# BMJ Paediatrics Open

# Update of a clinical prediction model for serious bacterial infections in preschool children by adding a host-protein-based assay: a diagnostic study

Chantal van Houten,[1] Josephine Sophia van de Maat,[2] Christiana Naaktgeboren,[3] Louis Bont,[1] R Oostenbrink[2]

[1]Paediatric infectious diseases, University Medical Centre Utrecht, Utrecht, The Netherlands
[2]General Paediatrics, Erasmus MC Sophia Childrens Hospital, Rotterdam, The Netherlands
[3]Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, The Netherlands

**Correspondence to**
Professor Louis Bont; l.bont@umcutrecht.nl

## ABSTRACT

**Objective** To determine whether updating a diagnostic prediction model by adding a combination assay (tumour necrosis factor-related apoptosis-inducing ligand, interferon γ induced protein-10 and C reactive protein (CRP)) can accurately identify children with pneumonia or other serious bacterial infections (SBIs).

**Design** Observational double-blind diagnostic study.

**Setting** Two hospitals in Israel and four hospitals in the Netherlands.

**Patients** 591 children, aged 1–60 months, presenting with lower respiratory tract infections or fever without source. 96 of them had SBIs. The original Feverkidstool, a polytomous logistic regression model including clinical variables and CRP, was recalibrated and thereafter updated by using the assay.

**Main outcome measures** Pneumonia, other SBIs or no SBI.

**Results** The recalibrated original Feverkidstool discriminated well between SBIs and viral infections, with a c-statistic for pneumonia of 0.84 (95% CI 0.77 to 0.92) and 0.82 (95% CI 0.77 to 0.86) for other SBIs. The discriminatory ability increased when CRP was replaced by the combination assay; c-statistic for pneumonia increased to 0.89 (95% CI 0.82 to 0.96) and for other SBIs to 0.91 (95% CI 0.87 to 0.94). This updated Feverkidstool improved diagnosis of SBIs mainly in children with low–moderate risk estimates of SBIs.

**Conclusion** We improved the diagnostic accuracy of the Feverkidstool by replacing CRP with a combination assay to predict pneumonia or other SBIs in febrile children. The updated Feverkidstool has the largest potential to rule out bacterial infections and thus to decrease unnecessary antibiotic prescription in children with low-to-moderate predicted risk of SBIs.

## INTRODUCTION

Suspicion of infectious disease is one of the most common causes of paediatric emergency department (ED) visits.[1] The proportion of bacterial infections in children with fever without source (FWS) and acute respiratory tract infections (RTIs) is however low (respectively 0.02%–13% and 26%–28%).[2–4] Next, identifying patients benefiting from antibiotic

### What is known about the subject?

► Identifying serious bacterial infections (SBIs) is a major diagnostic challenge.
► A combination assay (tumour necrosis factor-related apoptosis-inducing ligand, interferon γ induced protein-10 and C reactive protein) can differentiate accurately between viral and bacterial infections.
► Clinical prediction models that also include biomarkers, such as the Feverkidstool, are helpful in predicting SBIs.

### What this study adds?

► The diagnostic accuracy of the Feverkidstool increases when updated with the combination assay to predict SBIs.
► The Feverkidstool with the combination assay benefits children with low–moderate risk estimates for pneumonia or other SBIs in particular.
► Combining a clinical risk profile (like the Feverkidstool) with a new biomarker contributes to positioning that biomarker in the diagnostic workup.

treatment remains a major diagnostic challenge. Consequentially, children with acute RTI receive antibiotics almost twice as often as the estimated prevalence.[2 5 6] Antibiotic overuse is associated with increased antibiotic resistance, causing 25 000 deaths in Europe annually.[7 8] This underlines the need to better differentiate between viral and bacterial infections. Therefore, several prediction models have been developed.[9 10] The Feverkidstool, a clinical prediction model including both clinical parameters and C reactive protein (CRP), is a validated tool for supporting clinical decision-making on, for example, whether or not to start antibiotics.[11–13] However, further improvement of this diagnostic tool is warranted as it does not provide an accurate diagnosis for all patients.

We recently showed that a novel blood assay, combining concentrations of CRP with tumour necrosis factor-related apoptosis-inducing ligand (TRAIL) and interferon γ induced protein-10 (IP-10), could diagnose bacterial infections more accurately than CRP alone.[14–16]

The dynamics of TRAIL are complementary to traditionally studied bacteria-induced proteins; TRAIL concentrations decrease in bacterial infection and increase in viral infections.[15] The aim of this study is to investigate whether updating the Feverkidstool by replacing CRP with the combination assay can improve the diagnosis of serious bacterial infections (SBIs) in preschool children.

## PATIENTS AND METHODS

The current study builds on the prospective observational OPPORTUNITY Study performed in four hospitals in the Netherlands and two hospitals in Israel between 16 October 2013 and 28 January 2015. For detailed methods, we refer to the original publication.[14] In short, this study included clinical data, a host-protein-based assay, nasal swab PCR and 28-day follow-up data from children aged 1–60 months with lower RTI or FWS (n=777). The study was an observational double-blind diagnostic study. The analysis of the serum samples for the assay (index test) was done in the absence of any clinical or other patient-related information, and the expert panel (reference standard) was blinded to the decisions of their peers and to the results of the index test.

General inclusion and exclusion criteria can be found in the original publication that had 577 children in the primary analysis. We added to our current study data of children aged 1–2 months that in the original OPPORTUNITY Study were included for a subanalysis (n=28), as this study showed accurate results of the assay in this population. Children admitted to the intensive care (all referred from the wards of other hospitals, n=14) were excluded as the Feverkidstool was developed for febrile patients presenting at the ED (online supplementary figure 1). Parents gave written informed consent prior to sampling. This study was designed and analysed without patient involvement. Patients were not invited to contribute to the writing or editing of this document for readability or accuracy. This manuscript follows the TRIPOD (Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis) (online supplementary file 3) reporting guidelines.

### Prediction model

The Feverkidstool is a polytomous prediction model that predicts the risk of pneumonia or other SBIs, based on the following variables: age, body temperature, heart rate, respiratory rate, oxygen saturation, ill-appearance, peripheral capillary refill, chest wall retractions and CRP (definitions presented in the statistical analysis plan, online supplementary file 1).[12]

### Combined host-protein-based assay

The assay is currently ELISA-based (a point-of-care test is being developed) and combines the concentrations of TRAIL, IP-10 and CRP using a predetermined logistic regression formula to compute the likelihood score (scale 0–100) for a bacterial infection.[14 15] The assay was performed on coded serum samples in the absence of any patient-related information.

### Reference standard diagnosis

Currently, no single reference standard test exists for determining the aetiology of an infection.[17] Therefore, in the OPPORTUNITY Study, England's National Health Service standard for evaluating diagnostic tests was followed and an expert panel reference standard was composed.[14 18] Every recruited patient was diagnosed by three panel members affiliated to the country of recruitment using all available electronic Case Record Form (eCRF) information (clinical and laboratory information, including a 28-day follow-up), but blinded to the assay and Feverkidstool results and to the labels of their peers. Each expert assigned one of the following aetiologies to each patient: bacterial infection, viral infection, mixed infection (ie, bacterial and viral co-infection), non-infectious disease or indeterminate. Patients assigned as mixed infection were later classified as bacterial because they are clinically managed similarly. Patients with a bacterial reference standard diagnosis were divided into pneumonia or other SBIs (eg, meningitis, urinary tract infections, bacteraemia) based on the diagnosis at hospital discharge assigned by the attending physician.

### Statistical analysis

#### General approach

We compared the diagnostic accuracy to predict SBIs of the original Feverkidstool with the accuracy of the Feverkidstool updated with the assay (hereafter called updated Feverkidstool). Statistical analyses were performed in SPSS V.21.0 for Windows and R V.3.2.2. We used the Mann-Whitney U test for comparison of continuous variables. Categorical outcomes were analysed using the $\chi^2$ test or Fisher's exact test where expected cell counts were less than 5.

#### Model development and performance

First, we recalibrated the original Feverkidstool on our data using two separate logistic regressions, one for pneumonia and one for other SBIs (recalibrated Feverkidstool: *logit (pneumonia or other SBI)*= $\beta_0$ + $\beta_1$ (linear predictor Feverkidstool)). Then, we updated the Feverkidstool by adding the combination assay (updated Feverkidstool: *logit (pneumonia or other SBI)* = $\beta_0$ + $\beta_{2a}$ (linear predictor Feverkidstool) + $\beta_{2b}$ (score assay)). One element of the assay, CRP, was also a predictor in the linear predictor Feverkidstool. To be able to compare the effect of using the assay (updated Feverkidstool) instead of the CRP only (original Feverkidstool), we standardised the effect of the CRP coefficient (as part of the linear predictor Feverkidstool) in this updated Feverkidstool. Therefore the linear predictor Feverkidstool was based on the median CRP value for all participants. The actual CRP

value of patients was used to compute the score assay. Detailed models are presented in the statistical analysis plan (online supplementary file 1). The discriminative ability of the recalibrated original and updated models was expressed using pairwise c-statistics.[19]

### Predicted risk thresholds

To help interpret the potential benefit of the different models along the range of predicted probabilities, a decision curve analysis (DCA) was performed. The DCA assesses the relative harm of false positives and false negatives for different probability thresholds if the models were used to guide antibiotic prescription.[20] For predefined risk thresholds, we calculated sensitivity, specificity, positive and negative predictive values, and positive and negative likelihood ratios (LR+, LR−) for the original and updated models.

### Reclassification

To assess the potential added value of the updated Feverkidstool in correctly classifying SBIs and viral infections using defined thresholds, we did a head-to-head comparison for the updated Feverkidstool and the recalibrated original Feverkidstool in a reclassification table.[21]

### Missing values

Multiple imputation techniques enabled analysing all available data. Missing values in the variables, needed for the Feverkidstool, were imputed 10 times using the multivariate imputation by chained equations (MICE) algorithm in R statistical software.[22] Variables used in the imputation model are presented in the statistical analyses plan (online supplementary file 1). Analyses were performed separately in the 10 imputed data sets and combined using Rubin's rules.[23] The Feverkidstool variable peripheral capillary refill was not recorded in the OPPORTUNITY Study for any patient. Because we had no information on the distribution of values for capillary refill, we could not include this variable in the model for multiple imputation. Leaving the coefficient for capillary refill out of the Feverkidstool would have led to systematic lower predictions (biassed calibration), so we replaced this systematic missing variable by the mean prevalence of prolonged capillary refill in the initial Feverkidstool derivation cohort (=0.039, mean imputation).[24] Inconclusive diagnoses (n=71) were also imputed; a sensitivity analysis was performed leaving out patients with an inconclusive diagnosis.

## RESULTS
### Population characteristics

A total of 591 patients was available for analysis: 30 pneumonia, 66 other SBIs and 495 viral infections (online supplementary figure 1). Children with pneumonia were older than children with other SBIs or with viral infections (median age of 24.5 vs 15 months) and children with pneumonia or other SBIs were hospitalised more

often (73% and 77%) than children with viral infections (52%, table 1). Children with an inconclusive reference standard diagnosis differed from children with a conclusive diagnosis on age, biomarker values and antibiotic prescription (online supplementary table 1).

### Model performance

The recalibrated original Feverkidstool discriminated well between pneumonia and other infections (c-statistic 0.84, 95% CI 0.77 to 0.92), and between other SBIs and other infections (c-statistic 0.82, 95% CI 0.77 to 0.86, online supplementary figure 2). This performance is similar to previous Feverkidstool validations.[11 12] Updating the Feverkidstool with the assay improved discrimination between bacterial and other infections, reflected by an improved c-statistic for pneumonia to 0.89 (95% CI 0.82 to 0.96), and for other SBIs to 0.91 (95% CI 0.87 to 0.94) (online supplementary figure 2). A sensitivity analysis of the cohort without imputed reference standard diagnoses showed similar results, with improved prediction of pneumonia and other SBIs (online supplementary table 2).

### Predicted risk thresholds

The DCA shows the net benefit of starting antibiotics using predictions of the updated instead of the original Feverkidstool, depending on the choice of probability threshold. Children with low–moderate predicted risks ≤40% for pneumonia or other SBIs had the most benefit from the updated Feverkidstool (figure 1). For example, if all children with a predicted risk of >20% would be treated with antibiotics, the net benefit of the updated Feverkidstool increases with 0.2 from the 0.2 net benefit in the original feverkidstool for predicting pneumonia to 0.4 net benefit. For predicting other SBI the net benefit increases from 0.1 (original Feverkidstool) to 0.5 by the updated Feverkidstool. In absolute numbers this would mean that by using the updated Feverkidstool at a threshold of 20%, we achieve 20 or 50 more correct treatment decisions (out of 100 patients) than when using the original Feverkidstool. Table 2 gives more detailed insight on the effects of the updated model in diagnostic value using several thresholds. Using a rule of thumb of LR+ of 5 and LR− of 0.2,[25] thresholds of 10% and 2.5% using the updated model seem better applicable for ruling in and out of both pneumonia and other SBIs.

### Reclassification

Table 3 shows the clinical consequences if the updated Feverkidstool was used instead of the original Feverkidstool. In total, the updated Feverkidstool for pneumonia reduced the number of children with a falsely predicted high or intermediate risk from 234 to 143 (39%) compared with the original Feverkidstool, and from 427 to 197 (54%) for children with other SBIs.

**Table 1** Characteristics of patients included in the primary analysis

| | Pneumonia (n=30) | Other SBIs (n=66) | Viral (n=495) |
|---|---|---|---|
| *Predictor variables* | | | |
| **Age (months)** | 24.5 (12.7–41.3) | 15.0 (8.0–33.0) | 15.0 (7.0–28.0) |
| **Gender, male** | 19 (63%) | 32 (49%) | 280 (57%) |
| **Duration of fever (days)** | 3 (2–5) | 2 (1–4) | 2 (1–4) |
| **Temperature (°C)** | 38.6 (38.2–39.8) | 38.7 (37.8–39.4) | 38.5 (37.6–39.2) |
| | n=30 (100%) | n=66 (100%) | n=493 (99%) |
| **Respiratory rate** | 50 (34–70) | 40 (31–52) | 38 (30–52) |
| | n=15 (50%) | n=31 (47%) | n=252 (51%) |
| **Tachypnoea** | 12 (80%) | 17 (55%) | 130 (52%) |
| **Heart rate** | 160 (24) | 152 (27) | 151 (24) |
| | n=29 (97%) | n=57 (86%) | n=453 (92%) |
| **Tachycardia** | 21 (72%) | 29 (51%) | 232 (51%) |
| **Oxygen saturation (%O$_2$)** | 98 (97–99) | 99 (97–100) | 98 (96–100) |
| | n=28 (93%) | n=48 (73%) | n=417 (84%) |
| **Desaturation (<94%O$_2$)** | 1 (4%) | 0 (0%) | 24 (6%) |
| **Chest wall retractions** | 6 (20%) | 4 (6%) | 60 (12%) |
| | n=30 (100%) | n=63 (95%) | n=484 (98%) |
| **Ill appearance** | 13 (43%) | 25 (38%) | 141 (29%) |
| **C reactive protein (mg/l)** | 176 (72–224) | 102 (55–151) | 15 (5–36) |
| **Assay score** | 98 (76–100) | 88 (68–98) | 4 (1–26) |
| *Other variables* | | | |
| **Hospital admission** | 22 (73%) | 51 (77%) | 255 (52%) |
| Hospitalisation duration (days) | 4 (3–6) | 4 (3–5) | 3 (2–4) |
| **Antibiotic treatment prescribed** | 30 (100%) | 63 (96%) | 140 (29%) |
| **Recruiting site** | | | |
| Secondary care centre | 27 (90%) | 63 (96%) | 463 (94%) |
| Tertiary care centre | 3 (10%) | 3 (4%) | 32 (6%) |
| **Focus of infection** | | | |
| Central nervous system | 0 (0%) | 0 (0%) | 9 (2%) |
| Gastrointestinal tract | 0 (0%) | 1 (2%) | 19 (4%) |
| Other | 3 (10%) | 11 (17%) | 39 (8%) |
| Respiratory tract | 26 (87%) | 18 (27%) | 250 (50%) |
| Systemic | 0 (0%) | 2 (3%) | 9 (2%) |
| Unknown | 1 (3%) | 10 (15%) | 169 (34%) |
| Urinary tract | 0 (0%) | 24 (36%) | 0 (0%) |

Data are presented as n (%), median (IQR) or mean (SD). This table includes imputed reference standard diagnoses, data are based on 1 of the 10 imputed data sets. If data were not available for all patients, the total number of available data are noted. Clinical syndrome was based on the diagnosis of the attending physician at discharge from the hospital. LRTI included pneumonia and bronchiolitis; URTI included laryngitis, pharyngitis, otitis media, sinusitis and tonsillitis. LRTI: lower respiratory tract infection; URTI: upper respiratory tract infection
SBI, serious bacterial infection.

## DISCUSSION

In this study, we showed that when a clinical prediction model including CRP was updated with a combined host-protein-based assay, SBIs were predicted more accurately in children presenting with a lower RTI or FWS at the hospital. We showed that children with low-to-moderate predicted risk benefit most from this updated model.

## Strengths and weaknesses of the study

A strength of the present study is the used combination of variables. To our knowledge, this is the first study combining clinical parameters with both bacterial and viral biomarkers. A second strength is the use of an expert panel reference standard, which has the advantage of capturing a wide spectrum of illness severities, including
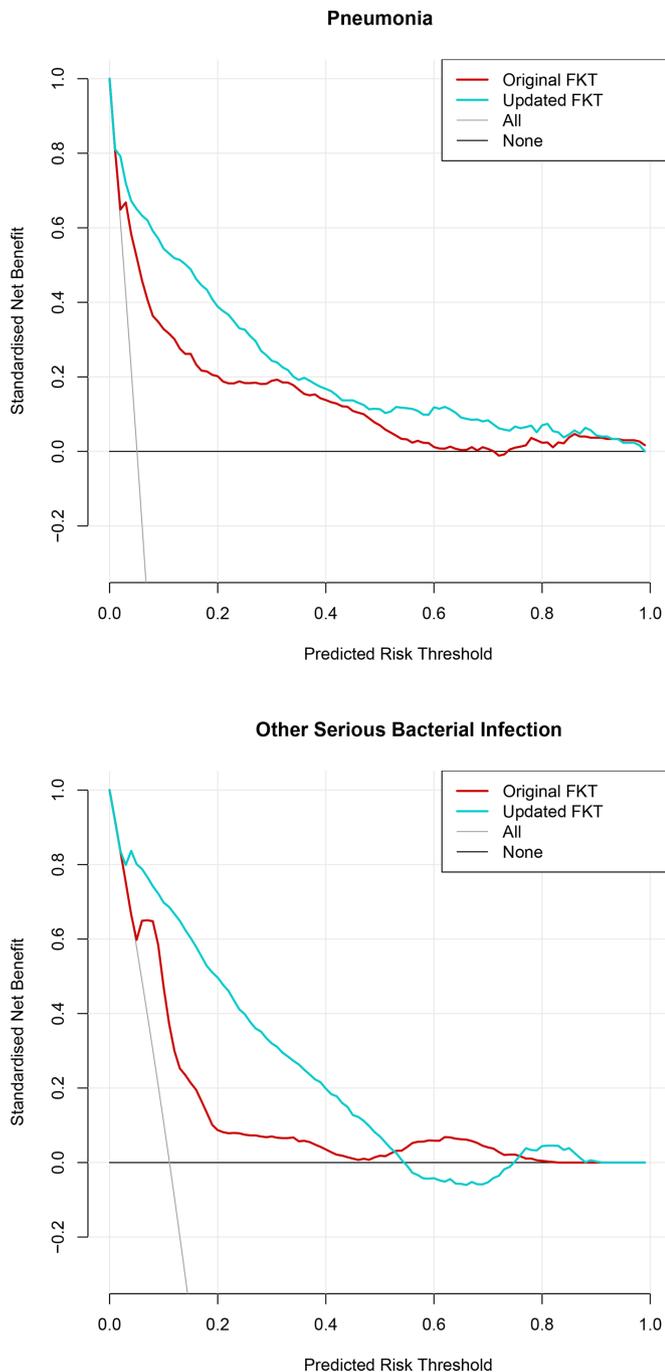
**Figure 1** Decision curve analysis with the net benefit of starting antibiotics to none of the patients (black line), to all patients (grey line), the original Feverkidstool (red line) and the updated Feverkidstool (blue line), depending on the choice of probability threshold for starting antibiotics for pneumonia (a) and other SBIs (b). FKT, Feverkidstool; SBI, serious bacterial infection.

difficult to diagnose cases, thereby reflecting the diagnostic process in clinical practice.[26] Another strength is the prospective patient recruitment. The ethics committees approved venous blood sampling even if blood sampling was not indicated for routine care. Therefore, a wide spectrum of illness severities was captured, including difficult to diagnose cases. In addition, a strength of the

Feverkidstool is its polytomous character. This enables the discrimination between different diseases: pneumonia and other SBIs versus no SBI.[27] Finally, in clinical practice different thresholds are needed depending on the setting. Therefore we performed DCA to help interpret the differences between the models along the wide range of predicted probabilities.[20]

Limitations of our study should also be addressed. First, the number of bacterial cases was relatively low. Therefore, it was not possible to refit the individual coefficients for all Feverkidstool variables. The aim of the current study was to see whether the new assay had additional value to the original well validated Feverkidstool rather than to build an optimal diagnostic model. Second, one of the Feverkidstool variables, capillary refill, was not available for any of the participants. Therefore, multiple imputation was impossible. Leaving capillary refill out by entering 0 for all capillary values would have resulted in an unfair reduction of model calibration. We think imputation of the mean was the best available option.[24] Even though this may have limited the discriminative performance of the model, we believe the influence is limited, because the dichotomous variable capillary refill has little diagnostic value within the Feverkidstool. In addition, the aim of the current study was to compare the original Feverkidstool with the updated model. Imputation of capillary refill values was performed similarly for both models, so this had no influence on the comparison of the models. Third, the reclassification is based on arbitrarily chosen thresholds (2.5% and 10%). These thresholds, however, mostly correspond with LR that have been reported to be meaningful in decisions for febrile children: LR+>5.0 for ruling in SBIs and LR−<0.2 for ruling out SBIs.[25] Fourth, the Feverkidstool includes important clinical variables that are used by every physician when deciding whether to start antibiotics or not. Following clinical care, during the expert panel reference standard process, the panel was provided with a wide range of clinical information, including the clinical Feverkidstool variables, but they were not informed about the algorithm. We, however, cannot exclude that incorporation bias—in which part of the test being assessed is included in the reference standard—has resulted in some degree of overestimation of the outcome. Fifth, due to the relatively low number of bacterial cases, some diseases, such as bacterial meningitis, were not observed in our study cohort. Information to discriminate sepsis from bacteraemia cases was not available. Sixth, in the current study, none of the ED patients were hospitalised at the intensive care unit. Further studies with higher numbers of patients with a more severe clinical presentation are warranted as especially in these patients early detection of a SBI can be critical. Seventh, in this study we wanted to study the added value of the assay on top of the model that had also undergone impact analysis[28] (ie, the original Feverkidstool (including CRP)), rather than developing a new model. As both the assay and the original Feverkidstool contain CRP, we had to solve this. Refitting the Feverkidstool's model without CRP will not provide the answer

**Table 2** Diagnostic performance of prediction models including clinical signs and symptoms and either C reactive protein or the assay

| | Sensitivity (95% CI) | Specificity (95% CI) | PPV (95% CI) | NPV (95% CI) | LR+ (95% CI) | LR− (95% CI) |
|---|---|---|---|---|---|---|
| **Original Feverkidstool** | | | | | | |
| *Pneumonia* | | | | | | |
| Risk of ≥2.5% | 0.85 (0.77 to 0.94) | 0.58 (0.54 to 0.63) | 0.10 (0.06 to 0.14) | 0.99 (0.98 to 0.99) | 2.06 (1.39 to 3.07) | 0.25 (0.08 to 0.76) |
| Risk of ≥5% | 0.81 (0.71 to 0.91) | 0.76 (0.73 to 0.79) | 0.15 (0.09 to 0.22) | 0.99 (0.98 to 0.99) | 3.41 (2.23 to 5.20) | 0.24 (0.10 to 0.62) |
| Risk of ≥10% | 0.62 (0.47 to 0.78) | 0.89 (0.87 to 0.92) | 0.24 (0.13 to 0.34) | 0.98 (0.97 to 0.99) | 5.83 (3.90 to 8.71) | 0.42 (0.25 to 0.69) |
| Risk of ≥15% | 0.48 (0.31 to 0.65) | 0.93 (0.91 to 0.95) | 0.27 (0.14 to 0.40) | 0.97 (0.96 to 0.98) | 6.80 (3.87 to 11.97) | 0.56 (0.37 to 0.84) |
| Risk of ≥20% | 0.40 (0.23 to 0.58) | 0.95 (0.94 to 0.97) | 0.31 (0.16 to 0.47) | 0.97 (0.96 to 0.98) | 8.52 (4.57 to 15.90) | 0.63 (0.45 to 0.87) |
| Risk of ≥30% | 0.27 (0.10 to 0.45) | 0.99 (0.98 to 1.00) | 0.60 (0.39 to 0.81) | 0.96 (0.95 to 0.98) | 28.09 (8.16 to 96.67) | 0.73 (0.57 to 0.94) |
| *Other SBIs* | | | | | | |
| Risk of ≥2.5% | 0.98 (0.97 to 1.00) | 0.19 (0.15 to 0.22) | 0.13 (0.10 to 0.16) | 0.99 (0.98 to 1.00) | 1.21 (1.13 to 1.30) | 0.09 (0.01 to 0.65) |
| Risk of ≥5% | 0.96 (0.94 to 0.99) | 0.41 (0.37 to 0.46) | 0.17 (0.13 to 0.21) | 0.99 (0.98 to 1.00) | 1.64 (1.48 to 1.83) | 0.08 (0.02 to 0.43) |
| Risk of ≥10% | 0.85 (0.78 to 0.92) | 0.70 (0.67 to 0.74) | 0.26 (0.20 to 0.33) | 0.98 (0.96 to 0.99) | 2.87 (2.37 to 3.48) | 0.21 (0.10 to 0.42) |
| Risk of ≥15% | 0.54 (0.42 to 0.66) | 0.82 (0.79 to 0.85) | 0.27 (0.19 to 0.36) | 0.94 (0.92 to 0.95) | 3.03 (2.24 to 4.12) | 0.56 (0.41 to 0.76) |
| Risk of ≥20% | 0.38 (0.26 to 0.50) | 0.89 (0.86 to 0.91) | 0.30 (0.19 to 0.40) | 0.92 (0.90 to 0.94) | 3.37 (2.23 to 5.09) | 0.70 (0.57 to 0.86) |
| Risk of ≥30% | 0.25 (0.13 to 0.36) | 0.94 (0.92 to 0.95) | 0.32 (0.18 to 0.46) | 0.91 (0.89 to 0.93) | 3.82 (2.15 to 6.77) | 0.81 (0.69 to 0.93) |
| **Updated Feverkidstool** | | | | | | |
| *Pneumonia* | | | | | | |
| Risk of ≥2.5% | 0.88 (0.80 to 0.95) | 0.75 (0.71 to 0.78) | 0.16 (0.09 to 0.22) | 0.99 (0.99 to 1.00) | 3.48 (2.39 to 5.07) | 0.16 (0.05 to 0.49) |
| Risk of ≥5% | 0.85 (0.76 to 0.94) | 0.83 (0.80 to 0.86) | 0.21 (0.13 to 0.29) | 0.99 (0.99 to 1.00) | 4.93 (3.55 to 6.84) | 0.18 (0.07 to 0.49) |
| Risk of ≥10% | 0.76 (0.64 to 0.88) | 0.89 (0.86 to 0.91) | 0.27 (0.17 to 0.37) | 0.99 (0.98 to 0.99) | 6.84 (4.49 to 10.41) | 0.27 (0.14 to 0.53) |
| Risk of ≥15% | 0.72 (0.60 to 0.85) | 0.92 (0.90 to 0.94) | 0.33 (0.21 to 0.45) | 0.98 (0.98 to 0.99) | 9.27 (5.47 to 15.72) | 0.30 (0.15 to 0.58) |
| Risk of ≥20% | 0.66 (0.52 to 0.80) | 0.94 (0.93 to 0.96) | 0.39 (0.25 to 0.53) | 0.98 (0.97 to 0.99) | 11.88 (7.07 to 19.97) | 0.36 (0.20 to 0.63) |
| Risk of ≥30% | 0.50 (0.34 to 0.67) | 0.97 (0.96 to 0.98) | 0.48 (0.31 to 0.65) | 0.97 (0.96 to 0.98) | 17.45 (8.88 to 34.27) | 0.50 (0.31 to 0.82) |
| *Other SBIs* | | | | | | |
| Risk of ≥2.5% | 0.97 (0.94 to 0.99) | 0.62 (0.58 to 0.66) | 0.24 (0.19 to 0.30) | 0.99 (0.99 to 1.00) | 2.57 (2.22 to 2.99) | 0.05 (0.01 to 0.26) |
| Risk of ≥5% | 0.92 (0.88 to 0.97) | 0.75 (0.72 to 0.79) | 0.32 (0.25 to 0.39) | 0.99 (0.98 to 1.00) | 3.71 (3.04 to 4.51) | 0.10 (0.04 to 0.26) |
| Risk of ≥10% | 0.88 (0.82 to 0.94) | 0.83 (0.80 to 0.86) | 0.39 (0.31 to 0.47) | 0.98 (0.97 to 0.99) | 5.05 (4.04 to 6.32) | 0.15 (0.07 to 0.30) |
| Risk of ≥15% | 0.84 (0.77 to 0.91) | 0.86 (0.84 to 0.89) | 0.43 (0.35 to 0.52) | 0.98 (0.97 to 0.99) | 6.16 (4.78 to 7.93) | 0.18 (0.10 to 0.34) |
| Risk of ≥20% | 0.80 (0.72 to 0.88) | 0.89 (0.86 to 0.91) | 0.46 (0.37 to 0.56) | 0.97 (0.96 to 0.98) | 6.95 (5.27 to 9.17) | 0.22 (0.14 to 0.37) |
| Risk of ≥30% | 0.67 (0.57 to 0.77) | 0.91 (0.89 to 0.93) | 0.49 (0.39 to 0.59) | 0.96 (0.94 to 0.97) | 7.70 (5.49 to 10.82) | 0.36 (0.25 to 0.52) |

LR+, positive likelihood ratio; LR−, negative likelihood ratio; NPV, negative predictive value; PPV, positive predictive value; SBI, serious bacterial infection.

**Table 3** Diagnostic reclassification by the updated Feverkidstool compared with the original Feverkidstool

| a. | | Reference standard diagnosis: Pneumonia (n=30) | | | Reference standard diagnosis: No pneumonia (n=561)* | | |
|---|---|---|---|---|---|---|---|
| | | Updated Feverkidstool predicted risk for pneumonia | | | Updated Feverkidstool predicted risk for pneumonia | | |
| | | Low | Intermediate | High | Low | Intermediate | High |
| *Original Feverkidstool predicted risk for pneumonia‡* | Low | 3 | 1 | 1 | 304 | 20 | 3 |
| | Intermediate | 1 | 2 | 4 | 100 | 46 | 27 |
| | High | 0 | 0 | 18 | 14 | 14 | 33 |

| b. | | Reference standard diagnosis: Other SBIs (n=66) | | | Reference standard diagnosis: No other SBIs (n=525)† | | |
|---|---|---|---|---|---|---|---|
| | | Updated Feverkidstool predicted risk for other SBIs | | | Updated Feverkidstool predicted risk for other SBIs | | |
| | | Low | Intermediate | High | Low | Intermediate | High |
| *Original Feverkidstool predicted risk for other SBIs§* | Low | 1 | 0 | 0 | 98 | 0 | 0 |
| | Intermediate | 1 | 2 | 6 | 190 | 55 | 25 |
| | High | 0 | 4 | 52 | 40 | 51 | 66 |

The Feverkidstool classification was compared with the majority reference standard for the prediction of pneumonia (A) and other SBIs (B). Head-to-head comparison of the original Feverkidstool and the updated Feverkidstool. Predicted risks of pneumonia (A) and other SBIs (B) were low if the predicted risk was below 2.5%, intermediate between 2.5% and 10%, and high above 10%.

*As the comparison in this table is between the presence or absence of pneumonia, it should be noted that 'no pneumonia' includes viral and other SBIs.

†As the comparison in this table is between the presence or absence of other SBIs, it should be noted that 'no other SBIs' includes viral and pneumonia.

‡Among the patients with pneumonia as determined by the majority in the expert panel, six patients were correctly reclassified as being at higher risk using the updated Feverkidstool instead of the original Feverkidstool and one patient was incorrectly reclassified. For patients in whom pneumonia was absent, these numbers are respectively 128 and 50. Reclassification improvement was 17% for patients with pneumonia (6 minus 1 of 30) and 14% for patients without pneumonia (128 minus 50 of 561).

§Among the patients with other SBIs as determined by the majority in the expert panel, six patients were correctly reclassified using the updated Feverkidstool instead of the original Feverkidstool and five patients were incorrectly reclassified. For patients in whom other SBIs are absent, these numbers are respectively 281 and 25. Reclassification improvement was 2% for patients with other SBIs (6 minus 5 of total 66) and 49% for patients without other SBIs (281 minus 25 of 525).

SBI, serious bacterial infection.

to our research question on the *added* value of the assay to the Feverkidstool. In addition, the number of bacterial cases was too low to refit the original Feverkidstool. We also did not want to simply add the assay as a predictor in the Feverkidstool model, as this would downwardly bias the assay's added value. Therefore, we essentially removed the effect of the CRP in the Feverkidstool by assigning the median value of CRP to the CRP single predictor in the Feverkidstool to all patients. To our knowledge, using the median CRP value was the best possible approach to avoid double counting of CRP. Finally, there were 71 patients for whom the expert panel could not assign a final diagnosis. Such inconclusive cases are inherent to studies using outcomes lacking a gold standard. To make optimal use of the data from all recruited patients, we have imputed these reference standard diagnoses. As the imputed diagnoses are used for both the original and the updated Feverkidstool, we do not expect this to have influenced the results of updating the Feverkidstool. In addition, a sensitivity analysis in which all cases with imputed diagnoses were excluded showed comparable results.

**Comparison with existing literature**

The study in which the Feverkidstool was developed and the OPPORTUNITY Study had comparable inclusion criteria. Most importantly, both studies included children suspected of infection based on increased temperature.[12 14] Differences in inclusion criteria should also be discussed. The Feverkidstool derivation cohort included Dutch children aged 1 month to 15 years, whereas the OPPORTUNITY Study included Dutch and Israeli children aged 1 month to 5 years. In addition, for the Feverkidstool development study, children who received antibiotics before the ED visit were excluded; for the OPPORTUNITY Study antibiotic use was no exclusion criterion.[12 14] As mentioned/argued in the methods, we additionally excluded children admitted to intensive care (all referred from wards of other hospitals, n=14), and added 28 children aged 1–2 months from the OPPORTUNITY subanalysis.

We recently confirmed the external validity of the Feverkidstool, but when procalcitonin (PCT) was added to the prediction model or when CRP was replaced by PCT the accuracy for predicting SBIs in febrile children did not improve.[11] Another study updated the Feverkidstool by adding PCT and resistin, resulting in minimal changes in LR+ and LR− for different risk thresholds.[13] In contrast to the two abovementioned Feverkidstool updates, our current study has shown that updating the Feverkidstool

with the assay does meaningfully improve the accuracy of the model. This further confirms our previous observation that the combination of CRP, TRAIL and IP-10 had higher diagnostic value in differentiating between bacterial and viral infections compared with PCT or CRP alone.[14]

### Implications for clinical practice

Clinical signs and symptoms play an important role when physicians diagnose febrile children, but do not sufficiently differentiate between viral and bacterial infections. Therefore, the use of diagnostic prediction models that include both clinical parameters and biomarkers is intuitive and helpful. The Feverkidstool provides two risk percentages; the risk for having pneumonia and the risk for having other SBIs. After further validation, a digital calculator should be constructed to facilitate potential clinical use. Both the assay and the Feverkidstool are designed to predict the outcome 'bacterial infection', and not the general level of illness of the child. Therefore, the outcome can be used to guide decisions on starting antibiotics or not, but cannot be used to guide decisions on how to administrate antibiotics or what type. We showed that the updated Feverkidstool has the most added value for patients in the low–moderate risk group, with predicted risk for SBIs below 40%. At thresholds of 2.5% and 10%, the reclassification table showed substantial improvement in diagnosis. The cases with predicted risks between low to moderate (2.5 vs 40%) may be characterised by having intermediate values of CRP with higher diagnostic uncertainty. Adding two viral biomarkers to the prediction rule will provide an extra dimension to the model and therefore improve diagnosis especially for those cases. As a point-of-care test is under development, the manufacturer has not given an indication to the eventual cost yet. To optimise cost-effective use of the combination test, our results suggest that the added value of the assay is the highest in children with a predicted risk <40% as predicted by the original Feverkidstool. Utility studies are needed to determine the cut-off for the best clinical utility and cost-effectiveness.

### Implications for future research

Since we have proven the accuracy of the updated Feverkidstool, the next step is to perform a prospective external validation and to evaluate its impact on resource use and antibiotic treatment. An important aspect is to define risk cut-offs for different settings in clinical practice, for example, young children and children with comorbidities. In addition, in order to optimise resource use, new biomarkers may benefit selected patient subgroups in particular (eg, selected on a set of clinical characteristics/ predicted risk) rather than in all febrile children. This targeted risk approach may also be applied to position the role of, for example, myxovirus resistance protein A and CRP.[29]

### CONCLUSION

In conclusion, a new blood assay including viral and bacterial biomarkers, combined with a clinical prediction model is, in this study cohort, superior to the model with CRP only for predicting SBIs in preschool children. In children with low-to-moderate predicted risk of SBIs in particular, the updated Feverkidstool with the assay has the potential to optimise targeted antibiotic prescription and to prevent unnecessary use of antibiotics.

### REFERENCES

1. 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.
2. Kronman MP, Zhou C, Mangione-Smith R. Bacterial prevalence and antimicrobial prescribing trends for acute respiratory tract infections. *Pediatrics* 2014;134:e956–65.
3. Chancey RJ, Jhaveri R. Fever without localizing signs in children: a review in the post-Hib and postpneumococcal era. *Minerva Pediatr* 2009;61:489–501.
4. Hubert-Dibon G, Danjou L, Feildel-Fournial C, *et al*. Procalcitonin and C-reactive protein may help to detect invasive bacterial infections in children who have fever without source. *Acta Paediatr* 2018;107:1262–9.
5. Van Boeckel TP, Gandra S, Ashok A, *et al*. Global antibiotic consumption 2000 to 2010: an analysis of national pharmaceutical sales data. *Lancet Infect Dis* 2014;14:742–50.
6. Kraus EM, Pelzl S, Szecsenyi J, *et al*. Antibiotic prescribing for acute lower respiratory tract infections (LRTI) – guideline adherence in the German primary care setting: an analysis of routine data. *PLoS One* 2017;12:e0174584.
7. Laxminarayan R, Duse A, Wattal C, *et al*. Antibiotic resistance—the need for global solutions. *Lancet Infect Dis* 2013;13:1057–98.
8. ECDC/EMEA. Technical report. the bacterial challenge: time to react, Sept 2009. Available: http://ecdc.europa.eu/en/publications/Publications/0909_TER_The_Bacterial_Challenge_Time_to_React.pdf
9. 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. *BMJ* 2010;340:c1594.
10. 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.
11. Nijman RG, Vergouwe Y, Moll HA, *et al*. Validation of the Feverkidstool and procalcitonin for detecting serious bacterial infections in febrile children. *Pediatr Res* 2017.
12. 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. *BMJ* 2013;346:f1706.

13. Irwin AD, Grant A, Williams R, *et al*. Predicting risk of serious bacterial infections in febrile children in the emergency department. *Pediatrics* 2017;140:e20162853.

14. van Houten CB, de Groot JAH, Klein A, *et al*. A host-protein based assay to differentiate between bacterial and viral infections in preschool children (opportunity): a double-blind, multicentre, validation study. *Lancet Infect Dis* 2017;17:431–40.

15. Oved K, Cohen A, Boico O, *et al*. A novel Host-Proteome signature for distinguishing between acute bacterial and viral infections. *PLoS One* 2015;10:e0120012.

16. Srugo I, Klein A, Stein M, *et al*. Validation of a novel assay to distinguish bacterial and viral infections. *Pediatrics* 2017;140:e20163453.

17. Bertens LCM, Broekhuizen BDL, Naaktgeboren CA, *et al*. Use of expert panels to define the reference standard in diagnostic research: a systematic review of published methods and reporting. *PLoS Med* 2013;10:e1001531.

18. Rutjes A, Reitsma J, Coomarasamy A, *et al*. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technol Assess* 2007;11:iii, ix–51.

19. Steyerberg E. *Clinical prediction models. A practical approach to development, validation and updating*. New York: Springer, 2009.

20. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 2016;352.

21. Pencina MJ, D'Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med* 2011;30:11–21.

22. Buuren Svan, Groothuis-Oudshoorn K. Mice: Multivariate Imputation by Chained Equations in R. *J Stat Softw* 2011;45:1–67.

23. Marshall A, Altman DG, Holder RL, *et al*. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol* 2009;9:57.

24. Janssen KJM, Vergouwe Y, Donders ART, *et al*. Dealing with missing predictor values when applying clinical prediction models. *Clin Chem* 2009;55:994–1001.

25. 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.

26. Reitsma JB, Rutjes AWS, Khan KS, *et al*. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol* 2009;62:797–806.

27. Biesheuvel CJ, Vergouwe Y, Steyerberg EW, *et al*. Polytomous logistic regression analysis could be applied more often in diagnostic research. *J Clin Epidemiol* 2008;61:125–34.

28. de Vos-Kerkhof E, Nijman RG, Vergouwe Y, *et al*. Impact of a clinical decision model for febrile children at risk for serious bacterial infections at the emergency department: a randomized controlled trial. *PLoS One* 2015;10:e0127620.

29. Engelmann I, Dubos F, Lobert P-E, *et al*. Diagnosis of viral infections using myxovirus resistance protein A (MxA). *Pediatrics* 2015;135:e985–93.