we believe that our main finding and conclusion still stands, as model selection aspects were considered in the internal validation procedure. In particular, if an external validation had not been performed, we would not have known that the model's performance after internal validation was no guarantee of its performance in future but comparable children.

Apart from the above-mentioned statistical drawbacks, we have tried to discover other reasons of this poor validation. Selection bias seems unlikely, because the same in- and exclusion criteria were used. Information bias seems improbable as well, since all data of both the derivation and the validation set were collected by the same persons and before any analysis had been started. There was a slight, statistically non-

significant, difference in frequency of serious bacterial infection between the derivation set (20%) and the validation set (25%), reflecting a difference in baseline risk (i.e. the intercept of the model). It is unlikely that this difference has affected the discriminative performance substantially, since the frequency does not directly influence estimation of the ROC area. The drop in R^2 may however partly be explained by this miscalibration. Besides the main analysis, we performed a subgroup analysis per hospital. No materially different results were found. In addition, we re-estimated the prediction model on the validation set to obtain insight in the maximally achievable performance of the model in the validation set. The ROC area of 0.70 and R^2 of 17% suggests that certain predictors in the original model still have some value in the validation set, although the magnitude of the associations (regression coefficients) has changed. Indeed, when exploring the validation set in further detail, the distribution of the signs and symptoms across children with and without a serious bacterial infection in the validation set appeared to be different from the derivation set. This suggests a difference in patient population, which was not exhibited by the comparison of the overall distribution (table 1). Such change in patient population might be an effect of a change in referral pattern by general practitioners, probably influenced by the introduction of the Hib vaccination (April 1993) for young infants. Besides, as the occurrence of Hib in the isolates was not appreciably changed (3 before and 0 after the introduction of the vaccine) and as those cases were excluded from the derivation set, it would be reasonable to consider the effect of vaccination on the generalisability minor. An alternative, though unlikely, explanation is that the poor validation is the result of an unfortunate data set. However, we can not exclude this contention. Furthermore, this lack of generalisability to future groups of children supports the view that clinical guidelines may not be durable with time and must be updated regularly (43).

In conclusion, after internal validation by bootstrapping of a diagnostic prediction model for serious bacterial infection in children with fever without apparent source, we had good expectations with respect to the performance of the model in the validation set. However, results from external validation showed an unexpected poor performance. This suggests that internal validation of prediction models may not be sufficient for relatively small data sets and that external validation is necessary before implementing predicting models in clinical practice.

Acknowledgement

We gratefully acknowledge Wilfried de Jong, and Femke Mineur, medical students for support in data collection. The Health Care Insurance Council of the Netherlands financially supported this project.

References

1. Harrell FE, Lee KL, Matchar DB, Reichert TA. Regression models for prognostic prediction: Advantages, problems, and suggested solutions. Cancer Treat Rep 1985;69:1071-7.
2. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine 1996;15(4):361-87.
3. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Ass 1983;78:316-31.
4. Houwelingen JC van, Le Cessie S. Predictive value of statistical models. Stat Med 1990;9(11):1303-25.
5. Justice AC, Covinsky KE, Berlin JA. Assessing the generalisability of prognostic information. Ann Intern Med 1999;130(6):515-24.
6. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. New England Journal of Medicine 1985;313(13):793-9.
7. Moons KG, Es GA van, Michel BC, Buller HR, Habbema JD, Grobbee DE. Redundancy of single diagnostic test evaluation. Epidemiology 1999;10(3):276-81.
8. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. Jama 1997;277(6):488-94.
9. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. Jama 2000;284(1):79-84.
10. Steyerberg EW, Eijkemans MJ, Habbema JD. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. J Clin Epidemiol 1999;52(10):935-42.
11. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med 2000;19(8):1059-79.
12. Efron B, Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. New York: Chapman & Hall; 1993.
13. Efron B, Tibshirani R. Improvements on cross-validation: the .632+ bootstrap method. JASA 1997;92:548-60.
14. Picard RR, Berk KN. Data splitting. Am Statistician 1990;44:140-7.

15. Sauerbrei W. The use of resampling methods to simplify regression models in medical statistics. J R Stat Soc C Appl Stat 1999;48:313-329.
16. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models. Efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54(8):774-81.
17. Knottnerus JA. Prediction rules: statistical reproducibility and clinical similarity. Med Decis making 1992;12:286-287.
18. McCarthy PL. Fever. Pediatr in Rev 1998;19:401-8.
19. Baraff LJ, Lee SI. Fever without source: management of children 3 to 36 months of age. Pediatr Infect Dis J 1992;11(2):146-51.
20. Steensel-Moll HA van, Jongkind CJ, Aarsen RSR, et al. [A problem-oriented patient classification system for general paediatrics II]. Tijdschr Kindergeneesk 1996;64(3):99-104.
21. Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteraemia in young febrile children. Ann Emerg Med 1998;31(6):679-87.
22. Baker MD, Bell LM, Avner JR. Outpatient management without antibiotics of fever in selected infants. N Engl J Med 1993;329(20):1437-41.
23. Baraff LJ, Oslund SA, Schriger DL, Stephen ML. Probability of bacterial infections in febrile infants less than three months of age: a meta-analysis. Pediatr Infect Dis J 1992;11(4):257-64.
24. Bleeker SE, Moons KGM, Derksen-Lubsen G, Grobbee DE, Moll HA. Predicting serious bacterial infection in young children with fever without apparent source. Acta Paediatrica 2001;90(11):1226-1232.
25. Oostenbrink R, Moons KGM, Donders ART, Grobbee DE, Moll HA. Prediction of bacterial meningitis in children with meningeal signs: reduction of lumbar punctures. Acta Paediatr 2001;90(6):611-7.
26. Oostenbrink R, Moons KGM, Theunissen CC, Derksen-Lubsen G, Grobbee DE, Moll HA. Signs of meningeal irritation at the emergency department: how often bacterial meningitis? Pediatr Emerg Care 2001;17(3):161-4.
27. Weller SC, Mann NC. Assessing rater performance without a "gold standard" using consensus theory. Med Decis Making 1997;17(1):71-9.
28. McCarthy PL, Lembo RM, Baron MA, Fink HD, Cicchetti DV. Predictive value of abnormal physical examination findings in ill-appearing and well-appearing febrile children. Paediatrics 1985;76(2):167-71.
29. McCarthy PL, Lembo RM, Fink HD, Baron MA, Cicchetti DV. Observation, history, and physical examination in diagnosis of serious illnesses in febrile children less than or equal to 24 months. J Pediatr 1987;110(1):26-30.
30. Kramer MS, Lane DA, Mills EL. Should blood cultures be obtained in the evaluation of young febrile children without evident focus of bacterial infection? A decision analysis of diagnostic management strategies. Paediatrics 1989;84(1):18-27.

31. Baraff LJ, Bass JW, Fleisher GR, et al. Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Paediatrics 1993;92:1-12.
32. Dagan R, Powell K, Hall C, Menegus M. Identification of infants unlikely to have serious bacterial infection although hospitalised for suspected sepsis. J Pediatr 1985;107:855-860.
33. Isaacman DJ, Shults J, Gross TK, Davis PH, Harper M. Predictors of bacteraemia in febrile children 3 to 36 months of age. Paediatrics 2000;106(5):977-82.
34. Hewson PH, Humphries SM, Roberton DM, McNamara JM, Robinson MJ. Markers of serious illness in infants under 6 months old presenting to a children's hospital. Arch Dis Child 1990;65(7):750-6.
35. Berger RM, Berger MY, Steensel-Moll HA van, Dzoljic-Danilovic G, Derksen-Lubsen G. A predictive model to estimate the risk of serious bacterial infections in febrile infants. Eur J Pediatr 1996;155(6):468-73.
36. Teach SJ, Fleisher GR. Duration of fever and its relationship to bacteraemia in febrile outpatients three to 36 months old. The Occult Bacteraemia Study Group. Pediatr Emerg Care 1997;13(5):317-9.
37. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analyses. Am J Epidemiol 1995;142(12):1255-64.
38. Little RA. Regression with missing X's: A review. J Am Stat Assoc 1992;87:1227-1237.
39. Nagelkerke N. A note on the general definition of the coefficient of determination. Biometrika 1991;78:691-2.
40. Houwelingen JC van. Shrinkage and penalized likelihood as methods to improve predictive accuracy. Stat Neerl 2001;55:17-34.
41. Steyerberg EW, Eijkemans MJC, Habbema JDF. Application of shrinkage techniques in logistic regression analysis: a case study. Stat Neerl 2001;55:76-88.
42. Hilden J, Habbema JD, Bjerregaard B. The measurement of performance in probabilistic diagnosis. II. Trustworthiness of the exact values of the diagnostic probabilities. Methods Inf Med 1978;17(4):227-37.
43. Shekelle PG, Eccles MP, Grimshaw JM, Woolf SH. When should clinical guidelines be updated? Bmj 2001;323:155-7.

4.3 A computer-based patient record with structured data entry for a comprehensive patient history and physical examination in paediatrics

SE Bleeker, G Derksen-Lubsen, AM van Ginneken, J van der Lei, HA Moll

Submitted

Abstract

In current practice, a diversity of paper charts is utilised and patient data are fragmentary documented. Furthermore, research on hand-written documents is labour-intensive. A computer-based patient record with structured data entry can reconcile these difficulties. Aim of this project was to develop a computer-based patient record for recording data from a comprehensive patient history and physical examination in paediatrics.

Open Record for Care–Structured Data Entry was used, in which the context of medical concepts is represented by a tree structure. Users' perspectives of the computer-based patient record were restrictedly evaluated.

In the paediatric computer-based patient record, patient history is described by twenty main concepts, the physical examination by ten. In total, the tree consists of 8200 nodes. Specific custom forms can be composed according to one's needs. The content of the tree can be adjusted easily and sharing records among different disciplines is possible. Conclusively, in paediatrics, a computer-based patient record with structured data entry supported by a tree was developed. In particular, collection of structured data from a comprehensive patient history and physical examination is possible in this computer-based patient record. Such a computer-based patient record offers many potential benefits and implementation might be feasible.

Introduction

At present, medical specialists still use the paper chart as the traditional medium for collecting and recording patient data. Each specialist records patient data in his own discipline and a great variety of paper charts is in use. In addition, fragmentation of patient data due to use of scattered sources endangers the continuity of care. Furthermore, research on non-standardised data on hand-written paper documents is labour-intensive and complicated by incompleteness, in particular regarding data from patient history and physical examination.

A computer-based patient record (CPR) with structured data entry can overcome these drawbacks (1). For optimal embedding of a CPR in daily practice the system has to be useful in practice, has to meet specific needs of the clinician in a tailor-made fashion, and has to require minimal extra work from the clinician (2). At present, such CPRs are sparsely available for concise subject areas, e.g. radiology and endoscopy (3-5). However, in literature we found little evidence of an applied CPR that supports structured data entry for patient data from a comprehensive case history and physical examination in a general and broad specialism, whereas it would be appreciated (6). The main challenging difficulty in the achievement of structured data entry for such a comprehensive case history and physical examination is the controversy of structuring complex descriptions of patients' problems with fixed components.

Aim of this project was to develop a CPR applicable for the general paediatric practice, in particular for recording data from patient history and physical examination, and which could be linked easily to existing hospital information systems. In this paper, we describe the achievement of structured data entry for such a CPR in paediatric practice. Briefly, we also report on the first experiences of physicians to assess the feasibility of implementation of this CPR in routine clinical practice.

Material and methods

Background

Recorded patient data can be divided into two main categories: specialism independent and specialism dependent data. Specialism independent implies that the attributes with which these data are described do not differ per discipline. For example, a laboratory test result (e.g. a serum glucose level) is described by the type of the test, its value and its unit. These descriptors are identical in each specialism.

On the other hand, specialism dependent means that data are described differently in each discipline. For example, a paediatric case history differs greatly from that of an orthopaedic surgeon. Specialism dependent data typically are data from patient history, physical examination and data from additional examinations, such as radiology or pathology reports. The relative importance and usefulness of different descriptors and values varies greatly between clinical specialisms.

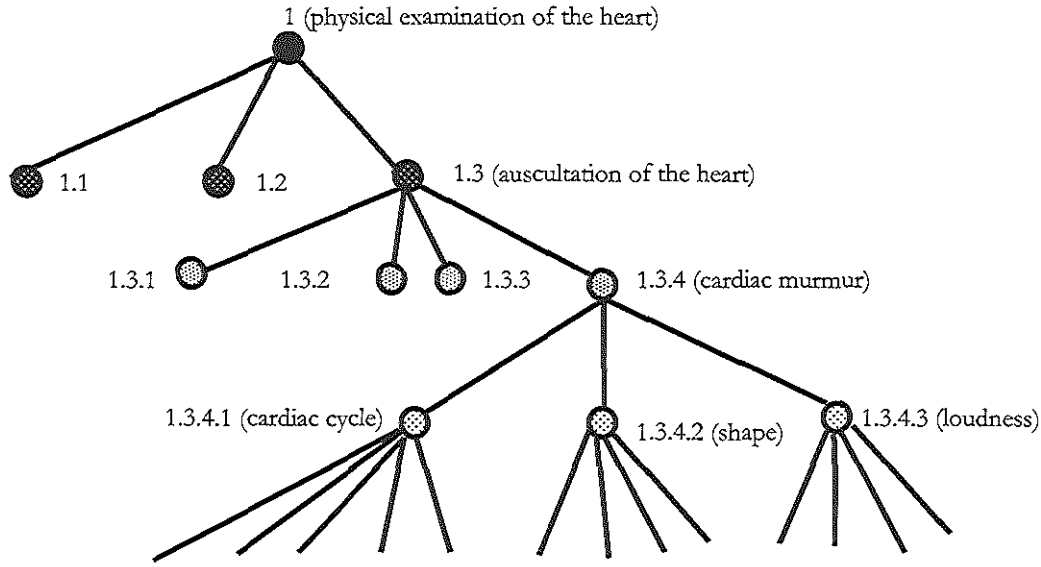
Materials

Open Record for Care–Structured Data Entry (ORCA-SDE), version 2, was used for the support of structured data entry for the CPR in paediatrics, as it appeared to fulfill our requirements: flexibility in content, uniformity of the underlying structure, and provision of unambiguous data that can be coded. In particular, ORCA-SDE is able to accommodate data from a broad specialism, in which the possibility of a detailed description of patient history and physical examination is necessary, and specific needs of each specialism can be tailor-made. It provides interchangeability of data among disciplines (shared records), and an eventual adjustment of the content covered by structured data entry does not require a change in database structure or software (7, 8). As recommended, ORCA-SDE presents data with explicit ordering, clear headings and subheadings, and custom views on data (9, 10).

In ORCA-SDE medical concepts are represented by a tree structure (figure 1). Each node in the tree represents a medical concept and its branches represent its descriptors. Medical descriptions are by nature hierarchical. For instance, physical examination of the heart (1) can be described by palpation (1.1), percussion (1.2) and auscultation (1.3) of the heart. Auscultation of the heart itself can be described by heart rate (1.3.1), heart sounds (1.3.2), heart rhythm (1.3.3), and a possible cardiac murmur (1.3.4). A cardiac murmur again can be described by its phase in the cardiac cycle (1.3.4.1), its shape (1.3.4.2), its loudness (1.3.4.3), etceteras. This tree represents the content in which medical concepts can be used to describe and store patient data. The task of ordering the medical concepts in the tree, and assigning specific parameters to the medical concepts, is called modelling.

ORCA-SDE can operate as a module in a clinical information system, to form a more complete CPR, with e.g. laboratory results, radiology reports, hospitalisation data, follow-up appointments, a problem-oriented patient classification system, and ICD9 classification.

Figure 1. A tree structure. The digits refer to the items listed in the text.



Methods

This study was approved by The Institutional Review Board of University Hospital Rotterdam. In close consultation with two paediatricians, the content of the paediatric history and physical examination was reviewed using the national standardised paediatric medical record and a paediatric textbook (11). Subsequently, this content was modelled into the ORCA-SDE tree structure. Special paediatric terminology and standard nomenclature were applied (11).

Five (trainee) paediatricians used the CPR with ORCA-SDE in a research setting. Each paediatrician was asked to select at least five patients, who had presented at ambulatory care with a new problem. Of those patients, data from the medical records were entered in the CPR. Prior to data entry, all participants received a demonstration of ORCA-SDE by the first author, and a printed description of the interface. To avoid any influence on the participants' behaviour no observer was present during data entry. After data entry, each participant filled in a questionnaire with respect to the users' perspectives of the paediatric CPR. These comprised the completeness (availability of medical concepts and its descriptors), user-friendliness (easiness to

familiarise oneself with the interface and tree structure) and the usefulness (capability to be used advantageously and for several purposes) of the structured data entry. Completeness was evaluated for the patient history and physical examination separately. Each item could be rated from 0 (worst) to 5 (best).

Results

Structure of the tree

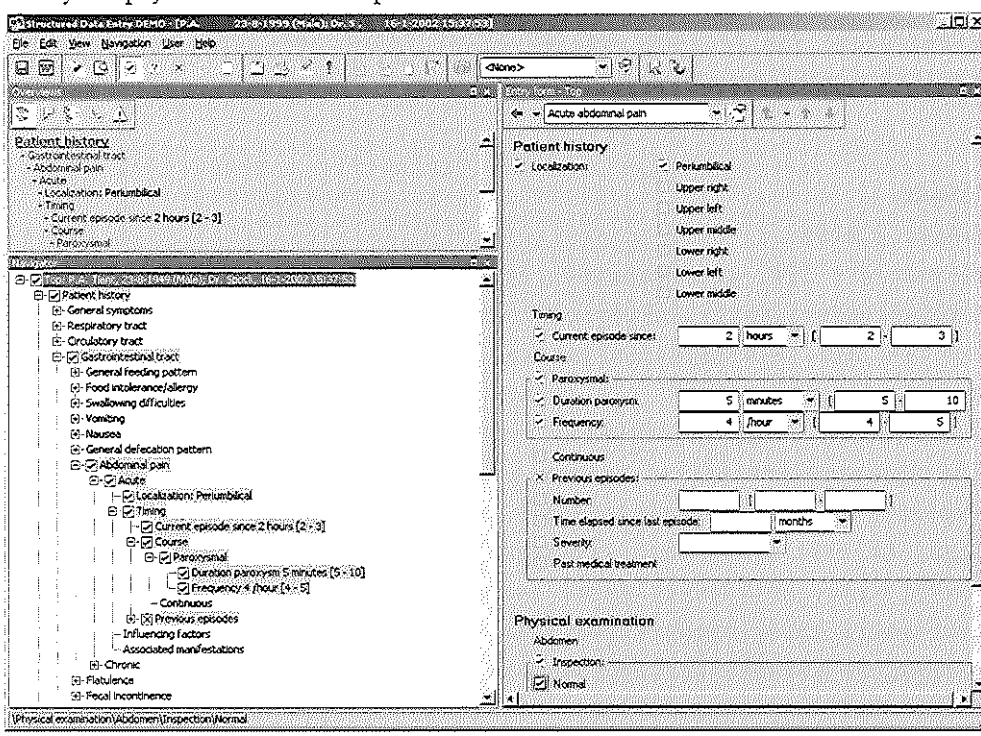
As part of a CPR, we constructed a tree for structured data entry of data from paediatric patient history and physical examination (figure 2). The top of the tree in this CPR starts with history and physical examination. History divides in branches with the specified concepts past medical history (including immunisation status), family history, allergies, social history, current medications and the chief complaint. The chief complaint exists of a general tract and fourteen specific history tracts (the respiratory-, circulatory-, gastrointestinal-, and urogenital tract, ear-nose-throat, skin, organs of sense, the nervous-, endocrine-, locomotor- and haematological system, feeding history, birth history, growth and developmental history). Each of these, in total twenty, history concepts splits again in five to twenty-five sub branches. For instance, the gastrointestinal tract subdivides in nine main concepts: general feeding pattern (including undernutrition, overnutrition, and appetite, which is linked to feeding history from the list of specific history tracts), intolerance- or allergy of food, swallowing difficulties, vomiting, nausea, general defecation pattern (including diarrhoea, constipation, mucous stools, painful defecation and bloody stools), abdominal pain (acute and chronic), flatulence and faecal incontinence. Commonly, all sub branches are in the end described by four to fifteen general attributes, like duration, severity, timing and setting of complaint, influencing factors of complaint and associated manifestations.

The physical examination splits in ten branches (general survey and vital signs, head and neck, thorax, abdomen, skin, limbs, genitalia, anus and rectum, nervous system, and spine), of which each forks as well. In total, the thesaurus consists of about 1800 items, used in 8200 nodes in the tree.

Data entering

When the user selects a concept in the tree, a data entry form is displayed with the descriptors of the selected concepts as options for data entry (figure 2). Data entry is

Figure 2: Computer-based patient record with structured data entry for data from patient history and physical examination in paediatrics.



accomplished primarily with the mouse: the user selects from pull-down boxes and pick-lists. Concepts can be marked as 'present', 'absent' or 'unknown'. Free text annotations can be added at each item. Values, e.g. body temperature, have to be entered by the keyboard. When an implausible value is entered, e.g. 118 degrees Fahrenheit for body temperature, the program shows a warning and the item will be presented in red (figure 3). However, clinicians are not obligated to correct it.

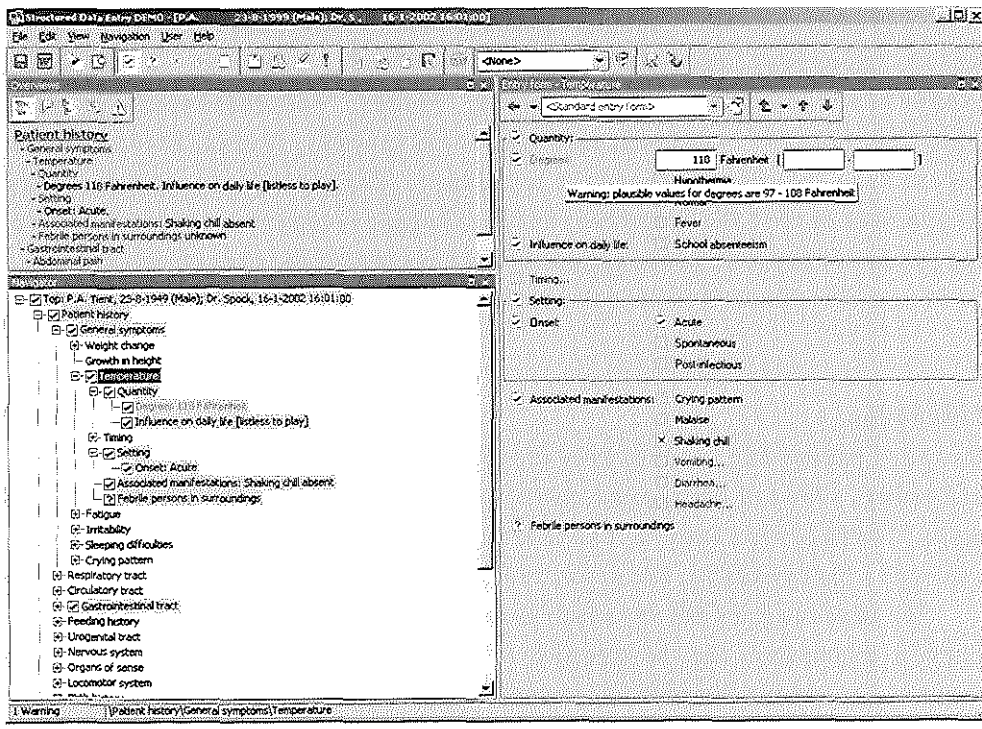
A search option is available to locate a specific item in the tree. Entering a search term by keyboard results in a number of hits, including synonyms. Each hit is displayed with its pathway, showing its position in the tree.

To improve workability, we made use of several shortcuts. For example, fever is described in the general history tract. In differential diagnostic thinking, however, fever is an important associated manifestation of many symptoms and has to be addressed in several history tracts. Through a shortcut, searching and scrolling

through the tree, and repeatedly recording of data in multiple places is avoided. For instance, in the history taking of coughing, which is described in the respiratory tract, fever has to be inquired about as well. When the clinician selects in the respiratory tract the shortcut 'fever', the program directly jumps to the descriptors of fever in the general history tract. Additional examples of shortcuts are 'irritability', 'nausea', 'appetite' and 'headache'.

At each moment the clinician is free to choose the starting point and endpoint (degree of detail) of data entry. Moreover, we developed an application, by which users can compose custom forms, tailoring it to the specific needs of a specialism, specific problem or disease entity. On such a custom form, medical concepts from different locations in the tree can be combined. For instance, one can develop a form for patients presenting with acute abdominal pain, by which signs and symptoms, like localisation and timing of the complaint, and physical examination of the abdomen, are displayed (figure 2). In fact, custom forms are user-defined views on the concepts

Figure 3. A warning at an implausible value.



in the tree. The possible number of custom forms to compose is unlimited.

The program provides the possibility to export entered data to a relatively simple internal text editor or to Microsoft Word. The exported data will comprise basic administrative patient data, date of data collection, name of the clinician who recorded the data and the collected data from history taking and physical examination in outline (indented) display. The text can be edited and completed by for instance a description of the interpretation of the presented clinical problem, a diagnosis or a treatment strategy. The resulting report may be used as a letter to the general practitioner.

The tree as used in ORCA-SDE is very flexible. The content of the tree can be restyled, expanded and adjusted over time easily, even without the expertise of a computer programmer. After any adjustment of the tree, previously stored data remain fully accessible.

Evaluation of users' perspectives of the CPR.

Each of the five participating (trainee) paediatricians entered data of at least five patients. The presenting problems of those patients included: general problems (excessive crying, malaise, icterus, constipation), infectious problems (fever and vomiting/diarrhoea, fever without apparent source, fever and dysuria, suspected sepsis), respiratory problems (coughing, dyspnoea, respiratory insufficiency), cardiovascular problems (cardiac arrest, souffle), endocrinological problems (thelarche, polyuria/ polydipsia) and behavioural problems (autism, attention deficit hyperactivity disorder, psychomotor retardation). The participants rated the users' perspectives of the CPR as specified (median (range)): completeness of the content of the patient history: 4 (3-5); completeness of the content of the physical examination: 5 (4-5); user-friendliness: 4 (1-4); and usefulness: 4 (3-5).

Discussion

Recorded patient data need to serve several purposes (diagnosis, reporting, research, and policy). Often, if using paper charts, the same efforts must be repeated all over again. Structured data entry in a CPR improves efficiency, in particular in recording data with a high variability, like in patient history and examination.

We developed a CPR by which data from a comprehensive case history and physical examination can be recorded in a structured fashion and which can be integrated to existing hospital information systems. ORCA-SDE was used by which

data entry is supported with a tree structure. We aimed to construct a tree with sufficient coverage and detail in the view of leading paediatricians. The 8200 nodes in the tree (1800 items) illustrate the degree of variability. However, if new insight would require adjustment or expansion of the tree, the content and structure of data in the CPR can be adapted easily, while previously stored data remain fully accessible. This flexibility is very important for acceptability of a CPR among future users. Furthermore, fixed forms are efficient though rigid with insufficient coverage, while interfaces with many data entry options are flexible though time-consuming. Hence, in order to meet both efficiency and flexibility, we created the option of composing tailor-made custom forms. Another essential benefit is the possibility that different specialists can share their records, while working with the same recognisable user interface. This enhances familiarity with the system and so the ultimate use of the CPR in practice. In this paper, the modelled content is of paediatric origin, though internal medicine could have been used as source as well. Currently, structured data entry based on ORCA-SDE is being applied to several medical domains (urology, neurology, burn care, psychiatry, radiology, pathology, cardiology) (12).

We exposed the CPR to a small number of paediatricians, to get a first impression of the evaluation of the users' perspectives. The opinions were rather supportive. However, the opinions on the user-friendliness were divided. We expect that this perspective coheres with experience and learning, and that acquaintance with the system will show improvement (1). The completeness of the content and the usefulness of the CPR were valued quite high. Nonetheless, this evaluation is too limited for definitive conclusions and necessary further exploration will be performed. However, based on this first impression, we belief that implementation of this CPR might be feasible.

In the Netherlands, almost all general practitioners use a computer to record patient data (2, 13). However, these records are character-based and data are entered as free text by keyboard, yielding a relatively crude structure, such as SOAP (Subjective-Objective-Assessment-Plan). As far as we are aware of, development of a CPR with structured data entry for a comprehensive case history and physical examination never has been realised before for a general and broad specialism such as paediatrics and internal medicine.

In general, using a CPR has several drawbacks. First, when introducing a CPR, the difficult problem of getting clinicians out of their traditional habit of using paper charts for recording patient data has to be faced. Then, the use of a CPR will be more

time-consuming at first and the clinician must invest energy in preparing his practice for computerisation to reach the optimal long-term gain. Besides, computer dependency can be a hindrance. Investment in equipment, maintenance, and proper security of a CPR is necessary. Moreover, clinical care requires that patients be described rather than labelled. A CPR rarely captures the detailed content of medicine, notwithstanding its comprehensiveness (6).

Nonetheless, the potential benefits of a CPR are many, both for patient care, quality assurance, research, and policy (1, 2, 13, 14). The main benefit of using a CPR is yielding bonded data, all gathered in the same patient record. Also, data will be more uniform, legible, and reporting of data (e.g. by a standard letter to the general practitioner) will be easier. The ability of access to the CPR at multiple locations overcomes the problem of physical paper records missing. Furthermore, electronic data interchange (e.g. for transferring patients), as well as more advanced decision support (e.g. by embedding clinical guidelines in the CPR), quality assessment, and patient-oriented clinical research could be realised. In the end, all these aspects together could result in a better and more complete documentation of patient data and a more efficient and better patient care (1, 14, 15).

Before widespread implementation of a CPR, the system should be rigorously evaluated and security and ethical aspects must be thought through carefully (14, 16, 17). Methodologies are being further developed, resulting in a possibly safer practice than in the paper era (2). With regard to both the physician acceptance and the patient acceptance, the final implementation must be guided carefully, although the latter appears to be greater than expected (18).

Special requirements for a CPR in paediatrics are either fulfilled or worked on (19). Among future plans for applications in our paediatric CPR with structured data entry are the embedding of age-based normal ranges of measurements for vital signs and other physiological parameters, clinical decision rules, and graphic display of growth data and special calculations of growth patterns. Further prospects are incorporation of reminder systems (e.g. for immunisation), selection of patient subgroups (e.g. for prevention strategies) and support of adequate drug dosing.

Conclusively, a CPR with structured data entry supported by a tree structure seems very valuable for practice and offers many potential benefits. This project resulted in an expandable CPR for paediatrics, in which in particular collection of structured data from a comprehensive patient history and physical examination is possible.

Acknowledgements

We express our gratitude to C. Nunnink, medical student for support in data collecting and M. de Wilde for technical support. We gratefully acknowledge C.B. Bouwhuis, MD, J.A. Hazelzet, MD PhD, W.J.W. Kollen, MD, PhD, J.M. Kouwenberg, MD and C.R. Lincke, MD for participating in the validation study of the CPR with structured data entry for paediatrics. The Health Care Insurance Council of the Netherlands financially supported this project.

References

1. Apkon M, Singhaviranon P. Impact of an electronic information system on physician workflow and data collection in the intensive care unit. Intensive Care Med 2001;27(1):122-30.
2. Knottnerus JA. Role of the electronic patient record in the development of general practice in the Netherlands. Methods Inf Med 1999;38(4-5):350-4.
3. Gouveia-Oliveira A, Salgado NC, Azevedo AP, et al. A unified approach to the design of clinical reporting systems. Methods Inf Med 1994;33(5):479-87.
4. Bell DS, Greenes RA. Evaluation of UltraSTAR: performance of a collaborative structured data entry system. Proc Annu Symp Comput Appl Med Care 1994:216-22.
5. Meijden MJ van der, Tange HJ, Boiten J, Troost J, Hasman A. An experimental electronic patient record for stroke patients. Part 2: system description. Int J Med Inf 2000;58-59:127-40.
6. Kirby J, Rector AL. The PEN&PAD data entry system: from prototype to practical system. Proc AMIA Annu Fall Symp 1996:709-13.
7. Ginneken AM van, Stam H, Mulligen EM van, Wilde M de, Mastrigt R van, Bemmel JH van. ORCA: the versatile CPR. Methods Inf Med 1999;38(4-5):332-8.
8. Ginneken AM van, Stam H, Moorman PW. A multi-strategy approach for medical records of specialists. Int J Biomed Comput 1996;42(1-2):21-6.
9. Nygren E, Wyatt JC, Wright P. Medical records: helping clinicians to find data and avoid delays. Lancet 1998;352:1462-6.
10. Wright P, Jansen C, Wyatt JC. Medical records: How to limit clinical errors in interpretation of data. Lancet 1998;352:1539-43.
11. Behrman RE, Kliegman RM, Arvin AM, Nelson WE, editors. Nelson textbook of paediatrics. 15th ed. Philadelphia: W.B. Saunders Company; 1996.
12. Pierik FH, Ginneken AM van, Timmers T, Stam H, Weber RF. Restructuring routinely collected patient data: ORCA applied to andrology. Methods Inf Med 1997;36(3):184-90.
13. Hiddema-van de Wal A, Smith RJ, Werf GT van der, Meyboom-de Jong B. Towards improvement of the accuracy and completeness of medication registration with the use of an electronic medical record (EMR). Fam Pract 2001;18(3):288-91.

14. Hunt DL, Haynes RB, Hanna SE, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review. Jama 1998;280(15):1339-46.
15. Schriger DL, Baraff LJ, Buller K, et al. Implementation of clinical guidelines via a computer charting system: effect on the care of febrile children less than three years of age. Journal of the American Medical Informatics Association 2000;7(2):186-95.
16. Mandl KD, Szolovits P, Kohane IS. Public standards and patients' control: how to keep electronic medical records accessible but private. Bmj 2001;322(7281):283-7.
17. Hodge JG, Jr., Gostin LO, Jacobson PD. Legal issues concerning electronic health information: privacy, quality, and liability. Jama 1999;282(15):1466-71.
18. Rethans JJ, Hoppener P, Wolfs G, Diederiks J. Do personal computers make doctors less personal? Br Med J (Clin Res Ed) 1988;296(6634):1446-8.
19. American Academy of Paediatrics: Task Force on Medical Informatics. Special requirements for electronic medical record systems in paediatrics. Paediatrics 2001;108(2):513-5.

Chapter 5

Summary

5.1 Summary

Children with fever (temperature ≥38°C) form a frequent diagnostic and therapeutic dilemma for clinicians. The majority of the children with fever have a viral infection. However, bacterial infections such as meningitis or sepsis are reported in 3% to 20% (1-4). Non-specific signs and symptoms in the young febrile child complicate an accurate estimation of the risk of serious bacterial infections and careful evaluation is necessary. At each step in this diagnostic and therapeutic evaluation the clinician has to outweigh risks and benefits and decide on the type of diagnostic tests, (antibiotic) treatment, and on type of care (in-hospital- or ambulatory care). The final diagnosis of a serious bacterial infection rests on a positive culture, which takes 12 to 24 hours or longer. Therefore, in most cases the decision to treat or not to treat has to be preliminarily, and any choice of antibiotics will be empirically. Several guidelines have been proposed, in which early discrimination of patients with and without a serious bacterial infection was aimed for (3-16). However, these recommendations were limited by 1) poor validation on new patients (15, 17); 2) development before the introduction of the Haemophilus influenzae type b (Hib) vaccine, disputing their current value (4); 3) cause of overtesting, overtreatment and rise of antibiotic resistance (16); 4) selection of patients on the diagnosis instead of on presenting problems, potentially leading to biased results (18); or 5) application of univariable rather than multivariable analysis (19).

The aim of this thesis was to assess whether the management of patients presenting with fever without apparent source could be optimised by diagnostic research. In particular with regard to the early segregation of the patients with serious bacterial infections from those not afflicted, using readily obtainable parameters. In addition, selected aspects of conducting diagnostic research were considered.

All studies in this thesis were performed on patients aged 1 month to 36 months. Children initially presented at the emergency department for the evaluation of acute fever without apparent source (including suspected sepsis). Retrospective identification of patients with fever without apparent source could only be achieved in the Sophia Children's University Hospital in Rotterdam (1988-1997). From 1998 onwards, patients from both the Sophia Children's University Hospital in Rotterdam and the Juliana Children's Hospital in The Hague were enrolled. Both hospitals are

large inner-city teaching hospitals in the Netherlands. Patient data were rettieved by means of a problem-oriented patient classification system, in which the main reason for encounter, according to the evaluation of the general practitioner or case history undertaken by the paediatrician, is classified (20, 21). Fever without apparent source was defined as a body temperature of at least thirty-eight degrees Celsius for which no clear focus was found after evaluation by the general practitioner, or after history taking by the paediatrician, i.e. before performance of the physical examination. This specific definition was used, as we aimed to distinguish patients with and without serious bacterial infections as early as possible in the diagnostic work-up, preferably only by means of taking a case history and by physical examination. Patients referred from other hospitals or with immune deficiencies were excluded.

Data were collected by reviewing the standardised medical records and included patients' general characteristics, and information on the patient's history, physical examination, diagnostic tests, treatment, and follow-up. Laboratory and radiographic test results were retrieved from the computer-documented hospital information system.

Outcome diagnosis was the presence or absence of a serious bacterial infection. A serious bacterial infection was defined based on the most consistent literature as the presence of bacterial meningitis, sepsis or bacteraemia, pneumonia, urinary tract infection, bacterial gastroenteritis, osteomyelitis or ethmoiditis (1, 2, 10, 11).

This study was conducted as part of a large ongoing study on paediatric diagnostic management and was approved by The Institutional Review Boards of both participating hospitals.

In chapter 2.1 two case histories illustrate the diagnostic and therapeutic dilemmas in young patients presenting with fever without apparent source. Previously described policy strategies for primary and secondary healthcare settings are discussed per age group in the light of current insights in the 'post-*Haemophilus influenzae* type b (Hib) era' (22). History taking and physical examination will continue to remain of first importance in discrimination of the patients with and without a serious bacterial infection. Readily obtainable and available laboratory tests will help with the additional selection of children at risk for invasive infection. For children aged 3 to 36 months, a more expectative strategy including less referral to paediatric care is suggested. In this respect, careful evaluation, clinical acumen, and time as a diagnostic tool all are key factors.

During the last decade, the management of febrile children was influenced by the introduction of the Hib-vaccine, the realisation of the growing emergence of antibiotic resistance, and the establishment of guidelines (3, 4, 6-16, 23-25). In chapter 2.2 trends in the management of referred children presenting at the emergency department with fever without apparent source are described. To this aim, patient data from two historical periods (1988-1992 and 1996-1998) were compared.

The prevalence of serious bacterial infections did not alter (21% versus 25%). The ratios of admitted patients, patients followed up on ambulatory basis, and of patients discharged after the initial consultation remained stable over time (approximately 1:2:1). Measurements of serum C-reactive protein (CRP) increased importantly towards 87%, while cultures of cerebrospinal fluid, blood or faeces were obtained significantly less often. We observed a significant diminishing in the antibiotic treatments. The admission rates remained stable at 26%, although the mean length of hospitalisation decreased considerably. The mean associated costs of the management per patient reduced valuably, mounting to a 46% reduction in the children eventually suffering from a serious bacterial infection. These figures confirm a trend towards a less invasive and more selective management in children with fever without apparent source.

Chapter 3 handles diagnosis in young children with fever without apparent source, aiming to timely discriminate between patients with and without serious bacterial infections. In chapter 3.1 a clinical rule was obtained to predict the presence of a serious bacterial infection in patients with fever without apparent source. The rule was developed on data from 231 patients, all referred by a general practitioner between 1996 and 1998. Using multivariable logistic regression and bootstrap resampling techniques, predictors from patient history and physical examination were selected. These included:

1. Duration of fever at presentation
2. Poor micturition
3. Vomiting
4. Age
5. Temperature $<36.7°C$ or $\geq 40°C$ at examination
6. Chest-wall retractions
7. Poor peripheral circulation

The discriminating performance (area under the receiver operating characteristic curve (95% confidence interval): ROC area (95%CI))) for the presence of a serious bacterial infection of this 'clinical model' was 0.75 (0.68-0.83). Strong predictors from laboratory tests with added diagnostic value were:

8. White blood cell count
9. Serum C-reactive protein
10. Presence of ≥70 white blood cells in urinalysis

The ROC area (95%CI) of the rule including all 10 predictors was 0.83 (0.77-0.89).

These two final models were then transformed to obtain a readily applicable diagnostic scoring rule. For each individual patient a 'clinical score' and a 'lab score' were estimated by assigning corresponding points for each variable present. With this rule, patients could be classified into categories of increased risks of having a serious bacterial infection. The risk stratification points out the value of clinical screening (range: 10% to 67%), which can be further modified by adding laboratory results (range: 6% to 92%).

Bypassing primary care by self-referral of particularly young children to paediatric care is on the increase, even up to 75% (26-30). In chapter 3.2, the important issue of discrimination of non-urgent patients and urgent patients early in presentation in these self-referred patients is dealt with. At first, it was assessed whether the diagnostic prediction rule constructed on referred patients (chapter 3.1) could be extended to a self-referred population (1997-1998, n=109). When applying the 'referred patients model' to the self-referred patients, the ROC area (95%CI) dropped from the original 0.75 (0.68-0.83) in the referred patients to 0.60 (0.47-0.73) in the self-referred children. Because of this poor discriminative ability, we concluded that the rule could not be generalised to the self-referred patients. This confirms that the type of patients is effected by the referral pattern (18). Referred patients already went through the diagnostic work-up of the general practitioner before presentation at paediatric care. The general practitioner only refers potentially urgent patients to paediatric care. In the self-referred group this shift by the general practitioner is absent, resulting in a more heterogeneous group visiting the paediatrician. Therefore, referred patients and self-referred patients should be considered as different patient populations.

Then, a specific prediction rule was developed for the self-referred patients. The 'clinical rule' comprised the following predictors (ROC area (95%CI): 0.70 (0.58-0.82)):

1. Duration of fever at presentation
2. Age
3. Degree of body temperature

After adding the univariably significant laboratory variables, the predictors containing additional diagnostic information were:

4. White blood cell count
5. Serum C-reactive protein

The ROC area (95%CI) increased significantly to 0.81 (0.71-0.91). The risk stratification ranged from 6% to 32% based on the 'clinical rule' only, and from 4% to 83% after additional laboratory evaluation.

In the study described in chapter 3.3, the diagnostic prediction rule for referred patients (chapter 3.1) was externally validated, and the predictive ability of the rule was improved. One hundred and fifty referred patients were prospectively enrolled between 2000 and 2001. For each patient, the data on the variables in the prediction rule were prospectively documented by the treating (trainee) paediatrician using a scoring form.

The generalisability of the prediction rule appeared too poor (ROC area (95%CI): 0.60 (0.49-0.70)). Subsequently, the predictive ability of the rule and the risk stratification were adjusted and improved using all available data of the referred patients with fever without source, i.e. data from 1996-1998 and 2000-2001 (n=381). Independent predictors from history and examination were (ROC area (95%CI): 0.69 (0.63-0.75)):

1. Duration of fever at presentation
2. Vomiting
3. Ill clinical appearance
4. Chest-wall retractions
5. Poor peripheral circulation

Recorded patient data need to serve several purposes (diagnosis, reporting, research, and policy). Often, if using paper charts, the same efforts must be repeated all over again. Structured data entry in a computer-based patient record (CPR) improves efficiency, in particular if recording data with a high variability, such as a patient's history and clinical examination. The main challenging difficulty in the achievement of structured data entry for such a comprehensive case history and physical examination is the controversy of structuring complex descriptions of patients' problems, by means of fixed components. At present, such CPR's that are actually applied in a general and broad specialism are sparsely available (54-57). In the study reported in chapter 4.3, we developed an expandable computer-based patient record with structured data entry, which is especially suitable for the registration of a comprehensive paediatric history and physical examination. This record can be incorporated with existing hospital information systems. ORCA-SDE was used, by which data entry is supported with a tree structure. Patient history is described by twenty main concepts, the physical examination by ten. In total, the tree consists of 8200 nodes and 1800 items, illustrating the degree of variability. Specific custom forms can be composed according to one's needs, however only with the previously defined concepts.

The CPR was exposed to a small number of paediatricians, to get a first impression of the evaluation of the users' perspectives (completeness, user-friendliness, and the usefulness). The opinions were rather positive.

Based on these first results, we conclude that implementation of this CPR might be feasible and that the many potential benefits of using a CPR outweigh the drawbacks.

Further special requirements for the developed CPR are worked on. Among future plans for applications in the paediatric CPR are the embedding of age-based normal ranges of measurements for vital signs and other physiological parameters, clinical decision rules, and graphic display of growth data and special calculations of growth patterns. Further prospects are incorporation of reminder systems (e.g. for immunisation), selection of patient subgroups (e.g. for prevention strategies) and support of adequate drug dosing.

References

1. Baker MD, Bell LM, Avner JR. Outpatient management without antibiotics of fever in selected infants. N Engl J Med 1993;329(20):1437-41.
2. Baraff LJ, Oslund SA, Schriger DL, Stephen ML. Probability of bacterial infections in febrile infants less than three months of age: a meta-analysis. Pediatr Infect Dis J 1992;11(4):257-64.
3. Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteraemia in young febrile children. Ann Emerg Med 1998;31(6):679-87.
4. Baraff LJ. Management of fever without source in infants and children. Ann Emerg Med 2000;36(6):602-14.
5. McCarthy PL, Lembo RM, Baron MA, Fink HD, Cicchetti DV. Predictive value of abnormal physical examination findings in ill-appearing and well-appearing febrile children. Paediatrics 1985;76(2):167-71.
6. McCarthy PL, Lembo RM, Fink HD, Baron MA, Cicchetti DV. Observation, history, and physical examination in diagnosis of serious illnesses in febrile children less than or equal to 24 months. J Pediatr 1987;110(1):26-30.
7. Kramer MS, Lane DA, Mills EL. Should blood cultures be obtained in the evaluation of young febrile children without evident focus of bacterial infection? A decision analysis of diagnostic management strategies. Paediatrics 1989;84(1):18-27.
8. Lieu TA, Schwartz JS, Jaffe DM, Fleisher GR. Strategies for diagnosis and treatment of children at risk for occult bacteraemia: clinical effectiveness and cost-effectiveness. J Pediatr 1991;118(1):21-9.
9. Downs SM, McNutt RA, Margolis PA. Management of infants at risk for occult bacteraemia: a decision analysis. Pediatr 1991;118(1):11-20.
10. Baraff LJ, Bass JW, Fleisher GR, et al. Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Paediatrics 1993;92:1-12.
11. Dagan R, Powell K, Hall C, Menegus M. Identification of infants unlikely to have serious bacterial infection although hospitalised for suspected sepsis. J Pediatr 1985;107:855-860.
12. Bonadio WA. Assessing patient clinical appearance in the evaluation of the febrile child. Am J Emerg Med 1995;13(3):321-6.
13. Isaacman DJ, Shults J, Gross TK, Davis PH, Harper M. Predictors of bacteraemia in febrile children 3 to 36 months of age. Paediatrics 2000;106(5):977-82.
14. Lee GM, Harper MB. Risk of bacteraemia for febrile young children in the post-Haemophilus influenzae type b era. Arch Pediatr Adolesc Med 1998;152(July):624-8.
15. Baker MD, Avner JR, Bell LM. Failure of infant observation scales in detecting serious illness in febrile, 4- to 8-week-old infants. Paediatrics 1990;85(6):1040-3.
16. Kramer MS, Shapiro ED. Management of the young febrile child: a commentary on recent practice guidelines. Paediatrics 1997;100:128-34.

17. Teach SJ, Fleisher GR. Efficacy of an observation scale in detecting bacteraemia in febrile children three to thirty-six months of age, treated as outpatients. Occult Bacteraemia Study Group. J Pediatr 1995;126(6):877-81.
18. Knottnerus JA, Leffers JP. The influence of referral patterns on the characteristics of diagnostic tests. J Clin Epidemiol 1992;45:1143-1154.
19. Moons KG, Es GA van, Michel BC, Buller HR, Habbema JD, Grobbee DE. Redundancy of single diagnostic test evaluation. Epidemiology 1999;10(3):276-81.
20. Derksen-Lubsen G, Jongkind CJ, Kraayenoord S, et al. [A problem-oriented patient classification system for general paediatrics I]. Tijdschr Kindergeneeskd 1996;64(3):93-8.
21. Steensel-Moll HA van, Jongkind CJ, Aarsen RSR, et al. [A problem-oriented patient classification system for general paediatrics II]. Tijdschr Kindergeneesk 1996;64(3):99-104.
22. Oostenbrink R, Groot R de, Moll HA. [Fever of unknown origin in a young child: diagnosis and treatment]. Ned Tijdschr Geneeskd 1999;143(4):185-90.
23. Neu HC. The crisis in antibiotic resistance. Science 1992;257(5073):1064-73.
24. Friedland IR, McCracken GH, Jr. Management of infections caused by antibiotic-resistant Streptococcus pneumoniae. N Engl J Med 1994;331(6):377-82.
25. Kaplan SL, Mason EO, Jr., Barson WJ, et al. Three-year multicenter surveillance of systemic pneumococcal infections in children. Paediatrics 1998;102(3 Pt 1):538-45.
26. Sharma V, Simon SD, Bakewell JM, Ellerbeck EF, Fox MH, Wallace DD. Factors influencing infant visits to emergency departments. Paediatrics 2000;106(5):1031-9.
27. Bedford HE, Jenkins SM, Shore C, Kenny PA. Use of an east end children's accident and emergency department for infants: a failure of primary health care? Qual Health Care 1992;1(1):29-33.
28. Rothschild M, Gilboa S, Sagi H, Berger I, Wolach B. Referral, admission, and discharge patterns in a paediatric emergency department in Israel. Pediatr Emerg Care 1993;9(2):72-6.
29. Fieselmann J, Hendryx M. Characteristics of self-reffered patients. Academic Medicine 1994;69(5):376-81.
30. Beebe SA, Casey R, Magnusson MR, Pasquariello PS, Jr. Comparison of self-referred and physician-referred patients to a paediatric diagnostic center. Clin Pediatr (Phila) 1993;32(7):412-6.
31. Harrell FE, Lee KL, Matchar DB, Reichert TA. Regression models for prognostic prediction: Advantages, problems, and suggested solutions. Cancer Treat Rep 1985;69:1071-7.
32. Steyerberg EW, Eijkemans MJ, Habbema JD. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. J Clin Epidemiol 1999;52(10):935-42.
33. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med 2000;19(8):1059-79.

34. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models. Efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54(8):774-81.
35. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine 1996;15(4):361-87.
36. Justice AC, Covinsky KE, Berlin JA. Assessing the generalisability of prognostic information. Ann Intern Med 1999;130(6):515-24.
37. Knottnerus JA. Prediction rules: statistical reproducibility and clinical similarity. Med Decis making 1992;12:286-287.
38. Laupacis A, Wells G, Richardson WS, Tugwell P. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. Jama 1994;272(3):234-7.
39. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. Jama 2000;284(1):79-84.
40. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. New England Journal of Medicine 1985;313(13):793-9.
41. 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 1994;271(5):389-91.
42. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. Jama 1994;271(9):703-7.
43. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. Jama 1997;277(6):488-94.
44. Begg CB. Biases in the assessment of diagnostic tests. Stat Med 1987;6(4):411-23.
45. Knottnerus JA. Application of logistic regression to the analysis of diagnostic data: exact modeling of a probability tree of multiple binary variables. Med Decis Making 1992;12:93-108.
46. Schouw YT van der, Dijk R van, Verbeek AL. Problems in selecting the adequate patient population from existing data files for assessment studies of new diagnostic tests. J Clin Epidemiol 1995;48(3):417-22.
47. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Ass 1983;78:316-31.
48. Houwelingen JC van, Le Cessie S. Predictive value of statistical models. Stat Med 1990;9(11):1303-25.
49. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19(4):453-73.

50. Efron B, Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. New York: Chapman & Hall; 1993.
51. Efron B, Tibshirani R. Improvements on cross-validation: the .632+ bootstrap method. JASA 1997;92:548-60.
52. Picard RR, Berk KN. Data splitting. Am Statistician 1990;44:140-7.
53. Sauerbrei W. The use of resampling methods to simplify regression models in medical statistics. J R Stat Soc C Appl Stat 1999;48:313-329.
54. Gouveia-Oliveira A, Salgado NC, Azevedo AP, et al. A unified approach to the design of clinical reporting systems. Methods Inf Med 1994;33(5):479-87.
55. Bell DS, Greenes RA. Evaluation of UltraSTAR: performance of a collaborative structured data entry system. Proc Annu Symp Comput Appl Med Care 1994:216-22.
56. Meijden MJ van der, Tange HJ, Boiten J, Troost J, Hasman A. An experimental electronic patient record for stroke patients. Part 2: system description. Int J Med Inf 2000;58-59:127-40.
57. Kirby J, Rector AL. The PEN&PAD data entry system: from prototype to practical system. Proc AMIA Annu Fall Symp 1996:709-13.

5.2 Samenvatting

Koorts (temperatuur ≥38°C) bij kinderen is vaak een diagnostisch en therapeutisch dilemma. Het merendeel van de kinderen met koorts heeft een virale infectie. In 3% tot 20% van deze kinderen wordt echter een bacteriële infectie, zoals bijvoorbeeld een meningitis of sepsis, vastgesteld (1-4). De aspecifieke klachtenpresentatie van het jonge kind met koorts maakt de inschatting van het risico op een ernstige bacteriële infectie nog moeilijker en dientengevolge is zorgvuldige evaluatie nodig.

In een dergelijke diagnostische en therapeutische evaluatie moeten de voor- en nadelen steeds tegen elkaar afgewogen worden. De arts moet besluiten nemen ten aanzien van de eventuele aanvullende diagnostiek, de (antibiotische) behandeling, en het type zorg (klinische of poliklinische zorg). De uiteindelijke diagnose van een ernstige bacteriële infectie berust op een positieve kweek, waarvan de uitslag na 12 à 24 uur of later bekend is. Derhalve is het besluit om wel of niet antibiotisch te behandelen empirisch, evenals de keuze van het antibioticum.

Verscheidene richtlijnen, met als doel vroegtijdig een onderscheid te maken in patiënten met en zonder een ernstige bacteriële infectie, zijn voorgesteld (3-16). Deze richtlijnen bleken echter beperkte waarde te hebben, namelijk: 1) toepassing van de richtlijn in nieuwe patiëntenpopulaties leverde matige resultaten op (15, 17); 2) de ontwikkeling van de desbetreffende richtlijn geschiedde vóór de introductie van de *Haemophilus influenzae* type b (Hib) vaccinatie, waardoor de huidige waarde in twijfel getrokken wordt (4); 3) de desbetreffende richtlijn leidt tot overmatige diagnostiek, behandeling en toename van de antibiotica resistentie (16); 4) geïncludeerde patiënten waren geselecteerd op einddiagnose in plaats van op het probleem van presentatie, mogelijk leidend tot vertekende resultaten (18); of 5) in plaats van multivariate analysen werden univariate analysen gebruikt (19).

Het doel van dit proefschrift was om vast te stellen of het beleid van kinderen die zich presenteren met koorts zonder duidelijk focus geoptimaliseerd kon worden met diagnostisch onderzoek. Hierbij lag het accent op het vroegtijdig onderscheiden van de patiënten met en zonder een ernstige bacteriële infectie, op basis van makkelijk te verkrijgen parameters. Daarnaast werd een aantal aspecten van de praktische uitvoering van diagnostisch onderzoek uiteengezet.

Alle studies beschreven in dit proefschrift zijn gebaseerd op gegevens van patiënten die 1 tot 36 maanden oud waren. De kinderen presenteerden zich initieel op de Acute Hulp voor de evaluatie van acute koorts zonder duidelijk focus (inclusief verdenking van sepsis). Patiënten met koorts zonder duidelijk focus konden alleen retrospectief in het Sophia Kinderziekenhuis in Rotterdam (1988-1997) achterhaald worden. Vanaf 1998 werden patiënten geïncludeerd uit zowel het Sophia Kinderziekenhuis in Rotterdam als het Juliana Kinderziekenhuis in Den Haag. Beide ziekenhuizen zijn opleidingsklinieken in Nederland. Patiëntgegevens werden achterhaald door middel van een probleemgeoriënteerd patiëntenclassificatiesysteem, waarin de hoofdklacht van presentatie, gebaseerd op de evaluatie van de huisarts of de anamnese van de kinderarts, wordt geclassificeerd (20, 21). Koorts zonder focus was gedefinieerd als een lichaamstemperatuur van achtendertig graden Celsius of hoger, waarvoor geen duidelijk focus werd gevonden na evaluatie door de huisarts of na anamnese door de kinderarts. Juist deze definitie werd gebruikt, aangezien we zo vroeg mogelijk in het diagnostische traject een schifting van de patiënten met en zonder ernstige bacteriële infecties nastreefden, bij voorkeur reeds op basis van anamnestische gegevens en bevindingen bij lichamelijk onderzoek. Patiënten die waren doorverwezen vanuit andere ziekenhuizen of die een immuundeficiëntie hadden werden geëxcludeerd.

Gegevens werden verzameld uit gestandaardiseerde medische dossiers en omvatte algemene patiëntkarakteristieken, anamnestische gegevens, bevindingen bij lichamelijk onderzoek en informatie over diagnostiek, behandeling en follow-up. Laboratoriumuitslagen en verslagen van beeldvormend onderzoek werden uit het Ziekenhuis Informatie Systeem gehaald.

De uitkomst ernstige bacteriële infectie was conform de literatuur gedefinieerd als de aanwezigheid van een bacteriële meningitis, sepsis, bacteriemie pneumonie, urineweginfectie, bacteriële gastro-enteritis, osteomyelitis of ethmoiditis (1, 2, 10, 11).

De studies in dit proefschrift werden met toestemming van de Medische Ethische Commissies van beide deelnemende ziekenhuizen uitgevoerd als onderdeel van het Ontwikkelingsgeneeskunde project 'Diagnostiek in de kindergeneeskunde op routinematig verzamelde gegevens'.

In hoofdstuk 2.1 worden de diagnostische en therapeutische dilemma's bij jonge kinderen met koorts zonder focus door twee ziektegeschiedenissen geïllustreerd. Per leeftijdsgroep worden eerder beschreven beleidsstrategieën voor de eerste- en

tweedelijnszorg in het kader van het huidige 'post-Hib tijdperk' bediscussieerd (22). Anamnese en lichamelijk onderzoek blijven van groot belang om de eerste schifting tussen kinderen met en zonder een ernstige bacteriële infectie te maken. Eenvoudig laboratoriumonderzoek draagt bij aan een tweede schifting. Voor kinderen van 3 tot 36 maanden oud wordt een meer expectatief beleid geadviseerd, waarbij minder snel naar de tweedelijnszorg verwezen behoeft te worden. Hierbij zijn zorgvuldige evaluatie, klinische ervaring, goede voorlichting aan ouders en de tijd als diagnosticum belangrijke factoren.

Gedurende de laatste tien jaar werd het beleid bij koortsende kinderen onder meer beïnvloed door de introductie van de Hib-vaccinatie, het besef van de toenemende ernst van antibioticaresistentie, en de introductie van richtlijnen (3, 4, 6-16, 23-25). In hoofdstuk 2.2 worden trends in het beleid van kinderen met koorts zonder duidelijk focus die door de huisarts naar de Acute Hulp verwezen zijn beschreven. Hiervoor werden patiëntgegevens uit 1988-1992 vergeleken met die uit 1996-1998.

Het percentage ernstige bacteriële infecties veranderde niet (21% versus 25%). De verhouding van het aantal opgenomen patiënten, patiënten met poliklinisch vervolg en patiënten die na het initiële consult werden ontslagen bleef gelijk gedurende de tijd (ongeveer 1:2:1). Het aantal C-reactief proteïne (CRP) metingen in het bloed steeg in belangrijke mate tot 87%, terwijl het aantal verrichte kweken van liquor, bloed of faeces significant afnam. We zagen een significante daling in het aantal antibiotische behandelingen. Het percentage opgenomen patiënten bleef 26%, alhoewel de gemiddelde opnameduur sterk verkortte. De gemiddelde kosten per patiënt ten gevolge van het beleid daalden aanzienlijk, oplopend tot een vermindering van 46% in de groep kinderen die uiteindelijk een ernstige bacteriële infectie bleek te hebben. Deze getallen bevestigen een trend in de richting van een minder invasief en meer selectief beleid bij kinderen met koorts zonder duidelijk focus.

Hoofdstuk 3 behandelt de diagnostiek bij kinderen met koorts zonder duidelijk focus, met als oogmerk patiënten met en zonder ernstige bacteriële infecties vroegtijdig te herkennen. In hoofdstuk 3.1 werd een klinische beslisregel ontwikkeld om bij patiënten met koorts zonder duidelijk focus de aanwezigheid van een ernstige bacteriële infectie te voorspellen. De regel was ontworpen op data van 231 patiënten, welke allen tussen 1996 en 1998 verwezen waren door de huisarts. Met behulp van

multivariate logistische regressie en bootstrap resampling technieken werden predictoren uit de anamnese en het lichamelijk onderzoek geselecteerd. Deze omvatten:

1. Duur van de koorts bij presentatie
2. Verminderde mictie
3. Braken
4. Leeftijd
5. Temperatuur <36.7°C of ≥40°C bij lichamelijk onderzoek
6. Intrekkingen ± tachypneu
7. Verminderde perifere circulatie

Van dit 'klinische model' was het discriminerend vermogen (oppervlakte onder de receiver operating characteristic curve (95% betrouwbaarheidsinterval): ROC oppervlakte (95%BI))) voor de aanwezigheid van een ernstige bacteriële infectie 0.75 (0.68-0.83). Sterke voorspellers met toevoegende diagnostische waarde uit het laboratoriumonderzoek waren:

8. Leukocytengetal in bloed
9. C-reactieve proteïne gehalte in bloed
10. Aanwezigheid van ≥70 leukocyten bij urine dipstickanalyse

De ROC oppervlakte (95%BI) van de regel inclusief alle 10 predictoren was 0.83 (0.77-0.89).

Deze twee eindmodellen werden omgezet naar makkelijk toepasbare diagnostische scoreregels. Door het toekennen van overeenkomende punten voor elke (aanwezige) variabele werd voor elke individuele patiënt een 'klinische score' en een 'lab score' berekend. Met behulp van deze regel konden patiënten geclassificeerd worden in categorieën van toenemend risico op de aanwezigheid van een ernstige bacteriële infectie. Deze risicostratificatie benadrukt de waarde van de klinische screening (range: 10% tot 67%), welke verder kan worden gespecificeerd door aanvullende laboratoriumonderzoekuitslagen (range: 6% tot 92%).

Het passeren van de eerstelijnszorg door zelfverwijzing naar de tweedelijnszorg van met name jonge kinderen neemt toe, zelfs tot 75% (26-30). In hoofdstuk 3.2 wordt het onderwerp van het vroegtijdig in de presentatie onderscheiden van patiënten met en zonder een ernstige bacteriële infectie in dergelijke zelfverwezen

patiënten belicht. Allereerst werd onderzocht of de diagnostische beslisregel die ontworpen was voor de verwezen patiënten (hoofdstuk 3.1) uitgebreid kon worden naar de niet-verwezen populatie (1997-1998, n=109). Bij toepassing van het 'verwezen patiënten model' op de zelfverwezen patiënten daalde de ROC oppervlakte (95%BI) van de oorspronkelijke 0.75 (0.68-0.83) in de verwezen patiënten naar 0.60 (0.47-0.73) in de niet-verwezen kinderen. Op basis van dit matig onderscheidend vermogen, concludeerden wij dat de beslisregel niet gegeneraliseerd kon worden naar de zelfverwezen patiënten. Dit bevestigt dat het type patiënten beïnvloed wordt door het verwijspatroon (18). Verwezen patiënten hebben reeds vóór presentatie in de tweedelijnszorg een diagnostisch work-up van de huisarts doorlopen. De huisarts verwijst alleen potentieel urgente patiënten door naar de kindergeneeskundige zorg. In de groep niet-verwezen patiënten ontbreekt deze schifting door de huisarts, resulterend in een meer gemêleerde groep. Derhalve moeten verwezen kinderen en niet-verwezen kinderen als verschillende patiëntenpopulaties beschouwd worden.

Vervolgens werd een beslisregel ontwikkeld specifiek voor de zelfverwezen patiënten. De 'klinische regel' bevatte de volgende voorspellers (ROC oppervlakte (95%BI): 0.70 (0.58-0.82)):

1. Duur van de koorts bij presentatie
2. Leeftijd
3. Hoogte van de lichaamstemperatuur

Na toevoeging van univariate significante laboratoriumvariabelen, bleken deze predictoren toevoegende diagnostische waarde te hebben:

4. Leukocytengetal in bloed
5. C-reactieve proteïne gehalte in bloed

De ROC oppervlakte (95%BI) nam significant toe tot 0.81 (0.71-0.91). De risicostratificatie liep van 6% tot 32% op basis van alleen de 'klinische regel' en van 4% tot 83% na aanvullende laboratoriumevaluatie.

In de studie beschreven in hoofdstuk 3.3 werd de diagnostische beslisregel voor verwezen patiënten (hoofdstuk 3.1) extern gevalideerd en het voorspellend vermogen van de regel werd verbeterd. Honderd en vijftig verwezen patiënten werden prospectief geïncludeerd tussen 2000 en 2001. Van elke patiënt werd informatie over

de variabelen in de beslisregel prospectief gedocumenteerd door de behandelend kinderarts (in opleiding) met behulp van een scoreformulier.

De generaliseerbaarheid van de beslisregel bleek te matig te zijn (ROC oppervlakte (95%BI): 0.60 (0.49-0.70)). Vervolgens werden het voorspellend vermogen van de regel en de risicostratificatie aangepast en verbeterd, gebruikmakend van alle beschikbare gegevens van de verwezen patiënten met koorts zonder duidelijk focus, d.w.z. data uit 1996-1998 en 2000-2001 (n=381). Onafhankelijke voorspellers uit anamnese en lichamelijk onderzoek waren (ROC oppervlakte (95%BI): 0.69 (0.63-0.75)):

1. Duur van de koorts bij presentatie
2. Braken
3. Zieke klinische indruk
4. Intrekkingen
5. Verminderde perifere circulatie

Wederom waren voorspellers uit laboratoriumonderzoek met aanvullende diagnostische waarde (ROC oppervlakte (95%BI): 0.83 (0.78-0.88)):

6. Leukocytengetal in bloed
7. C-reactieve proteïne gehalte in bloed
8. Aanwezigheid van ≥70 leukocyten bij urine dipstickanalyse

Alle predictoren van deze 'geoptimaliseerde regel', behalve zieke klinische indruk, kwamen overeen met de voorspellers in de beslisregel zoals beschreven in hoofdstuk 3.1. Alhoewel het relatieve gewicht van elke predictor nu anders was, suggereert het weer terugkomen van deze predictoren in de selectie dat het stabiele voorspellers zijn. Daarentegen waren de voorspellers die initieel wel geselecteerd waren maar bij de heranalysen buiten de selectie vielen (verminderde mictie, leeftijd boven 1 jaar en lichaamstemperatuur) waarschijnlijk juist instabiel. Ondanks toepassing van bootstrap technieken was het initieel afgeleide model zeer waarschijnlijk overfit, aangezien we stapsgewijze selectiemethoden in een relatief klein sample gebruikten om een model te definiëren (31-34).

De herziene risicostratificatie reikte van 4% tot 54%. Wederom laat de risicostratificatie zien dat informatie uit anamnese en lichamelijk onderzoek waardevol is als een eerste screeningsinstrument en dat aanvullend laboratoriumonderzoek het individuele risico kan toespitsen. Tengevolge van de heranalysen moet de

generaliseerbaarheid van de herziene regel nog geëvalueerd worden in een nieuwe externe validatie procedure alvorens de regel in praktijk toe te passen (35-40).

De drie beslisregels zoals beschreven in hoofdstukken 3.1, 3.2 en 3.3 bevatten overlappende voorspellers. Waarschijnlijk is dit een gevolg van de heterogeniteit van de uitkomst ernstige bacteriële infectie bij kinderen met koorts zonder duidelijk focus, die inherent is aan de niet-specifieke presentatie van deze kinderen. Derhalve werden meer algemene dan specifieke voorspellers geselecteerd, met waardevolle informatie, die alle specifieke diagnoses van de uitkomst ernstige bacteriële infectie overkoepelt. Heterogeniteit in de uitkomst bemoeilijkt het zuiver scheiden van patiënten met en zonder de betreffende uitkomst, wat geïllustreerd werd door het imperfecte discriminatievermogen van alle drie de scoreregels. Desalniettemin verschaften deze enkele klinische waarde, aangezien de *'post test'* kansen verschilden van de *'a priori'* kansen (41-43).

In het algemeen blijft het beleid bij jonge kinderen met koorts zonder duidelijk focus een moeilijk probleem, omdat het risico niet geheel naar nul gereduceerd kan worden. De beschreven diagnostische beslisregels moeten gebruikt worden als extra hulpmiddel naast de klinische expertise van de behandelend arts.

In hoofdstuk 4 worden enkele aspecten van de methoden om (diagnostisch) onderzoek uit te voeren uiteengezet. In de praktijk is het stellen van een diagnose een gefaseerd proces, dat begint met een bepaalde patiënt met een bepaalde klacht of een symptoom, verdacht voor een bepaalde ziekte. Er bestaat discussie over de betrouwbaarheid van de resultaten dat diagnostisch onderzoek, gebruikmakend van gegevens van routinematige patiëntenzorg, oplevert (44-46). Hoofdstuk 4.1 weidt uit over de voor- en nadelen van dergelijk diagnostisch onderzoek op basis van routinematige verzamelde gegevens en over hoe om te kunnen gaan met bijkomende problemen. Vereiste selectie van patiënten op hun klinische presentatie in plaats van op hun uiteindelijke diagnose wordt beschreven. Voor een dergelijke selectie op het klinische probleem is een probleemgeoriënteerd patiëntenclassificatiesysteem nodig (20, 21). Tevens worden het vóórkomen van missende gegevens in routinematige verzamelde data en verificatie bias en blinderen in diagnostisch onderzoek belicht.

Hoofdstuk 4.2 illustreert de beperkingen van interne validatie om de generaliseerbaarheid van een diagnostisch model naar toekomstige situaties in te

schatten. Het discriminerend vermogen van predictiemodellen neigt te verminderen op nieuwe patiëntgegevens (validatie set) ten opzichte van de patiëntgegevens waarop het model ontwikkeld was (derivatie set) (19, 31, 35, 36, 39, 40, 43, 47-49), met name in kleine datasets (32, 33). Bootstrap resampling technieken zijn derhalve aangedragen om de regressiecoëfficiënten aan te passen voor optimisme (interne validatie) (34, 35, 47, 50-53). Resultaten worden veelal geaccepteerd zonder dat er genoeg aandacht is geschonken aan de externe validiteit, waarmee de accuratesse van het model in patiënten uit een andere maar nagenoeg overeenkomende populatie wordt belicht (36, 37, 39).

In dit hoofdstuk werden de resultaten van een interne en een externe validatie van een beslisregel voor kinderen met koorts zonder duidelijk focus vergeleken. Een beslisregel voor de aanwezigheid van een ernstige bacteriële infectie in kinderen met koorts zonder duidelijk focus was afgeleid en intern gevalideerd gebruikmakend van bootstrap resampling technieken. Vervolgens werd het model extern gevalideerd.
In de derivatie set (1988-1992, n=376) werden negen onafhankelijke voorspellers gedefinieerd. Het onderscheidend vermogen (ROC oppervlakte (95%BI)) van het model was 0.83 (0.78-0.87) en 0.76 (0.67-0.85) na bootstrapping. In de validatie set (1997-1998, n=179) daalde het onderscheidend vermogen naar 0.57 (0.47-0.67). Het model vertoonde een matige calibratie in de validatie set, met name in de 'laag risico' groepen.

Deze studie illustreert en bevestigt dat interne validiteit op zich geen garantie is voor generaliseerbaarheid van predictiemodellen en dus geen vervanging is voor externe validatie (36). Externe validatie is derhalve essentieel voor implementatie in de klinische praktijk, zeker als de modellen zijn afgeleid van relatieve kleine datasets.

Verzamelde patiëntgegevens dienen meerdere doeleinden (diagnose, verslaglegging, onderzoek en beleidsvorming). Vaak moet bij gebruik van papieren statussen dezelfde inspanning meerdere keren geleverd worden. Gestructureerde gegevensinvoer in een elektronisch medisch dossier (EMD) bevordert de efficiëntie, met name als het data met een hoge variabiliteit betreft, zoals gegevens uit anamnese en lichamelijk onderzoek. De grootste uitdaging in het bewerkstelligen van gestructureerde gegevensinvoer voor een uitgebreide anamnese en lichamelijk onderzoek ligt in de tegenstelling van het structureren van complexe beschrijvingen van patiëntenproblemen door middel van gefixeerde componenten. Tot op heden zijn dergelijke EMD's die daadwerkelijk in praktijk in een algemeen en breed specialisme

worden toegepast beperkt beschikbaar (54-57). In de studie weergegeven in hoofdstuk 4.3, hebben we een uitbreidbaar EMD met gestructureerde gegevensinvoer ontwikkeld, dat speciaal geschikt is voor het registreren van een uitgebreide en gedetailleerde kindergeneeskundige anamnese en lichamelijk onderzoek. Dit EMD kan geïncorporeerd worden met bestaande Ziekenhuis Informatie Systemen. ORCA-SDE werd gebruikt, waarbij gegevensinvoer wordt ondersteund met een boomstructuur. Anamnese wordt beschreven door twintig hoofdconcepten, het lichamelijk onderzoek door tien. In totaal bestaat de boom uit 8200 vertakkingen en 1800 items; dit illustreert de mate van variabiliteit. Specifieke formulieren kunnen naar ieders behoefte worden samengesteld, echter alleen op basis van de reeds gedefinieerde concepten.

Het EMD werd geïntroduceerd aan een klein aantal kinderartsen om een eerste indruk te verkrijgen van de gebruikersperspectieven (volledigheid, gebruikersvriendelijkheid en de bruikbaarheid). De meningen waren redelijk positief.

Gebaseerd op deze eerste resultaten concluderen we dat implementatie van dit EMD haalbaar kan zijn en dat de vele potentiële voordelen van het gebruik van een EMD zwaarder wegen dan de nadelen.

Aan verdere speciale vereisten aan het EMD wordt gewerkt. Enkele toekomstplannen voor applicaties in het kindergeneeskundig EMD omvatten het inbouwen van leeftijdsafhankelijke normaalwaarden voor metingen van vitale en andere fysiologische parameters, klinische beslisregels, grafische weergave van groeicurven en speciale berekeningen van groeipatronen. Verdere ideeën beslaan incorporatie van herinneringsystemen (b.v. voor vaccinaties), selectie van patiëntensubgroepen (b.v. voor preventiestrategieën) en ondersteuning van adequate geneesmiddelendosering.

Referenties

1. Baker MD, Bell LM, Avner JR. Outpatient management without antibiotics of fever in selected infants. N Engl J Med 1993;329(20):1437-41.
2. Baraff LJ, Oslund SA, Schriger DL, Stephen ML. Probability of bacterial infections in febrile infants less than three months of age: a meta-analysis. Pediatr Infect Dis J 1992;11(4):257-64.
3. Kuppermann N, Fleisher GR, Jaffe DM. Predictors of occult pneumococcal bacteraemia in young febrile children. Ann Emerg Med 1998;31(6):679-87.
4. Baraff LJ. Management of fever without source in infants and children. Ann Emerg Med 2000;36(6):602-14.

5. McCarthy PL, Lembo RM, Baron MA, Fink HD, Cicchetti DV. Predictive value of abnormal physical examination findings in ill-appearing and well-appearing febrile children. Paediatrics 1985;76(2):167-71.
6. McCarthy PL, Lembo RM, Fink HD, Baron MA, Cicchetti DV. Observation, history, and physical examination in diagnosis of serious illnesses in febrile children less than or equal to 24 months. J Pediatr 1987;110(1):26-30.
7. Kramer MS, Lane DA, Mills EL. Should blood cultures be obtained in the evaluation of young febrile children without evident focus of bacterial infection? A decision analysis of diagnostic management strategies. Paediatrics 1989;84(1):18-27.
8. Lieu TA, Schwartz JS, Jaffe DM, Fleisher GR. Strategies for diagnosis and treatment of children at risk for occult bacteraemia: clinical effectiveness and cost-effectiveness. J Pediatr 1991;118(1):21-9.
9. Downs SM, McNutt RA, Margolis PA. Management of infants at risk for occult bacteraemia: a decision analysis. Pediatr 1991;118(1):11-20.
10. Baraff LJ, Bass JW, Fleisher GR, et al. Practice guideline for the management of infants and children 0 to 36 months of age with fever without source. Paediatrics 1993;92:1-12.
11. Dagan R, Powell K, Hall C, Menegus M. Identification of infants unlikely to have serious bacterial infection although hospitalised for suspected sepsis. J Pediatr 1985;107:855-860.
12. Bonadio WA. Assessing patient clinical appearance in the evaluation of the febrile child. Am J Emerg Med 1995;13(3):321-6.
13. Isaacman DJ, Shults J, Gross TK, Davis PH, Harper M. Predictors of bacteraemia in febrile children 3 to 36 months of age. Paediatrics 2000;106(5):977-82.
14. Lee GM, Harper MB. Risk of bacteraemia for febrile young children in the post-Haemophilus influenzae type b era. Arch Pediatr Adolesc Med 1998;152(July):624-8.
15. Baker MD, Avner JR, Bell LM. Failure of infant observation scales in detecting serious illness in febrile, 4- to 8-week-old infants. Paediatrics 1990;85(6):1040-3.
16. Kramer MS, Shapiro ED. Management of the young febrile child: a commentary on recent practice guidelines. Paediatrics 1997;100:128-34.
17. Teach SJ, Fleisher GR. Efficacy of an observation scale in detecting bacteraemia in febrile children three to thirty-six months of age, treated as outpatients. Occult Bacteraemia Study Group. J Pediatr 1995;126(6):877-81.
18. Knottnerus JA, Leffers JP. The influence of referral patterns on the characteristics of diagnostic tests. J Clin Epidemiol 1992;45:1143-1154.
19. Moons KG, Es GA van, Michel BC, Buller HR, Habbema JD, Grobbee DE. Redundancy of single diagnostic test evaluation. Epidemiology 1999;10(3):276-81.
20. Derksen-Lubsen G, Jongkind CJ, Kraayenoord S, et al. [A problem-oriented patient classification system for general paediatrics I]. Tijdschr Kindergeneeskd 1996;64(3):93-8.
21. Steensel-Moll HA van, Jongkind CJ, Aarsen RSR, et al. [A problem-oriented patient classification system for general paediatrics II]. Tijdschr Kindergeneesk 1996;64(3):99-104.

22. Oostenbrink R, Groot R de, Moll HA. [Fever of unknown origin in a young child: diagnosis and treatment]. Ned Tijdschr Geneeskd 1999;143(4):185-90.
23. Neu HC. The crisis in antibiotic resistance. Science 1992;257(5073):1064-73.
24. Friedland IR, McCracken GH, Jr. Management of infections caused by antibiotic-resistant Streptococcus pneumoniae. N Engl J Med 1994;331(6):377-82.
25. Kaplan SL, Mason EO, Jr., Barson WJ, et al. Three-year multicenter surveillance of systemic pneumococcal infections in children. Paediatrics 1998;102(3 Pt 1):538-45.
26. Sharma V, Simon SD, Bakewell JM, Ellerbeck EF, Fox MH, Wallace DD. Factors influencing infant visits to emergency departments. Paediatrics 2000;106(5):1031-9.
27. Bedford HE, Jenkins SM, Shore C, Kenny PA. Use of an east end children's accident and emergency department for infants: a failure of primary health care? Qual Health Care 1992;1(1):29-33.
28. Rothschild M, Gilboa S, Sagi H, Berger I, Wolach B. Referral, admission, and discharge patterns in a paediatric emergency department in Israel. Pediatr Emerg Care 1993;9(2):72-6.
29. Fieselmann J, Hendryx M. Characteristics of self-reffered patients. Academic Medicine 1994;69(5):376-81.
30. Beebe SA, Casey R, Magnusson MR, Pasquariello PS, Jr. Comparison of self-referred and physician-referred patients to a paediatric diagnostic center. Clin Pediatr (Phila) 1993;32(7):412-6.
31. Harrell FE, Lee KL, Matchar DB, Reichert TA. Regression models for prognostic prediction: Advantages, problems, and suggested solutions. Cancer Treat Rep 1985;69:1071-7.
32. Steyerberg EW, Eijkemans MJ, Habbema JD. Stepwise selection in small data sets: a simulation study of bias in logistic regression analysis. J Clin Epidemiol 1999;52(10):935-42.
33. Steyerberg EW, Eijkemans MJ, Harrell FE, Jr., Habbema JD. Prognostic modelling with logistic regression analysis: a comparison of selection and estimation methods in small data sets. Stat Med 2000;19(8):1059-79.
34. Steyerberg EW, Harrell FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models. Efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54(8):774-81.
35. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Statistics in Medicine 1996;15(4):361-87.
36. Justice AC, Covinsky KE, Berlin JA. Assessing the generalisability of prognostic information. Ann Intern Med 1999;130(6):515-24.
37. Knottnerus JA. Prediction rules: statistical reproducibility and clinical similarity. Med Decis making 1992;12:286-287.

38. Laupacis A, Wells G, Richardson WS, Tugwell P. Users' guides to the medical literature. V. How to use an article about prognosis. Evidence-Based Medicine Working Group. Jama 1994;272(3):234-7.
39. McGinn TG, Guyatt GH, Wyer PC, Naylor CD, Stiell IG, Richardson WS. Users' guides to the medical literature: XXII: how to use articles about clinical decision rules. Evidence-Based Medicine Working Group. Jama 2000;284(1):79-84.
40. Wasson JH, Sox HC, Neff RK, Goldman L. Clinical prediction rules. Applications and methodological standards. New England Journal of Medicine 1985;313(13):793-9.
41. 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 1994;271(5):389-91.
42. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. Jama 1994;271(9):703-7.
43. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules. A review and suggested modifications of methodological standards. Jama 1997;277(6):488-94.
44. Begg CB. Biases in the assessment of diagnostic tests. Stat Med 1987;6(4):411-23.
45. Knottnerus JA. Application of logistic regression to the analysis of diagnostic data: exact modeling of a probability tree of multiple binary variables. Med Decis Making 1992;12:93-108.
46. Schouw YT van der, Dijk R van, Verbeek AL. Problems in selecting the adequate patient population from existing data files for assessment studies of new diagnostic tests. J Clin Epidemiol 1995;48(3):417-22.
47. Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. J Am Stat Ass 1983;78:316-31.
48. Houwelingen JC van, Le Cessie S. Predictive value of statistical models. Stat Med 1990;9(11):1303-25.
49. Altman DG, Royston P. What do we mean by validating a prognostic model? Stat Med 2000;19(4):453-73.
50. Efron B, Tibshirani R. An introduction to the bootstrap. Monographs on statistics and applied probability. New York: Chapman & Hall; 1993.
51. Efron B, Tibshirani R. Improvements on cross-validation: the .632+ bootstrap method. JASA 1997;92:548-60.
52. Picard RR, Berk KN. Data splitting. Am Statistician 1990;44:140-7.
53. Sauerbrei W. The use of resampling methods to simplify regression models in medical statistics. J R Stat Soc C Appl Stat 1999;48:313-329.
54. Gouveia-Oliveira A, Salgado NC, Azevedo AP, et al. A unified approach to the design of clinical reporting systems. Methods Inf Med 1994;33(5):479-87.

55. Bell DS, Greenes RA. Evaluation of UltraSTAR: performance of a collaborative structured data entry system. Proc Annu Symp Comput Appl Med Care 1994:216-22.
56. Meijden MJ van der, Tange HJ, Boiten J, Troost J, Hasman A. An experimental electronic patient record for stroke patients. Part 2: system description. Int J Med Inf 2000;58-59:127-40.
57. Kirby J, Rector AL. The PEN&PAD data entry system: from prototype to practical system. Proc AMIA Annu Fall Symp 1996:709-13.

List of co-authors

G Derksen-Lubsen, MD, PhD
Department of General Paediatrics
Juliana Children's Hospital
The Hague, The Netherlands

ART Donders, PhD
Julius Centre for General Practice and Patient Oriented Research/ Department of BioStatistics
University Medical Centre Utrecht
Utrecht, The Netherlands

AM van Ginneken, MD, PhD
Department of Medical Informatics
Erasmus University Rotterdam
Rotterdam, The Netherlands

DE Grobbee, MD, PhD
Julius Centre for General Practice and Patient Oriented Research
University Medical Centre Utrecht
Utrecht, The Netherlands

R de Groot, MD, PhD
Department of Immunology and Infectious Diseases
University Hospital Rotterdam/ Sophia Children's Hospital
Rotterdam, The Netherlands

J van der Lei, PhD
Department of Medical Informatics
Erasmus University Rotterdam
Rotterdam, The Netherlands

HA Moll, MD, PhD
Department of General Paediatrics
University Hospital Rotterdam/ Sophia Children's Hospital
Rotterdam, The Netherlands

KGM Moons, PhD
Julius Centre for General Practice and Patient Oriented Research
University Medical Centre Utrecht
Utrecht, The Netherlands

JB Oostenbrink, MSc
Institute for Medical Technology Assessment
Erasmus University Rotterdam
Rotterdam, The Netherlands

R Oostenbrink, MD, PhD
Department of General Paediatrics
University Hospital Rotterdam/ Sophia Children's Hospital
Rotterdam, The Netherlands

WK Redekop, PhD
Institute for Medical Technology Assessment
Erasmus University Rotterdam
Rotterdam, The Netherlands

EW Steyerberg, PhD
Center for Clinical Decision Sciences
Department of Public Health
Erasmus University Rotterdam
Rotterdam, The Netherlands

Dankwoord

Promoveren doe je gelukkig niet in je eentje. Daarom wil ik iedereen die mee heeft gewerkt aan de totstandkoming van dit proefschrift bij deze bedanken en een aantal mensen met name.

De hele projectgroep Ontwikkelingsgeneeskunde: Henriëtte Moll, Arda Derksen-Lubsen, Rick Grobbee, Carl Moons, Rianne Oostenbrink, Rogier Donders, Jan Oostenbrink, Ken Redekop, Willy Atema, Johan van der Lei, professor Büller en Minke Twijnstra. Bedankt voor alle nuttige vergaderingen door het hele land, alle goede ideeën en inzet, en de ondersteuning waar nodig: zonder jullie had het niet gekund. Rianne, bedankt voor het plaveien van de wegen en voor alle uitleg. Ook wil ik alle studenten die betrokken waren bij dit project bedanken voor hun inzet: Wilfried de Jong, Femke Mineur, Haley Kang en Jantine de Penning.

In het bijzonder wil ik mijn copromotor Henriëtte Moll bedanken. Lieve Henriëtte, dank je wel voor je fantastische begeleiding. Ik kon vrijwel altijd (ondanks dat je parttime werkt) bij je terecht voor nuttig prikkelend overleg, waarbij je me altijd aanzette tot kritisch nadenken. Daarnaast wil ik je bedanken voor de vrijheid die je me gaf, zowel in het uitwerken van het onderzoek, als in de privé-sfeer toen dat nodig was.

Dr. G. Derksen-Lubsen, Arda, bedankt voor je altijd zo nuchtere en doortastende blik op de zaken. Verder wil ik je bedanken voor je vertrouwen in mij, door jou ging het balletje rollen en kwam ik op mijn plek terecht. En nu ben je 'mijn' opleider!

Dr. K.G.M. Moons, beste Carl, jij bracht mij de beginselen van de epidemiologie bij en leerde mij artikelen schrijven. Onze discussies over de moeilijkere epidemiologische materie waren leerzaam en hielden ons scherp. Dank je wel voor jouw bemoedigende woorden bij tegenslag.

Mijn promotoren Prof.dr. D.E. Grobbee en Prof.dr. R. de Groot dank ik allereerst voor hun begeleiding en opbouwende kritiek. Beste Rick, jouw visie op de materie was altijd zeer helder en duidelijk, zo ook je commentaren. Ik heb veel van je geleerd! Beste Ronald, dank je wel voor de gezellige uitjes en congressen met 'jouw' subafdeling infectie-en immuunziekten waar ik als 'geadopteerde' onderzoeker mee naar toe mocht.

Alle verpleegkundigen, arts-assistenten en kinderartsen uit de regio Rotterdam en Den Haag die meewerk(t)en aan de prospectieve studiefasen dank ik hartelijk voor hun inzet.

Dr. E.W. Steyerberg, beste Ewout, bedankt voor je verfrissende (alhoewel soms zeer ingewikkelde) blik op de zaken. Overleg met jou was altijd zeer informatief en statistisch significant.

Rianne Oostenbrink, Carola Bouwhuis, Ezra Piepers en Liesbeth Duijts, mijn kamergenoten door de tijd heen, bedank ik voor de gezellige tijden. Carola, ik mis onze eigenzinnige ochtend-, middag-, en zelfs avond- 'kamergenootrituelen'. De laatste loodjes waren soms best zwaar, maar met ons tweeën een heel stuk lichter! Nu voor jou de allerlaatste loodjes nog, zet 'm op!

Annemarie F., Annemarie van R., Barbara, Carola, Clementien, Daphne, Debbie, Dederieke, Dick, Edwin, Ellen, Ester, Ezra, Floor, Geert, Gwenda, Hannerieke, Hestien, Hettie, Inge, Jeroen, Jessie, Jochem, José, Karen, Laurens, Leon, Liesbeth, Marieke den B., Marieke E., Marije, Marloes, Nicolette, Pieter, Rianne, Robert, Sander, Saskia, Sophie, Venje, Ward, Wendy, Yvonne: thanks voor de gezellige lunchuurtjes, borrels, promotiefeesten, Aio-weekends en andere uitjes.

Irene, dank je wel voor alle gezellige dingetjes samen, zowel op het werk als daarnaast en voor het grondig nalezen van het manuscript.

Mijn paranimfen, lieve Carola en Barbara. Wat heb ik een fijne tijd met jullie gehad. Jullie waren de beste winkelmaatjes, lunchpartners, theeleuten en borrelende kroegtijgers met luisterende oren en opbeurende woorden. Dank jullie wel, ik hoop dat dit blijft!

En dan mijn lieve ouders, mijn broer Marc en Victoria, mijn zusje Lize en Guido, en mijn dierbare vrienden: hier ging het nu allemaal om, hierdoor heb ik jullie allemaal de laatste maanden veel minder gezien dan mij lief is! Maar ik heb jullie niet vergeten. Dank voor al jullie steun op allerlei gebieden.

Lieve, lieve Dirk, m'n makkertje. Dank je wel voor je onvoorwaardelijke steun en toeverlaat en vooral voor je enorme geduld en begrip. Ik hoop dat we samen nog een lange gelukkige tijd voor de boeg hebben. Laten we er een feest van maken!

Curriculum vitae

Sacha Elisabeth Bleeker was geboren (10 juli 1972) en getogen in Apeldoorn. In 1990 behaalde zij haar VWO diploma aan het Stedelijk Gymnasium Apeldoorn. Aansluitend studeerde zij aan de Rijksuniversiteit Leiden, door uitloting aanvankelijk Biomedische Wetenschappen en vanaf 1991 Geneeskunde. In 1992 behaalde zij haar propedeutisch examen. Tijdens haar studie participeerde zij in migraine-onderzoeken op de afdeling Neurologie (Leids Universitait Medisch Centrum) en werkte zij als student-assistent op de bloedafname van de Bloedbank Leiden. Na haar afstudeeronderzoek (afdeling Endocrinologie, Leids Universitair Medisch Centrum) doorliep zij aansluitend haar keuzecoschap in Australië op de afdeling Obstetrie en Gynaecologie. Zij behaalde haar doctoraalexamen begin 1996. Na haar artsexamen in juni 1998 werkte zij als artsonderzoeker in het Juliana Kinderziekenhuis te Den Haag aan een tweetal projecten. Hiervan groeide één uit tot een promotieonderzoek (Ontwikkelingsgeneeskunde: 'Evaluatie van de waarde van diagnostiek bij kinderen') in het Sophia Kinderziekenhuis te Rotterdam en het Julius Centrum voor Huisartsgeneeskunde en Patiëntgebonden Onderzoek (Universiteit Utrecht), waar dit proefschrift het produkt van is. Tijdens dit promotietraject behaalde zij de graad Master of Science in de Klinische Epidemiologie (Nihes, Erasmus Universiteit Rotterdam). Per juni 2002 is zij begonnen aan de opleiding Kindergeneeskunde in het Juliana Kinderziekenhuis (opleider: mw. G. Derksen-Lubsen).

List of abbreviations

90%BI: 90% betrouwbaarheidsinterval
95%BI: 95% betrouwbaarheidsinterval
90%CI: 90% confidence interval
95%CI: 95% confidence interval
ß: regression coefficient
CPR: computer-based patient record
CRP: C-reactive protein/ C-reactief proteïne
CSF: cerebrospinal fluid
ED: emergency department
EMD: elektronisch medisch dossier
Hib: *Haemophilus influenzae* type b
JCH: Juliana Children's Hospital
LR: likelihood ratio
ORCA-SDE: Open Record for Care–Structured Data Entry
ROC area: area under the receiver operating characteristic curve
ROC oppervlakte: oppervlakte onder de receiver operating characteristic curve
SBI: serious bacterial infection
SCH: Sophia Children's University Hospital
WBC: white blood cell