

Validation of the Feverkidstool and procalcitonin for detecting serious bacterial infections in febrile children

Ruud G Nijman¹, Yvonne Vergouwe², Henriëtte A Moll¹, Frank J Smit³, Floor Weerkamp⁴, Ewout W Steyerberg², Johan van der Lei⁵, Yolanda B de Rijke⁶ and Rianne Oostenbrink¹

BACKGROUND: To validate the Feverkidstool, a prediction model consisting of clinical signs and symptoms and C-reactive protein (CRP) to identify serious bacterial infections (SBIs) in febrile children, and to determine the incremental diagnostic value of procalcitonin.

METHODS: This prospective observational study that was carried out at two Dutch emergency departments included children with fever, aged 1 month to 16 years. The prediction models were developed with polytomous logistic regression differentiating “pneumonia” and “other SBIs” from “non-SBIs” using standardized, routinely collected data on clinical signs and symptoms, CRP, and procalcitonin.

RESULTS: A total of 1,085 children were included with a median age of 1.6 years (interquartile range 0.8–3.4); 73 children (7%) had pneumonia and 98 children (9%) had other SBIs. The Feverkidstool showed good discriminative ability in this new population. After adding procalcitonin to the Feverkidstool, c-statistic for “pneumonia” increased from 0.85 (95% confidence interval (CI) 0.76–0.94) to 0.86 (0.77–0.94) and for “other SBI” from 0.81 (0.73–0.90) to 0.83 (0.75–0.91). A model with clinical features and procalcitonin performed similar to the Feverkidstool.

CONCLUSION: This study confirms the external validity of the Feverkidstool, with CRP and procalcitonin being equally valuable for predicting SBI in our population of febrile children. Our findings do not support routine dual use of CRP and procalcitonin.

Fever is among the most common presenting signs of illness in children (1). Identifying those febrile children with a serious infectious disease among the large majority who will have a benign, self-limiting infectious disease continues to be challenging. Many clinical signs and symptoms and biomarkers have been proposed as diagnostic predictors of serious bacterial infections (SBIs), but not one has yet proven useful for ruling in or ruling out the presence

of SBI as a solitary predictor (2–5). Prediction models that combine predictors of SBI appear more useful (6). A systematic review included the Feverkidstool, a clinical prediction model including both clinical signs and symptoms and the biomarker C-reactive protein (CRP), as a potential tool for supporting physicians in their decision-making for safely discharging febrile children from an emergency department (ED) (7). This Feverkidstool assists in distinguishing children with pneumonia and children with other SBIs from children without SBI (8). However, external validity of such tools needs to be confirmed to ensure their usefulness elsewhere, a step that is often lacking in pediatric research (9). In addition, procalcitonin (PCT), another biomarker that has been linked with the presence of SBI, had not been considered in the original study (3,4,10–12). Hence, questions on the diagnostic value of PCT in addition to clinical signs and symptoms as well as the comparability of PCT and CRP remain unanswered. In this study, we aimed to validate the Feverkidstool, and to determine the incremental diagnostic value of PCT in a prospective observational cohort of children with fever at risk of having an SBI at the ED.

METHODS

Design

First, we performed a prospective observational study to validate the Feverkidstool, a previously presented prediction model consisting of clinical signs and symptoms and CRP, (8) in a cohort of febrile children who were included from two Dutch EDs. Next, we determined the incremental value of PCT to this prediction model. In addition, we looked at the diagnostic properties of a model with clinical signs and symptoms and PCT, and compared these with the diagnostic properties of the Feverkidstool.

Setting and Participants

This study was conducted at two EDs in Rotterdam, the Netherlands, and it included children aged 1 month to 16 years. Children were included from the Erasmus MC—Sophia Children’s Hospital between February 2009 and May 2012, and from the Maastad Hospital between May 2011 and May 2012 (ref. 13). The Erasmus MC is a university hospital and its pediatric ED is visited by ~9,000 children annually, of whom 90% require basic pediatric care (14).

¹Department of Pediatrics, Erasmus MC—Sophia Children’s Hospital, Rotterdam, The Netherlands; ²Department of Public Health, Centre of Medical Decision Making, Erasmus MC—University Medical Center Rotterdam, Rotterdam, The Netherlands; ³Department of Pediatrics, Maastad Hospital, Rotterdam, The Netherlands; ⁴Department of Clinical Chemistry, Maastad Hospital, Rotterdam, The Netherlands; ⁵Department of Medical Informatics, Erasmus MC—University Medical Center Rotterdam, Rotterdam, The Netherlands; ⁶Department of Clinical Chemistry, Erasmus MC—University Medical Center Rotterdam, Rotterdam, The Netherlands. Correspondence: Ruud G. Nijman (r.nijman@imperial.ac.uk)

Received 15 November 2016; accepted 16 July 2017; advance online publication 8 November 2017. doi:10.1038/pr.2017.216

The Maasstad Hospital is a teaching hospital with a mixed adult and pediatric ED that has some 10,000 pediatric medical emergency consultations yearly. At both study sites a study protocol was introduced that involved routinely recruiting patients at the moment of triage, measuring PCT and CRP during their stay at the ED, recording standardized clinical signs and symptoms in all eligible patients, and standardized follow-up after discharge from the ED. Fever was defined as a body temperature $\geq 38.5^\circ\text{C}$ at triage, fever recorded at home within 24 h before consultation, fever as reason for referral by the general practitioner, or fever as positive discriminator in the Manchester Triage System. Children with an underlying chronic condition requiring specialist pediatric follow-up that is expected to last for at least 1 year were excluded (15). Well-appearing febrile children with a clear focus of an upper airway infection not undergoing additional investigations were also excluded as they were at such a low overall risk of SBI that diagnostic tools would not contribute to clinical management; this concerned with only well appearing children where the nurse felt there was a clear focus of an upper respiratory tract infection, and no other warning signs present, while being in the ED; children with any persisting diagnostic uncertainty were eligible for the study. Additional diagnostic tests were performed at the discretion of the attending physician. Children who revisited the ED within 5 days after their initial visit were only eligible with their initial visit, although data from all visits were considered in establishing final diagnosis.

Ethics Approval

The study protocol was approved by the institutional ethical committees of both the Erasmus MC (MEC-2007-066) and the Maasstad Hospital (2010/64). Informed consent was required and written informed consent was obtained from all children who had CRP and/or PCT tests performed.

Data Collection

In the Erasmus MC, trained nurses recorded clinical signs and symptoms of all children with fever in a standardized electronic patient form that was tailored specifically to this study's purposes, and which was incorporated in the patient's digital hospital record. Triage data and patient characteristics were also recorded as a part of the ongoing data collection (16). Simultaneously, physicians reported on predictor variables by completing a predefined electronic proforma. In the Maasstad Hospital, patient characteristics, data on vital signs, triage data, and data on diagnostics and follow-up were collected prospectively in a standardized digital patient file by the triaging nurse. Additional clinical data were recorded by the attending physician on a paper research form. Nurses and physicians were trained in data collection and informed of the study, but were blinded for the specific outcome measures. Regular feedback on data quality and inclusion rates was provided by the researchers for both settings.

CRP and PCT

CRP was measured either by a bedside test (84%) or by a traditional laboratory test (16%). CRP values were available to the physician for further clinical evaluation. In the Erasmus MC laboratory CRP was determined in heparin plasma with an immunoturbidimetric CRP assay using a Modular system (Roche, Mannheim, Germany). In the Maasstad laboratory CRP was determined in heparin plasma with a nephelometric CRP assay using a Dimension VISTA 1500 system (Siemens, Erlangen, Germany). The Afinion AS100 (Axis-Shield PoC AS, Oslo, Norway, distributed by Clindia Benelux BV) was used for CRP bedside testing, and was used at both EDs. Before the study was carried out, validity and comparability of these tests were assessed and approved (13). PCT samples from both study sites were analyzed in the Erasmus MC using the Brahms Kryptor Assay (13). PCT values were available for research purposes only and were not available to the physician for further clinical evaluation.

Predictor Variables

The original Feverkidstool included clinical variables and CRP (Table 1); variables were defined as previously described (8). That is, age was modeled piece-wise linear, which accounted for continuously differentiating risk profiles in children < 1 years old and of children aged 1 year or older. Duration of fever was truncated to the 97th percentile at a maximum of 6 days. Body temperature was measured in degrees Celsius and was measured rectally in children of up to 5 years, and tympanic in older children at the Erasmus MC—Sophia, and either rectally or tympanic at the Maasstad at the discretion of the nurse. Heart rate (beats per minute) and respiratory rate (breaths per minute) were recorded according to standards of the Advanced Pediatric Life Support group (APLS), and categorized using APLS age-specific thresholds (17). Oxygen saturation was dichotomized at $< 94\%$ O_2 , and, when not recorded, the oxygen saturation was assumed to be normal at $\geq 94\%$ O_2 . Ill appearance was defined on a 2 point scale according to the nurse's impression at triage. If missing, ill appearance was completed by the physician's impression (47%) and ultimately (in 5% of the eligible children) by using the Young infant (children aged ≤ 3 months) or Yale Observational scale (children aged > 3 months) score of ≥ 8 (refs 18,19). A prolonged peripheral capillary refill was defined as ≥ 3 s. Chest wall retractions were defined as present if subcostal retractions, or intercostal retractions, or tracheal tug were observed. We truncated CRP at a value of 225 mg/l and PCT at 6.0 ng/ml, both at the 97th centiles, and we used the natural logarithms of CRP and PCT (Ln(CRP) and Ln(PCT)).

Outcome Measures

We defined SBI according to a predefined reference standard based on positive cultures from otherwise sterile locations, focal abnormal radiologic findings, and expert clinical consensus by the investigators (R.O., E.K., and R.N.; *Supplementary Table S1* online; (8,20)). We subsequently categorized SBI into children with pneumonia and children with other SBI vs. children with no SBI to account for the heterogeneity of presenting signs and symptoms similar to the definitions used previously (8). Other SBIs consisted of urinary tract infections, septicemia, meningitis, orbital cellulitis, erysipelas, bacterial gastroenteritis, bacterial arthritis, bacterial upper airway infection, and bacterial osteomyelitis. Final diagnoses were coded by trained medical students, who were blinded for the study's specific aims and PCT values, but not for clinical data or CRP, and all diagnoses were reviewed for consistency by the investigators. Consensus diagnoses were mandatory for all children not completely fulfilling the reference standard, children with an atypical clinical course, and for children with a prolonged hospital admission or a prolonged course of antibiotics. These included children who received antibiotics before cultures were taken, most importantly in children with abnormal urinalyses and negative cultures, children with pleiocytosis in cerebral spinal fluid and negative cultures (i.e., aseptic meningitis), and children with radiographic evidence of musculoskeletal infection and negative cultures. Follow-up of all eligible children included a standardized telephone call after 72 h or follow-up visit at the ED of discharged patients, and checking for revisits in the medical system to ensure that no cases of SBI were missed, thus minimizing verification bias (21).

Missing Values

Children with values for both CRP and PCT were available for this study. Missing variables for clinical signs and symptoms were imputed 10 times using the MICE algorithm, R statistical software (*Supplementary Table S2*; (ref. 22)). The imputation model included the Feverkidstool variables, PCT, the outcome measure SBI, and descriptors of case-mix, such as arrival time, season of presentation, triage urgency, antibiotics prescription, hospitalization, and a cohort variable (23). Multiple imputation techniques enabled analyzing all available data, without adding information to the original available data, and thus not adding information on the incremental value of

Table 1. General characteristics of children with C-reactive protein and procalcitonin available

	Study population (n = 1,085)
<i>Predictor variables</i>	
Age (years); median (25th–75th percentile)	1.59 (0.84–3.45)
Gender; female	479 (44%)
Duration of fever (days)	
Total	1,084 (100%)
Median (25th–75th percentile)	2 (1–4)
Temperature (°C)	
Total	1,068 (98%)
Median (25th–75th percentile)	38.9 (38.2–39.5)
Respiratory rate (breaths per minute)	
Total	885 (82%)
Median (25th–75th percentile)	32 (24–40)
Heart rate (beats per minute)	
Total	994 (92%)
Median (25th–75th percentile)	143 (120–164)
Oxygen saturation; <94% O ₂	37 (3%)
Capillary refill time	
Total	1,064 (98%)
≥ 3 s	128 (12%)
Chest wall retractions	
Total	1,071 (99%)
Present	141 (13%)
Ill appearance	
Total	1,072 (99%)
Ill	415 (39%)
C-reactive protein (mg/l)	
Median (25th–75th percentile)	15 (8–40)
Procalcitonin (ng/ml)	
Median (25th–75th percentile)	0.19 (0.10–0.58)
<i>Case-mix descriptors</i>	
Manchester triage system urgency ^a	
Total	1,074 (99%)
Urgency level “immediate”	22 (2%)
Urgency level “very urgent”	458 (43%)
Urgency level “urgent”	464 (43%)
Urgency level “non-urgent”	128 (12%)
Urgency level “standard”	2 (0%)
Shift at presentation	
07.00–15.00	123 (11%)
15.00–23.00	486 (45%)
23.00–07.00	476 (44%)
Season of presentation	
Winter	371 (34%)

Table 1 Continued

	Study population (n = 1,085)
Spring	244 (22%)
Summer	200 (18%)
Autumn	270 (25%)
Antibiotics prescription; yes	436 (40%)
Hospitalization	
Yes	325 (30%)
<i>Outcome measure</i>	
Serious bacterial infection, present	171 (16%)
Pneumonia	73 (7%)
Other SBIs	
Urinary tract infection	56 (5%)
Septicemia/meningitis	8 (1%)
SBI other	34 (3%)
No SBI	914 (84%)

If data were not available for all patients, the total number of available data are noted.

SBI, serious bacterial infection.

^aManchester triage system urgency classification: Erasmus MC, MTS version 2 (including updates for children with fever), Maasstad, MTS version 1.

PCT (24). Analyses were performed separately in the 10 imputed data sets and were combined using Rubin’s rules (25).

Statistical Analysis

First, we validated the original prediction model by determining discriminative ability and calibration, thus validating the diagnostic performance of the Feverkidstool in this new population. Next, we refitted the Feverkidstool variables and added PCT to the model. In addition, we refitted a model containing clinical signs and symptoms and PCT, but without CRP. For all analyses we used polytomous regression analyses differentiating pneumonia vs. no SBI and other SBIs vs. no SBI. Polytomous regression analyses allow for distinguishing more than two outcome categories while considering the total population of febrile children and their potential final diagnoses. Discriminative ability was expressed by pairwise c-statistics and the polytomous discrimination index (PDI). Pairwise c-statistics were estimated as one outcome category (i.e., pneumonia or other SBIs) vs. one other outcome category (here: no SBI), equaling receiver operating characteristic curves. The PDI assesses discriminative ability considering all possible outcome categories with a value of 0.33 (1/number of outcome categories, here 1/3), indicating a non-informative test and a value of one indicating a perfect discriminating test. The PDI is the average of category-specific polytomous discrimination indices each of which separately reflects the ability to predict a case better than non-cases for all outcome categories of interest (26). Bootstrapping (100 iterations) was used to determine 95% confidence intervals (CIs) of the PDI. Calibration reflects the predicted risks vs. the true observed risks for children with either pneumonia or other SBIs vs. no SBI. Incremental diagnostic value of PCT was determined by its predictive strength in the model with clinical signs and symptoms and CRP (27), and by differences in the discriminative ability between the models assessed as the differences in pairwise C-statistics and PDI (26,28,29). We estimated sensitivities, specificities, and positive and negative likelihood ratios for several risk thresholds of the refitted prediction models (adjusted for

model optimism) by comparing one outcome category vs. the rest. All analyses were performed using R statistical software (22).

RESULTS

Population Characteristics

In all, 2,703 febrile children were included in this study, of whom 1,085 had both CRP and PCT performed and were available for analysis (Table 1 and Figure 1); their median age was 1.6 years (interquartile range: 0.8–3.4), and 73 children (7%) had pneumonia, and 98 children (9%) had other SBIs (Table 2). Overall, children with CRP and PCT measures had a slightly higher incidence of SBI, were hospitalized more often, and were prescribed antibiotic therapy more frequently compared with the total included population, (Supplementary Table S3). Seventy-four percent of the children had their CRP measured: characteristics of children with both CRP and PCT were similar to the children with only CRP.

Validation of the Original Prediction Model

The Feverkidstool showed good external validity: using the original coefficients of the Feverkidstool (Supplementary Appendix S4), pairwise c-statistics were 0.84 (95% CI: 0.74–0.94) for the prediction of pneumonia and 0.82 (95% CI:

0.73–0.91) for other SBI; PDI was 0.66 (95% CI 0.60–0.70). The predicted risks for having pneumonia were clearly higher for children with pneumonia than for those without pneumonia; similar findings were observed for the risk of having other SBIs vs. children without other SBIs (Figure 2). Calibration was reasonable with an underestimation of the predicted risks in the highest risk quartile, mostly because of sparsity of cases in higher-risk categories (Figure 3).

Incremental Value

The predictive strength of PCT when added to a refitted model was reflected by regression coefficients of Ln(PCT) of 0.34 (SE 0.13) for pneumonia and 0.45 (SE 0.11) for other SBIs; (Supplementary Table S5). The likelihood ratio test of adding PCT to the model was significant with a P value <0.001 (difference in deviance between the two polytomous models was 23.4, with 2 degrees of freedom). The addition of PCT did not alter regression coefficients of the other variables substantially, but lessened the predictive effect of CRP (Supplementary Table S5). The model that included both CRP and PCT had similar discriminative ability as the refitted Feverkidstool: for the prediction of pneumonia the pairwise

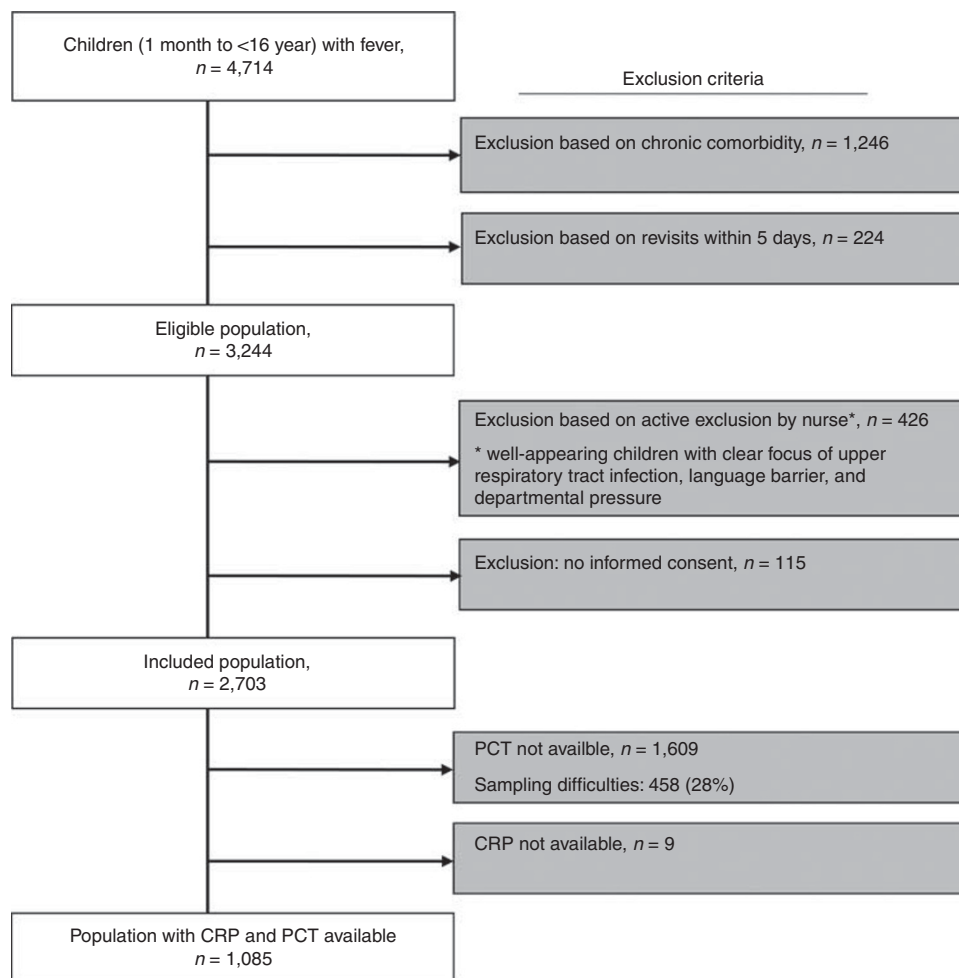


Figure 1. Flowchart of included population.

Table 2. Feverkidstool: predictor distribution for the three outcome categories

	No SBI (n=914)	Pneumonia (n=73)	Other SBIs ^a (n=98)
Age, years	1.50 (0.84–3.14)	3.39 (1.59–6.02)	1.72 (0.60–5.28)
Sex, female	379 (41%)	39 (53%)	61 (62%)
Duration of fever, days ^b	2.0 (1.0–4.0)	3.0 (2.0–4.0)	2.0 (1.0–5.0)
Temperature, °C	38.9 (38.1–39.5)	39.2 (38.5–40.0)	39.1 (38.3–39.8)
Tachypnoea, present ^c	274 (37%)	43 (63%)	19 (28%)
Tachycardia, present ^c	348 (41%)	49 (68%)	39 (47%)
Oxygen saturation, <94% O ₂	25 (3%)	12 (17%)	0 (0%)
Capillary refill time, ≥ 3 s	110 (12%)	7 (10%)	11 (11%)
Chest wall retractions, present	110 (12%)	30 (42%)	1 (1%)
Ill appearance, ill	334 (37%)	47 (65%)	34 (35%)
Procalcitonin, ng/ml	0.16 (0.09–0.41)	0.77 (0.24–4.84)	0.81 (0.28–2.62)
C-reactive protein, mg/l	12 (8–30)	51 (15–143)	56 (23–133)

SBI, serious bacterial infection.

Median and interquartile range, and absolute number with percentage where appropriate.

^aOther SBIs include septicemia/meningitis, urinary tract infections, and others.

^bMaximum duration of fever is 6 days.

^cDefined according to age-specific Advanced Pediatric Life Support threshold values.

c-statistic increased from 0.85 (95% CI: 0.76–0.94) to 0.86 (95% CI: 0.77–0.94) and for other SBI from 0.81 (95% CI: 0.73–0.90) to 0.83 (95% CI: 0.75–0.91). The PDI increased from 0.65 (95% CI: 0.59–0.70) to 0.66 (95% CI: 0.61–0.71; **Table 3**). A refitted model with clinical signs and symptoms and PCT, but without CRP, achieved similar model performance (**Table 3**). Distributions of predicted risks were similar for the three refitted models (**Figure 4**). All models had similar strength for ruling in and ruling out SBI (**Table 4**): risk thresholds ≥10% were, based on high specificities and positive likelihood ratios, useful for ruling in SBI, and risk thresholds ≤2.5% were, based on high sensitivities and negative likelihood ratios, useful for ruling out SBIs. A single value of CRP or PCT was not as strong for predicting SBI as the prediction model and lacked predictive value for all types of SBI (**Supplementary Table S6**; (ref. 13)). Overall, models including a biomarker performed notably better than a model including just clinical signs and symptoms, in particular for other SBIs (**Table 3**).

DISCUSSION

Principal Findings

This is the first study to confirm the external validity of the Feverkidstool, a clinical prediction model with clinical signs and symptoms and CRP that calculates the risk of having SBI in febrile children at the ED. Among those children with both CRP and PCT available, the incremental value of PCT was significant; however, it showed only marginal improvement of the discriminative ability of the Feverkidstool. Similarly, when looking at the diagnostic performance at several risk thresholds, a model including both CRP and PCT did not outperform such models containing either CRP or PCT. Models with clinical signs and symptoms and one of the two biomarkers performed similar, but better than a model with clinical signs and symptoms alone.

Comparison with Existing Literature

In contrast to our findings, a systematic review claimed that PCT as a single marker was a better biomarker for detecting SBI in febrile children compared with CRP, and was especially useful for ruling out the presence of SBI (4). Other recent papers also concluded that PCT was superior to CRP in identifying SBI in febrile children, in particular for detecting more invasive bacterial infections in young infants (10,11,30). However, these studies were mostly performed in highly selected patient populations, such as well-appearing young infants with fever, ill-appearing children, or hospitalized children, as opposed to our broader population of febrile children at the ED with diagnostic uncertainty (3,10,11,30). In addition, many of the studies that evaluated the diagnostic value of PCT looked at the predictive ability of PCT as a single predictor not considering the clinical context or value of CRP (3,4). One prediction model that combined CRP and PCT together with urinalysis, the Lab-score, showed promising diagnostic value for detecting SBI in febrile children (31,32). However, clinical signs and symptoms were not incorporated and the additional value of one biomarker in relation to the other was unclear. Moreover, two validation studies failed to replicate the diagnostic performance for ruling out SBI (13,33). Furthermore, although in most previous research SBIs are questionably traditionally grouped in a single homogenous outcome category, we assessed the diagnostic value of PCT in a polytomous manner estimating the value of PCT to predict pneumonia and other SBIs simultaneously. Next, some of the clinical signs and symptoms considered in the prediction model, most importantly age and duration of fever, have been associated with differentiated predictive ability of CRP and PCT. For example, concentration levels of PCT are elevated at an earlier stage of disease than levels of CRP (11). Likewise, PCT appears to be a better predictor of invasive bacterial infections in young infants, in particular in

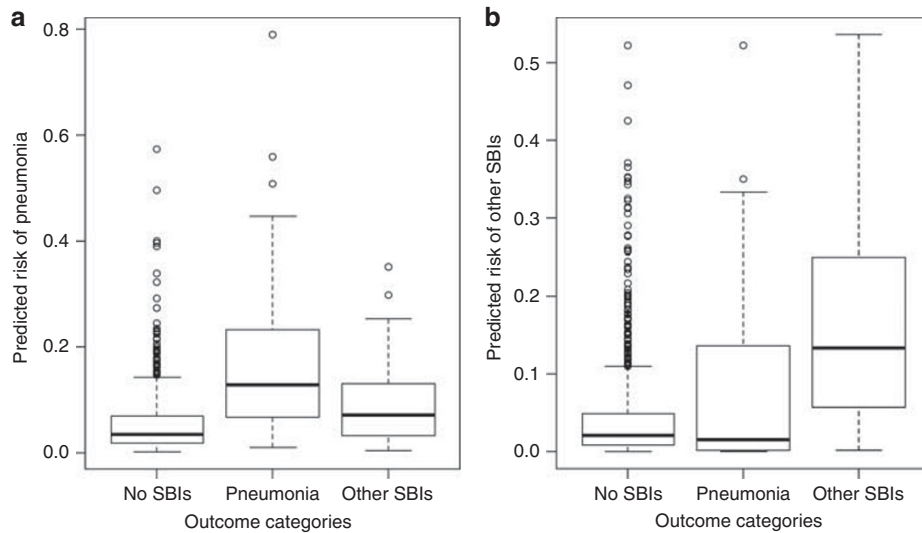


Figure 2. External validation of the original prediction model with clinical signs and symptoms and C-reactive protein. Predicted risks of pneumonia (a) and of other SBI (b) for children diagnosed with pneumonia, other SBI, or no SBI. SBI, serious bacterial infection.

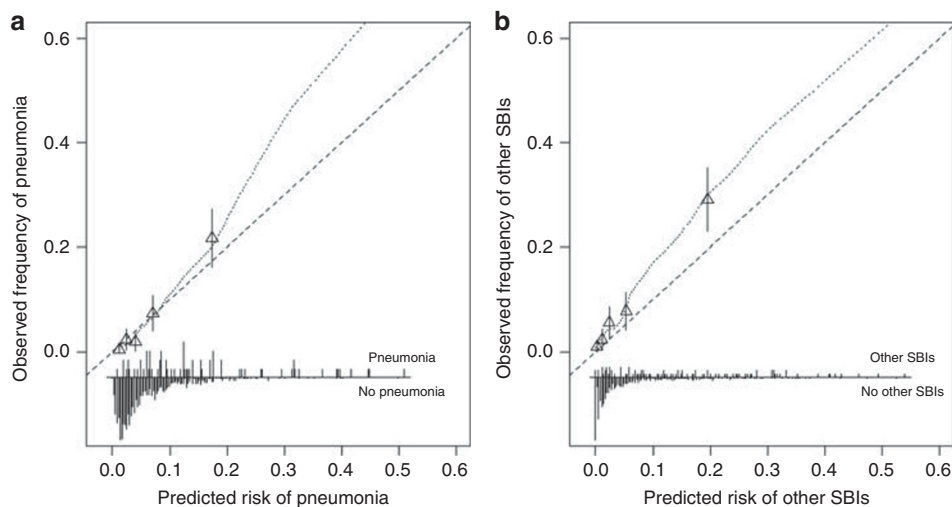


Figure 3. External validation of the original prediction model with clinical signs and symptoms and C-reactive protein. Calibration plots for the risk of having pneumonia (a) and risk of having other SBIs (b). Calibration plot of the predicted risk (x axis) and the observed frequency of pneumonia (a) and other SBIs (b) (with 95% CI, y axis). The triangles represent the mean (predicted vs. observed) risk estimates by quintiles of predicted risks. The dashed diagonal line represents ideal calibration. The distribution of the predicted risks of patients with pneumonia ($n=73$ (a)), other SBIs ($n=98$, (b)), and other patients ($n=1,012$ (a) and $n=987$ (b)) is shown in the bottom of the graph, parallel to the x axis. CI, confidence interval; SBI, serious bacterial infection.

those with a short onset of fever (10). Although our model combined age and duration of fever together with CRP and PCT, thus accounting for potential confounding, our data were not specifically tailored to look at the group of young febrile infants with a short duration of fever. In our cohort only 70 children were aged <3 months with 15 children having SBI (21%); of these children 49 (70%) had fever for ≤ 24 h, with 11 children having an SBI. In a previous study we showed that duration of fever had no significant effect on the interpretation of diagnostic values of CRP and PCT in a general population of febrile children up to 16 years of age in pediatric emergency care (13).

Clinical Implications

Altogether, our data suggest that routinely determining both CRP and PCT in the evaluation of febrile children at the ED does not appear to be warranted. The near identical diagnostic performances at the different risk thresholds across the models do not seem to justify potential differences in clinical management for most children at present. However, PCT was still a significant predictor of SBI and an additional PCT value might yet be useful in individual patients with persisting diagnostic uncertainty. This would depend on individualized clinical risk thresholds that, for example, would imply starting antibiotic therapy or

Table 3. Discriminative ability of the refitted prediction models

	Clinical signs	Clinical signs and C-reactive protein (Feverkidstool)	Clinical signs, C-reactive protein and procalcitonin	Clinical signs and procalcitonin
Polytomous discrimination index (95% confidence interval) ^a	0.59 (0.53–0.64)	0.65 (0.59–0.70)	0.66 (0.61–0.71)	0.66 (0.70–0.61)
<i>Pairwise C-statistic (95% confidence interval)^a</i>				
Pneumonia vs. no SBI	0.83 (0.74–0.92)	0.85 (0.76–0.94)	0.86 (0.77–0.94)	0.84 (0.74–0.93)
Other SBI vs. no SBI	0.74 (0.65–0.84)	0.81 (0.73–0.90)	0.83 (0.75–0.91)	0.81 (0.73–0.89)
<i>R square^a</i>				
Pneumonia vs. no SBI	0.21	0.32	0.33	0.29
Other SBI vs. no SBI	0.08	0.24	0.27	0.22

SBI, serious bacterial infection.

^aModel performance adjusted for optimism with bootstrap procedure (100 iterations).

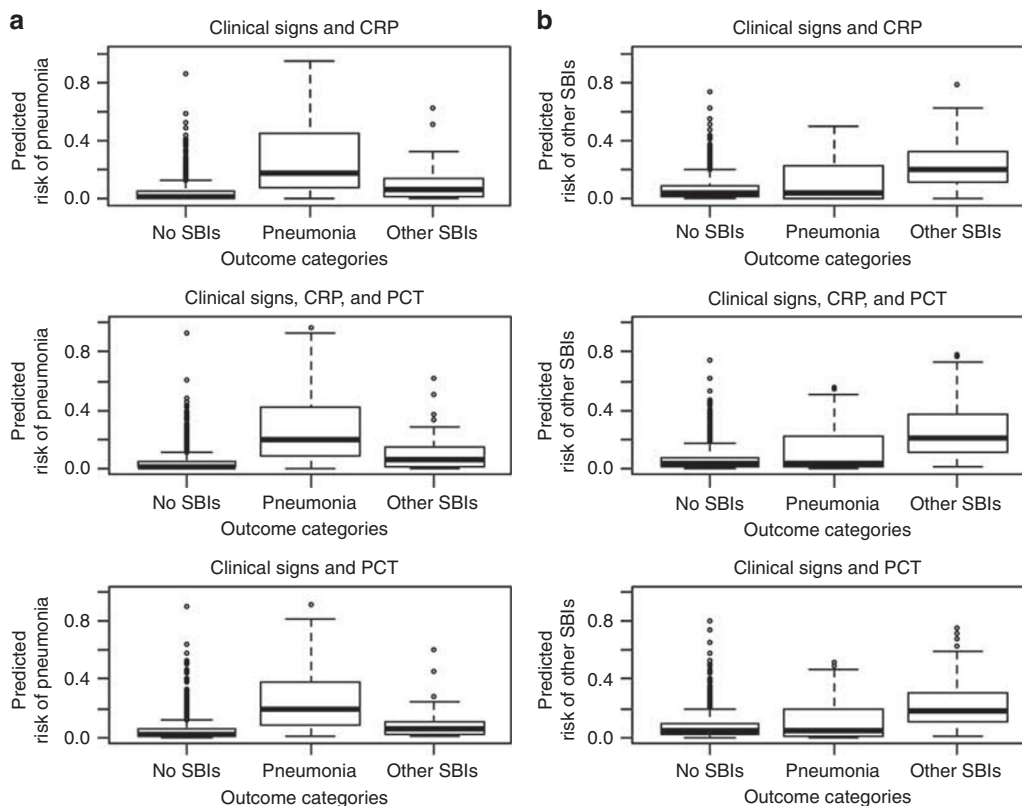


Figure 4. Distribution of predicted risks for pneumonia and other SBIs. Predicted risks of pneumonia (a) and other SBIs (b) for children diagnosed with pneumonia, other SBIs, or no SBI for the three refitted models. SBI, serious bacterial infection.

performing additional diagnostic tests. Yet, these thresholds currently remain elusive for the topic of febrile children and should be the focus of future research (23,34). Our study adds to the conclusions of a systematic review that the Feverkidstool can be used as a validated tool for supporting the decision of safely discharging febrile children from the ED (7). As always, debate will continue on how

reliable a diagnostic test needs to be to account for the potentially severe consequences of missing a diagnosis of SBI, but our findings imply a guiding diagnostic role for the Feverkidstool at least. This study also provides more evidence that CRP and PCT can be used interchangeably in an emergency care setting, depending on local preferences for one or the other.

Table 4. Diagnostic performances of prediction models including clinical signs and symptoms and either C-reactive protein or procalcitonin and a model including clinical signs and symptoms and both C-reactive protein and procalcitonin

	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)
<i>Clinical signs and C-reactive protein (Feverkidstool)</i>				
<i>Pneumonia</i>				
Risk of ≥ 2.5%	0.88 (0.79–0.97)	0.56 (0.53–0.60)	1.99 (1.79–2.22)	0.21 (0.11–0.43)
Risk of ≥ 5%	0.83 (0.73–0.93)	0.71 (0.68–0.74)	2.79 (2.44–3.20)	0.24 (0.13–0.45)
Risk of ≥ 10%	0.63 (0.50–0.75)	0.84 (0.82–0.87)	3.83 (3.08–4.77)	0.44 (0.31–0.64)
Risk of ≥ 15%	0.47 (0.35–0.59)	0.91 (0.89–0.93)	4.90 (3.58–6.71)	0.58 (0.45–0.75)
Risk of ≥ 30%	0.31 (0.20–0.42)	0.97 (0.96–0.98)	10.64 (6.42–17.63)	0.71 (0.60–0.84)
<i>Other SBIs</i>				
Risk of ≥ 2.5%	0.93 (0.87–1.00)	0.39 (0.35–0.43)	1.53 (1.41–1.66)	0.16 (0.07–0.37)
Risk of ≥ 5%	0.84 (0.75–0.92)	0.58 (0.55–0.62)	1.99 (1.76–2.23)	0.28 (0.17–0.46)
Risk of ≥ 10%	0.75 (0.66–0.85)	0.76 (0.73–0.78)	3.06 (2.60–3.60)	0.33 (0.22–0.48)
Risk of ≥ 15%	0.60 (0.49–0.71)	0.84 (0.82–0.87)	3.85 (3.08–4.80)	0.47 (0.36–0.62)
Risk of ≥ 30%	0.27 (0.18–0.36)	0.95 (0.93–0.96)	4.76 (3.10–7.30)	0.77 (0.67–0.89)
<i>Clinical signs, C-reactive protein and procalcitonin</i>				
<i>Pneumonia</i>				
Risk of ≥ 2.5%	0.88 (0.79–0.97)	0.56 (0.53–0.60)	2.01 (1.81–2.23)	0.21 (0.10–0.42)
Risk of ≥ 5%	0.82 (0.72–0.92)	0.73 (0.70–0.76)	2.98 (2.59–3.43)	0.24 (0.14–0.43)
Risk of ≥ 10%	0.63 (0.51–0.75)	0.85 (0.82–0.87)	4.03 (3.20–5.07)	0.44 (0.31–0.62)
Risk of ≥ 15%	0.55 (0.42–0.67)	0.91 (0.89–0.93)	5.98 (4.48–8.00)	0.49 (0.37–0.66)
Risk of ≥ 30%	0.34 (0.22–0.45)	0.97 (0.96–0.99)	12.90 (7.73–21.51)	0.68 (0.57–0.82)
<i>Other SBIs</i>				
Risk of ≥ 2.5%	0.94 (0.88–1.00)	0.42 (0.39–0.45)	1.61 (1.50–1.73)	0.14 (0.05–0.37)
Risk of ≥ 5%	0.84 (0.75–0.92)	0.61 (0.58–0.65)	2.13 (1.89–2.41)	0.26 (0.16–0.44)
Risk of ≥ 10%	0.75 (0.66–0.85)	0.78 (0.75–0.81)	3.40 (2.90–3.99)	0.32 (0.21–0.47)
Risk of ≥ 15%	0.60 (0.49–0.70)	0.85 (0.83–0.88)	3.92 (3.13–4.91)	0.47 (0.36–0.63)
Risk of ≥ 30%	0.31 (0.22–0.41)	0.94 (0.93–0.96)	5.61 (3.72–8.47)	0.73 (0.62–0.85)

Table 4 Continued

	Sensitivity (95% confidence interval)	Specificity (95% confidence interval)	Positive likelihood ratio (95% confidence interval)	Negative likelihood ratio (95% confidence interval)
<i>Clinical signs and procalcitonin</i>				
<i>Pneumonia</i>				
Risk of ≥ 2.5%	0.89 (0.80–0.99)	0.54 (0.50–0.57)	1.93 (1.74–2.13)	0.19 (0.08–0.43)
Risk of ≥ 5%	0.79 (0.68–0.90)	0.71 (0.68–0.74)	2.66 (2.29–3.09)	0.30 (0.18–0.49)
Risk of ≥ 10%	0.65 (0.53–0.77)	0.85 (0.83–0.88)	4.37 (3.51–5.44)	0.41 (0.29–0.57)
Risk of ≥ 15%	0.59 (0.47–0.72)	0.91 (0.89–0.93)	6.75 (5.15–8.86)	0.45 (0.33–0.61)
Risk of ≥ 30%	0.34 (0.22–0.45)	0.97 (0.96–0.99)	13.52 (8.02–22.81)	0.68 (0.56–0.82)
<i>Other SBIs</i>				
Risk of ≥ 2.5%	0.95 (0.89–1.01)	0.36 (0.32–0.39)	1.46 (1.36–1.57)	0.14 (0.05–0.44)
Risk of ≥ 5%	0.87 (0.79–0.95)	0.55 (0.51–0.59)	1.91 (1.71–2.14)	0.23 (0.12–0.45)
Risk of ≥ 10%	0.70 (0.60–0.80)	0.76 (0.73–0.79)	2.94 (2.47–3.51)	0.39 (0.27–0.55)
Risk of ≥ 15%	0.54 (0.43–0.64)	0.84 (0.82–0.87)	3.37 (2.67–4.27)	0.55 (0.43–0.70)
Risk of ≥ 30%	0.25 (0.16–0.35)	0.96 (0.94–0.97)	6.02 (3.78–9.58)	0.78 (0.68–0.89)

SBI, serious bacterial infection.

Strengths and Limitations

This study is the first to address the diagnostic value of PCT in combination with clinical signs and symptoms and CRP: our suggested approach of using a validated prediction model and evaluating the incremental value of a new diagnostic test in its clinical context makes a strong case for the validity of our findings. A strength is the large, prospective cohort of febrile children of whom we collected standardized clinical signs and symptoms, and in whom we attempted to routinely perform CRP and PCT tests. Nevertheless, a significant amount of the total included children did not have PCT measured, and our findings might represent a selected population of febrile children at higher risk of SBI, requiring future studies in other settings to confirm our findings. Missing PCT values were largely because of prospective sampling difficulties, possibly indicating a restricting factor for the routine use of this test. If anything, it seems that the presented population is at higher risk for SBI than the total included population, with sampling being more successful likely due to a higher proportion of venesections rather than capillary sampling methods. Yet, we feel our cohort is still more representative of a population of febrile children at risk for SBI in the ED than most previous studies (10–12,35–37). The children with only CRP approximated the characteristics of children with both CRP and PCT closely, diffusing any selection bias. Another strength is the strong discriminative ability and reasonable calibration of the Feverkidstool in this new population, supporting its external validity. Previously, in a validation study conducted at several European emergency care departments, prediction rules for detecting SBIs that comprised only clinical signs lacked discriminative value to be useful in clinical practice (38).

Differences in data collection and population characteristics between the two settings were a limitation, but this heterogeneity also forms a strong test of external validity. Next, as is inherently related to large prospective observational studies, some data on clinical signs and symptoms were missing. The amount of missing data was limited (Table 1), and by multiplying imputing missing values the observed effects of all available data could be used, estimating the incremental value of PCT more precisely (39). In addition, CRP levels, as well as other clinical data, were available for the clinician, whereas PCT values were not. Although CRP was not among the main criteria of the reference standard, it might have influenced clinical management and requests of additional diagnostic tests, as well as final coding of the outcome categories. Another limitation was posed by the absence of a valid, and ethically justifiable, gold standard for bacterial pneumonia (40,41). Focal infiltrates on chest radiographs in combination with suiting clinical features, as used in this study, arguably constitute the best proxy (42), at least including all children with definite lower respiratory tract infections (43). Other limitations associated with the development of the original prediction model were addressed previously (8). For example, the lack of certain reference tests was substituted with follow-up reducing the risk of any

missed SBI, as such dealing with the ethical dilemma of not being able to perform all reference tests in all patients. The effect of missing reference tests remains important to discuss. In our departments, antibiotics were typically only started after senior clinical review and after performing reference tests following local and national guidelines. In addition, the standardized follow-up tackled the problem of false-negatives, i.e., children with SBI who were discharged without antibiotic treatment or additional diagnostic testing. Still, some children with SBI might have gone undetected. Noteworthy, although carefully reviewed for their impact in children with serious infections, the use of antibiotics before presentation to the ED was not used as an exclusion criteria. This might have influenced the coding of the outcome diagnosis of children with negative culture results. However, antibiotics use in the Netherlands for uncomplicated infections is fairly low (44), and reference tests, for example, for children suspected of UTI, will mostly be taken before starting antibiotics in both primary and secondary care. In addition, children were only included if the research laboratory markers were taken at the same time as the blood cultures and before intravenous antibiotics were given. In addition, children on antibiotic prophylaxis for an underlying chronic illness were not eligible for the study. Another limitation was inter-rater variability, while still needing to be taken into account, was reduced by considering mostly objective clinical parameters and was considered of relatively limited importance for this robust validation study. Finally, a number of other performance measures have been developed to report on incremental diagnostic values, most notably the net reclassification index, decision curve analysis, and net benefit (23). However, at present these measures for dichotomous outcomes have not been fully developed for polytomous models.

CONCLUSION

The Feverkidstool can now be considered as a validated tool for supporting clinical decision-making, such as to safely discharge febrile children from the ED. Our data suggest that CRP and PCT were equally valuable predictors in addition to clinical signs and symptoms for predicting SBI in our population of febrile children. Using CRP and PCT in the diagnostic management of febrile children at the ED appears useful, but our findings do not support routine dual use of both CRP and PCT.

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/pr>

ACKNOWLEDGMENTS

We thank Barry Koelwijjn, Sascha Smit, members of the point of care testing team of the Erasmus MC—Sophia Children's Hospital, and all nurses at both study sites for their valuable contributions to the study. Moreover, we acknowledge Marjolein Neele for contributing to the data collection and for analyzing the laboratory samples at the Maasstad hospital. We thank Evelien Kerkhof for supporting data collection and defining final diagnoses. Finally, we acknowledge Marcel de Wilde for

designing the electronic data entry forms at the Erasmus MC. None had access to the data or received additional financial compensation. The Afinion AS100 analyzer and Afinion AS100 C-Reactive Protein bedside kits were provided by Axis-Shield PoC AS, Norway, and were distributed by Clindia Benelux BV. The procalcitonin kits were supplied by Brahms.

AUTHOR CONTRIBUTIONS

All authors substantially contributed to the writing (i.e., drafting and/or critical revision) of the manuscript. R.G.N. is the main author of the paper and was involved in development and implementation of the study's protocol, data acquisition, patient inclusion, and data analysis. H.A.M. contributed to obtaining funding for the study, protocol development, and data collection and monitoring. Y.V. was responsible for statistical analyses and interpretation of the results. F.J.S. was responsible for protocol development, data collection, and study implementation. F.W. was responsible for analyzing laboratory samples and data collection. E.W.S. contributed to obtaining funding, protocol development, statistical analyses, and interpretation of the results. J.v.d.L. contributed to obtaining funding and protocol development, supervising the data collection and monitoring, and providing technical support. Y.B.d.R. was involved in protocol development, laboratory analysis, and data collection. R.O. was mainly responsible for obtaining funding for the study, supervised the research project, and was responsible for protocol development, data collection, and interpretation of the results. All authors have read and approved the final version of the manuscript; R.G.N. and R.O. had full access to all data and take responsibility for the integrity of the data and accuracy of the data analysis.

STATEMENT OF FINANCIAL SUPPORT

R.G.N. is supported by ZonMW, a Dutch organization for health research and development, and Erasmus MC Doelmatigheid; Y.V. is funded by the Netherlands Organization for Scientific Research (Grant 917.11.383); and R. O. is supported by a fellowship grant of the European Society of Pediatric Infectious Diseases in 2010.

DISCLAIMER

Neither funding sources, nor Axis-Shield PoC AS, Clindia Benelux BV, or Brahms were involved in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclosure: The authors declare no conflict of interest.

REFERENCES

- Alpern ER, Stanley RM, Gorelick MH, et al. Epidemiology of a pediatric emergency medicine research network: the PECARN Core Data Project. *Pediatr Emerg Care* 2006;22:689–99.
- Van den Bruel A, Haj-Hassan T, Thompson M, Buntinx F, Mant D. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *Lancet* 2010;375:834–45.
- Van den Bruel A, Thompson MJ, Haj-Hassan T, et al. Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review. *Br Med J* 2011;342:d3082.
- Yo CH, Hsieh PS, Lee SH, et al. Comparison of the test characteristics of procalcitonin to c-reactive protein and leukocytosis for the detection of serious bacterial infections in children presenting with fever without source: a systematic review and meta-analysis. *Ann Emerg Med* 2012;60:591–600.
- Thompson M, Van den Bruel A, Verbakel J, et al. Systematic review and validation of prediction rules for identifying children with serious infections in emergency departments and urgent-access primary care. *Health Technol Assess* 2012;16:1–100.
- Craig JC, Williams GJ, Jones M, et al. The accuracy of clinical symptoms and signs for the diagnosis of serious bacterial infection in young febrile children: prospective cohort study of 15 781 febrile illnesses. *Br Med J* 2010;340:c1594.

7. Irwin AD, Wickenden J, Le Doare K, Ladhani S, Sharland M. Supporting decisions to increase the safe discharge of children with febrile illness from the emergency department: a systematic review and meta-analysis. *Arch Dis Child* 2015;101:259–66.
8. Nijman RG, Vergouwe Y, Thompson M, et al. Clinical prediction model to aid emergency doctors managing febrile children at risk of serious bacterial infections: diagnostic study. *Br Med J* 2013;346:f1706.
9. Maguire JL, Kulik DM, Laupacis A, Kuppermann N, Uleryk EM, Parkin PC. Clinical prediction rules for children: a systematic review. *Pediatrics* 2011;128:e666–77.
10. Gomez B, Bressan S, Mintegi S, et al. Diagnostic value of procalcitonin in well-appearing young febrile infants. *Pediatrics* 2012;130:815–22.
11. Luaces-Cubells C, Mintegi S, Garcia-Garcia JJ, et al. Procalcitonin to detect invasive bacterial infection in non-toxic-appearing infants with fever without apparent source in the emergency department. *Pediatr Infect Dis J* 2012;31:645–7.
12. van Rossum AM, Wulkan RW, Oudesluys-Murphy AM. Procalcitonin as an early marker of infection in neonates and children. *Lancet Infect Dis* 2004;4:620–30.
13. Nijman RG, Moll HA, Smit FJ, et al. C-reactive protein, procalcitonin and the lab-score for detecting serious bacterial infections in febrile children at the emergency department: a prospective observational study. *Pediatr Infect Dis J* 2014;33:e273–9.
14. Bouwhuis CB, Kromhout MM, Twijnstra MJ, Buller HA, Moll HA. Few ethnic differences in acute pediatric problems: 10 years of acute care in the Sophia Children's Hospital in Rotterdam. *Ned Tijdschr Geneesk* 2001;145:1847–51.
15. Hardelid P, Dattani N, Gilbert R. Programme Board of the Royal College of P, Child H, Child Death Overview Working G. Estimating the prevalence of chronic conditions in children who die in England, Scotland and Wales: a data linkage cohort study. *Br Med J Open* 2014;4:e005331.
16. van Veen M, Steyerberg EW, Ruige M, et al. Manchester triage system in paediatric emergency care: prospective observational study. *Br Med J* 2008;337:a1501.
17. Turner NM. *Advanced Paediatric Life Support (APLS)*, 2nd edn. Maarsse, the Netherlands: Reed Business, 2006.
18. Bonadio WA, Hennes H, Smith D, et al. Reliability of observation variables in distinguishing infectious outcome of febrile young infants. *Pediatr Infect Dis J* 1993;12:111–4.
19. McCarthy PL, Sharpe MR, Spiesel SZ, et al. Observation scales to identify serious illness in febrile children. *Pediatrics* 1982;70:802–9.
20. Bleeker SE, Derksen-Lubsen G, Grobbee DE, Donders AR, Moons KG, Moll HA. Validating and updating a prediction rule for serious bacterial infection in patients with fever without source. *Acta Paediatr* 2007;96:100–4.
21. Reitsma JB, Rutjes AW, Khan KS, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol* 2009;62:797–806.
22. R Foundation for Statistical Computing. R: a language and environment for statistical computing Vienna (Austria), 2006 (30-04-2014). (<http://www.R-project.org/>).
23. Steyerberg EW. In: M Gail KK, J Samet, A Tsiatis, W Wong, eds. *Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating*, 1st edn. Chapter 7. New York, NY, USA: Springer, 2009, pp 113–138.
24. Moons KG, Donders RA, Stijnen T, Harrell FE Jr. Using the outcome for imputation of missing predictor values was preferred. *J Clin Epidemiol* 2006;59:1092–1.
25. Marshall A, Altman DG, Holder RL, Royston P. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009;9:57.
26. Van Calster B, Van Belle V, Vergouwe Y, Timmerman D, Van Huffel S, Steyerberg EW. Extending the c-statistic to nominal polytomous outcomes: the Polytomous Discrimination Index. *Stat Med* 2012;31:2610–26.
27. Vickers AJ, Cronin AM, Begg CB. One statistical test is sufficient for assessing new predictive markers. *BMC Med Res Methodol* 2011;11:13.
28. Van Calster B, Vergouwe Y, Looman CW, Van Belle V, Timmerman D, Steyerberg EW. Assessing the discriminative ability of risk models for more than two outcome categories. *Eur J Epidemiol* 2012;27:761–70.
29. Van Hoorde K, Vergouwe Y, Timmerman D, Van Huffel S, Steyerberg EW, Van Calster B. Simple dichotomous updating methods improved the validity of polytomous prediction models. *J Clin Epidemiol* 2013;66:1158–65.
30. Milcent K, Faesch S, Gras-Le Guen C, et al. Use of procalcitonin assays to predict serious bacterial infection in young febrile Infants. *JAMA Pediatrics* 2016;170:62–9.
31. Galetto-Lacour A, Zamora SA, Andreola B, et al. Validation of a laboratory risk index score for the identification of severe bacterial infection in children with fever without source. *Arch Dis Child* 2010;95:968–73.
32. Lacour AG, Zamora SA, Gervaix A. A score identifying serious bacterial infections in children with fever without source. *Pediatr Infect Dis J* 2008;27:654–6.
33. Bressan S, Gomez B, Mintegi S, et al. Diagnostic performance of the lab-score in predicting severe and invasive bacterial infections in well-appearing young febrile infants. *Pediatr Infect Dis J* 2012;31:1239–44.
34. Mallett S, Halligan S, Thompson M, Collins GS, Altman DG. Interpreting diagnostic accuracy studies for patient care. *Br Med J* 2012;345:e3999.
35. Thayyil S, Shenoy M, Hamaluba M, Gupta A, Frater J, Verber IG. Is procalcitonin useful in early diagnosis of serious bacterial infections in children? *Acta Paediatr* 2005;94:155–8.
36. Manzano S, Bailey B, Gervaix A, Cousineau J, Delvin E, Girodias JB. Markers for bacterial infection in children with fever without source. *Arch Dis Child* 2011;96:440–6.
37. Yu CW, Juan LI, Wu MH, Shen CJ, Wu JY, Lee CC. Systematic review and meta-analysis of the diagnostic accuracy of procalcitonin, C-reactive protein and white blood cell count for suspected acute appendicitis. *Br J Surg* 2013;100:322–9.
38. Verbakel JY, Van den Bruel A, Thompson M, et al. How well do clinical prediction rules perform in identifying serious infections in acutely ill children across an international network of ambulatory care datasets? *BMC Med* 2013;11:10.
39. Vergouwe Y, Royston P, Moons KG, Altman DG. Development and validation of a prediction model with missing predictor data: a practical approach. *J Clin Epidemiol* 2010;2 :205–14.
40. McIntosh K. Community-acquired pneumonia in children. *N Engl J Med* 2002;346:429–37.
41. Wilkins TR, Wilkins RL. Clinical and radiographic evidence of pneumonia. *Radiol Technol* 2005;77:106–10.
42. Lynch T, Bialy L, Kellner JD, et al. A systematic review on the diagnosis of pediatric bacterial pneumonia: when gold is bronze. *PLoS ONE* 2010;5: e11989.
43. Nijman RG, Thompson M, van Veen M, Perera R, Moll HA, Oostenbrink R. Derivation and validation of age and temperature specific reference values and centile charts to predict lower respiratory tract infection in children with fever: prospective observational study. *Br Med J* 2012;345: e4224.
44. Elshout G, van Ierland Y, Bohnen AM, et al. Alarm signs and antibiotic prescription in febrile children in primary care: an observational cohort study. *Br J Gen Pract* 2013;63:e437–44.