BMJ Open

Seeking diagnostic and prognostic biomarkers for childhood bacterial pneumonia in sub-Saharan Africa: study protocol for an observational study

Clarissa Valim [1], Yekin Ajauoi Olatunji,[2] Yasir Shitu Isa,[2] Rasheed Salaudeen,[2] Sarwar Golam,[2] Edward F Knol,[3] Sheriffo Kanyi,[4] Abdoulie Jammeh,[5] Quique Bassat,[6,7] Wilco de Jager,[3,8] Alejandro A Diaz,[9,10] Roger C Wiegand,[11] Julio Ramirez,[12] Marsha A Moses,[13,14] Umberto D'Alessandro,[15,16] Patricia L Hibberd,[17] Grant A Mackenzie [2,18]

For numbered affiliations see end of article.

Correspondence to
Dr Clarissa Valim;
cvalim@bu.edu

ABSTRACT

Introduction Clinically diagnosed pneumonia in children is a leading cause of paediatric hospitalisation and mortality. The aetiology is usually bacterial or viral, but malaria can cause a syndrome indistinguishable from clinical pneumonia. There is no method with high sensitivity to detect a bacterial infection in these patients and, as result, antibiotics are frequently overprescribed. Conversely, unrecognised concomitant bacterial infection in patients with malarial infections occur with omission of antibiotic therapy from patients with bacterial infections. Previously, we identified two combinations of blood proteins with 96% sensitivity and 86% specificity for detecting bacterial disease. The current project aimed to validate and improve these combinations by evaluating additional biomarkers in paediatric patients with clinical pneumonia. Our goal was to describe combinations of a limited number of proteins with high sensitivity and specificity for bacterial infection to be incorporated in future point-of-care tests. Furthermore, we seek to explore signatures to prognosticate clinical pneumonia.

Methods and analysis Patients (n=900) aged 2–59 months presenting with clinical pneumonia at two Gambian hospitals will be enrolled and classified according to criteria for definitive bacterial aetiology (based on microbiological tests and chest radiographs). We will measure proteins at admission using Luminex-based immunoassays in 90 children with definitive and 160 with probable bacterial aetiology, and 160 children classified according to the prognosis of their disease. Previously identified diagnostic signatures will be assessed through accuracy measures. Moreover, we will seek new diagnostic and prognostic signatures through machine learning methods, including support vector machine, penalised regression and classification trees.

Ethics and dissemination Ethics approval has been obtained from the Gambia Government/Medical Research Council Unit The Gambia Joint Ethics Committee (protocol 1616) and the institutional review board of Boston University Medical Centre (STUDY00000958). Study results will be disseminated to the staff of the study hospitals, in scientific seminars and meetings, and in publications.

Strengths and limitations of this study

► This study will be conducted in a resource-limited country for future development of a point-of-care test for these settings.

► We will quantify known and recently identified biomarkers that will be combined using different machine learning approaches.

► Children will be followed up for 30 days after discharge to yield mid-term prognostic assessment of the clinical pneumonia.

► A difficulty shared by studies aiming to identify biomarkers to diagnose bacterial infection is lack of an optimal gold standard for bacterial infection; thus, we will use a composite criterion based on microbiological and radiological evidence of infection (primary criteria) and probable criteria based on expert evaluation of clinical and laboratory data.

► Although the sample size is relatively small, results will establish a benchmark against which the identified signatures in populations in other areas and in different subgroups, such as children with undernutrition, can be compared.

Trial registration number H-38462.

INTRODUCTION

Clinical pneumonia may be caused by viral or bacterial infections[1] and, where malaria is endemic, malaria can cause a syndrome indistinguishable from clinical pneumonia.[2] There has been a growing number of publications on biomarkers for the diagnosis of bacterial infections and to prognosticate outcomes in patients with clinical pneumonia.[3–37] Such increased attention is driven by the need for low cost point-of-care (POC) tests able to accurately, reliably and promptly diagnose bacterial infections that can guide

treatment.[38–41] All existing tests have delayed turnaround times or require specialised facilities that are often unavailable in resource-limited settings. POC tests can support timely decisions about treatment and ameliorate the overprescription of antibiotics to treat viral infections, thus slowing the progression of global antibiotic resistance.[42–46] In malaria endemic areas, antibiotics may also be 'underprescribed' and children with bacterial pneumonia may be mistakenly treated for malaria and not receive antibiotic therapy.[43 47] Prognostic POCs could also guide referrals and allow prioritisation of intensive treatment measures that are particularly scarce in resource-limited settings.

Symptoms of bacterial and viral diseases and of malaria often overlap.[48 49] Radiographical evidence of pneumonia has been used as an endpoint in studies of pneumococcal vaccine, but chest X-rays, particularly in resource-limited settings, when available have moderate reliability[50–58] and may result in both false-positive and false-negative results. Laboratory tests have low accuracy. Blood or pleural fluid cultures are highly specific for diagnosing bacterial infections but have low sensitivity, and results are not promptly available.[59 60] Bacterial PCR tests require specialised resources and have low sensitivity. Viral PCR also requires specialised facilities and does not establish a diagnosis of viral pneumonia.[49 61–63] Furthermore, in children, antigen detection in the urine has limited utility, and induced sputum samples are difficult to obtain and likely to be contaminated with the normal flora.[64–67] Where malaria is endemic, standard malaria diagnostic tests can detect the parasite in blood. Nevertheless, malaria may coexist with other infections while not causing the respiratory syndrome.[68 69]

Generally, previous biomarker studies to diagnose bacterial diseases or to prognosticate disease in patients with pneumonia or sepsis focused on procalcitonin (PCT) and/or C reactive protein (CRP) and concluded that the markers had moderate accuracy (area under the receiver operating characteristic curve (AUC-ROC)<0.80).[9–14 34–36 70–93] Accuracy of a single biomarker in bacterial pneumonia is probably low since many biomarkers may be in the inflammatory pathway triggered by several pathogens.[94] More recently, studies screened through novel biomarkers and identified gene transcription signatures able to distinguish patients with bacterial and viral infections.[95 96] Additionally, a biomarker signature based on haptoglobin (HP) and lipocalin (NGAL) has been identified to differentiate malaria and probable bacteria aetiology in paediatric patients with clinical pneumonia in a resource-limited country.[97] Two other promising signatures to differentiate viral from bacterial infections in high-income countries (HICs) have been identified. The first (ImmunoXpert) is based on CRP, tumour necrosis factor-related apoptosis-inducing ligand (TRAIL) and interferon gamma-induced protein 10 (IP-10) and has been validated in HICs, with sensitivity ranging from 87% to 94% and specificity from 90% to 94%.[5 98–101] The second is based on CRP and Mixovirus Resistance Protein A and has been implemented in a POC (FebriDx) in HICs, with sensitivity and specificity varying from 80% to 96% and from 67% to

94%, respectively.[4 102 103] Although promising, these two signatures need to be validated in resource-limited settings where inflammatory responses may vary because of malaria, coinfections and undernutrition.[104]

In Mozambique, we scanned through several biomarkers and combined them through machine learning methods, aiming to detect bacterial infection (vs no bacterial infection) in paediatric patients with clinical pneumonia. Considering the potential of coexisting malaria infections, we sought biomarker signatures by comparing children with malaria, bacterial and viral infections and validated accurate signatures by comparing patients with and without bacterial infections.[105] Signatures with 3–5 proteins were identified that had high sensitivity for the diagnosis of bacterial pneumonia (96%–91%) while misdiagnosing few viral cases. The selected signatures performed better than signatures including commonly collected clinical and laboratory markers, as well as pneumonia with consolidation on X-ray. The current study builds on our previous findings and aims to validate previous diagnostic signatures identified by us and other groups, in addition to improving on signatures by scanning through novel biomarkers in a different population in rural Gambia.

Biomarkers for signatures will be selected and combined using a comprehensive set of data mining approaches. To validate and improve on prior signatures and seek prognostic signatures, we will add some novel promising biomarkers such as resistin and lactoferrin.[37 106 107] At the end of this study, we expect to select signatures with excellent accuracy (≥93% sensitivity) that have the highest prognostic utility and composed of a limited number of biomarkers that can be developed in a POC test. Our focus is on identifying patients who need antibiotic therapy for bacterial diseases.

Study objectives
Specifically, this study aimed to

► Assess combinations of blood proteins to accurately and reliably diagnose bacterial pneumonia in paediatric patients.
  We expect that protein combinations that were identified in previous studies as accurate for diagnosing bacterial infection in patients with clinical pneumonia will have high sensitivity, specificity and AUC-ROC in an African paediatric population different from the Mozambican children. Biomarkers of these signatures will be absent in healthy children. Moreover, we hypothesise that accuracy of previous signatures can be improved by seeking to combine them with novel candidate proteins.
► Explore prognostic biomarker signatures in patients with clinical pneumonia caused by any pathogen.
  We anticipate that combination of inflammatory proteins in the blood measured at admission will accurately identify patients who will clinically deteriorate.

METHODS AND ANALYSIS

Design

To test and identify existing and novel diagnostic biomarker signatures for bacterial disease, we will measure proteins in the blood at admission and collect clinical and laboratory information at admission to ascertain diagnostic groups. Diagnostic signatures will be sought by comparing children in the target diagnosis groups. To seek prognostic biomarker signatures, we will prospectively follow clinical outcomes collected after admission and discharge. When studying prognostic signatures in the primary analysis, proteins will be compared across three groups with poor, moderate and good prognoses. The study started in February of 2019 and is expected to be concluded by November 2021.

Study environment

The study will be conducted at the Boston University School of Public Health (BUSPH), the Medical Research Council Unit The Gambia (MRCG) at London School of Hygiene & Tropical Medicine, and the Multiplex Core Facility of the Centre of Translational Immunology from the University Medical Centre Utrecht. Patients will be recruited at Basse and Bansang Hospitals. These hospitals serve a population of 178 510 (224 villages) and 99 113 (217 villages), respectively, with 19% of the population ag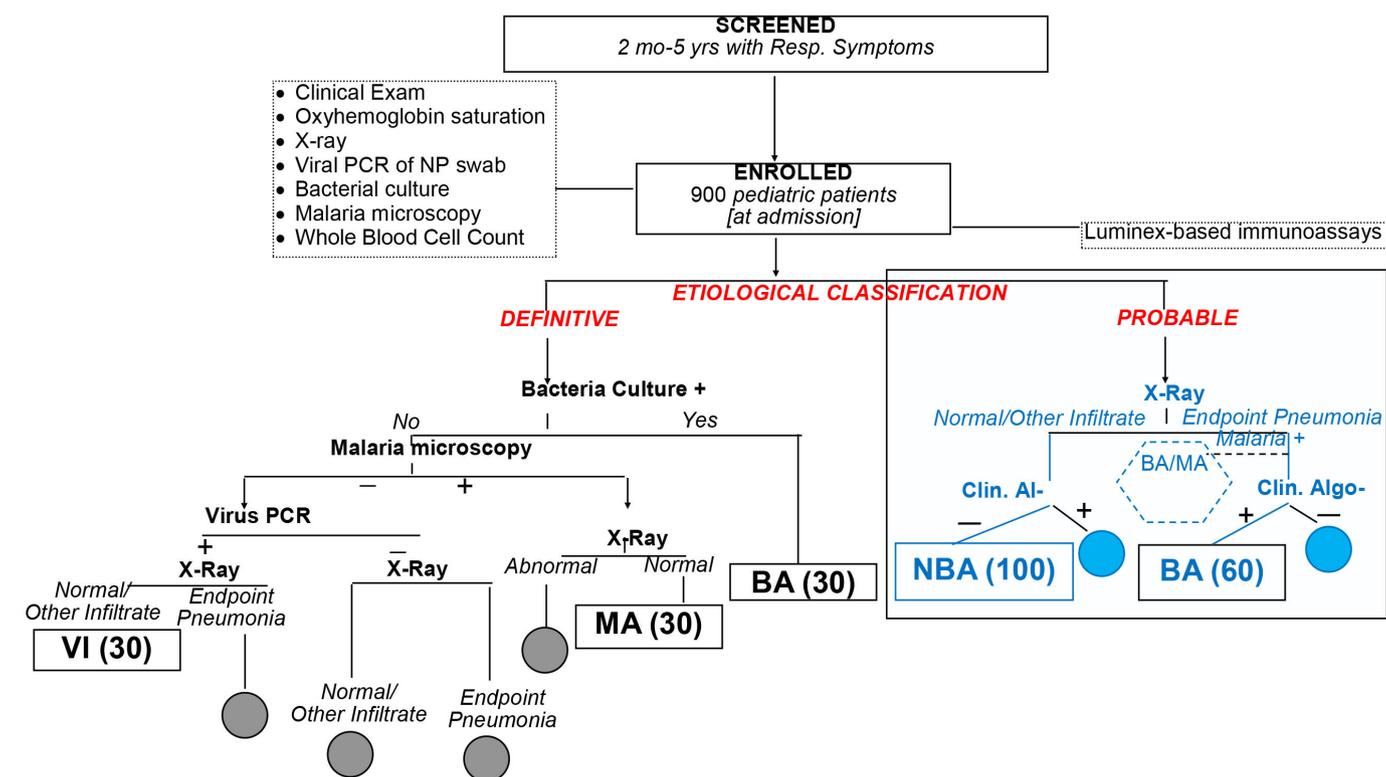ed <5 years. Diagnostic laboratory tests will be conducted at the MRCG, Basse Field Station. The multiplex bead-based immunoassay will be conducted at the Center for Translational Immunology/Utrecht Medical Center and analysis will be conducted at BUSPH.

Outcomes

Two sets of outcomes will be evaluated, one for each study aim.

Objective 1: diagnostic group classification

The primary objective of the diagnostic group classification was to identify patients in need of antibiotic therapy for a bacterial infection and not to identify the pathogen responsible for the respiratory symptoms. Patients will be assigned to bacterial, malarial and viral groups according to a definitive classification (primary analysis) (figure 1) when seeking signatures, and bacterial and non-bacterial infection groups when testing signatures. Patients will also be categorised into bacterial and non-bacterial infection groups based on a probable classification criterion (secondary analysis). Patients with ambiguous diagnosis who do not meet our criteria will be excluded from the study of objective 1. By incorporating the two outcomes, we will assess the validity of the identified biomarker signatures in patients with and without detected bacteraemia and provide data to support the generalisability of the inflammatory biomarker signature.

Figure 1 Definitive and probable diagnostic classification criteria (gold standard of diagnostic signature aim) to assign patients to the MA, BA and VI groups (definitive composite criteria) and to the BA and NBA groups. BA, bacterial; MA, malaria; NBA, non-bacterial; NP, nasopharyngeal; VI, viral.

For the definitive classification, a patient will be assigned to the bacterial (BA) group if he/she has a positive bacterial pathogen culture of fluid from a normally sterile site (eg, blood or pleural fluid). Patients will be assigned to the viral (VI) group if they have negative bacteria microbiological tests, negative malaria blood slides, X-ray without 'endpoint pneumonia' (consolidation or pleural effusion[51]), no evidence of fungal infection, and positive PCR for a viral pathogen from nasopharyngeal swabs (NPS). Patients will be assigned to the malarial (MA) group if they have normal X-ray (with neither infiltrates nor endpoint pneumonia), no bacterial infection and >0/µL asexual *Plasmodium falciparum* parasites if they are aged <1 year or >2500 asexual parasites/µL of blood if they are aged >1 year. Patients who are admitted with viral infections but who develop a secondary bacterial pneumonia during hospitalisation will be excluded from the VI group.

Patients with HIV infection or exposure (<1 year old) will be categorised as fungal pneumonia if they have chest X-ray with bilateral/symmetrical reticular or granular opacities without lobar consolidation and severe hypoxia (≤85% at admission) or positive NPS PCR for *Pneumocystis jiroveci* or a blood culture positive to any pathogenic fungal micro-organism.

For the diagnostic categorisation of probable bacterial pneumonia, patients will be assigned to the *probable BA* group based on review of data by an expert panel including paediatricians and infectious diseases specialists with at least 10 years of experience. Data will include clinical information collected at and during admission, chest X-ray results, full blood cell counts (FBCs), malaria microscopy and/or rapid diagnostic test (mRDT), bacteria culture, viral NPS, and other laboratory test results, when available, for example, urinalysis. Two experts will review patient data, independently and blindly, classifying patients as 'certainly' BA or non-bacterial (NBA), 'probable' BA or NBA, 'unclear' BA or NBA. Only patients with consensual certainly or probable diagnosis will be included in this analysis.

Objective 2: poor prognosis

Poor prognosis will be primarily based on a three-category outcome recorded during and after hospitalisation. Additionally, we will explore secondary outcomes associated with a poor evolution of the clinical pneumonia during the first 5 days of hospitalisation, at discharge and 30 days after admission.

For the *three-category outcome*, regardless of the original infection, patients will be selected from the following three groups: group a, children who die during or up to 30 days after admission; group b, children with prolonged hospital stay (>7 days) or who need to have antibiotic therapy changed within 48 hours of admission, or who are readmitted within 30 days after the first admission with symptoms associated with the first admission; and group c, children discharged well within 3 days of admission and without the need for a change in antibiotic therapy after

admission. Secondary outcomes of clinical deterioration will be (1) area under the first 5 days of the oxygen saturation curve; (2) necessity to change antibiotic therapy within the first 2 days from admission or to receive oxygen therapy when it was not required at admission and (3) duration of hospital admission.

Participants

We will enrol 900 patients with clinical pneumonia aged 2–59 months by January 2021, assuming we will have approximately 30 patients with a definitive diagnosis of bacterial infection, 30 with malaria and 30 with viral infections (figure 1). Among the 900 patients, we anticipate enrolling 60 and 100 patients with and without probable bacterial infection, respectively.

For the study of prognostic biomarkers (figure 2), we will include 160 patients. All deaths during or up to 30 days from admission will be selected, and the remaining sample size will be divided between the other two category outcomes, that is, group b, prolonged admission, readmission or changed antibiotics during admission, and group c, discharged within 3 days with improvement of symptoms. To meet the target sample size in the groups b and c, eligible children will be randomly selected.

A total of 20 community healthy control children aged 2–59 months without any symptoms or signs and no malaria infection will be enrolled in vaccination clinics in Basse to assess precision of signatures.

Patient and public involvement

Patients and the public were not involved in the design of this research.

Inclusion and exclusion criteria

Paediatric patients aged 2–59 months will be screened and will be enrolled if they present with cough or difficult breathing, the guardians provide informed consent, are referred for probable admission, and have at least one of the following signs: increased respiratory rate for age based on the WHO criteria, lower chest indrawing, peripheral arterial oxygen saturation <93% (measured by pulse oximetry), grunting, nasal flaring or undernutrition. The WHO criteria for increased respiratory rate are respiratory rate ≥50 breaths/min for children 2–11 months old; and respiratory rate ≥40 breaths/min for children 1–5 years old.[108] Undernutrition will be defined as visible wasting or middle upper arm circumference <11.5 cm if aged ≥6 months or weight-for-height Z-scores of <−3. Patients will be excluded if they have suspected tuberculosis based on a cough lasting >2 weeks, were admitted to a hospital in the previous 2 weeks and have evidence of any condition that could be worsened by collection of blood.

Healthy community controls will be enrolled if an mRDT is negative; vaccines have not been administered that day; there are no symptoms of any disease; and there is no history of hospital admission in the previous 4 weeks. Healthy controls will be excluded if they are later found to have malaria infection detected by a positive microscopy.

3 PROGNOSIS GROUPS (N = 160)		
Poor prognosis	**Adverse progress of disease**	**Good Prognosis**
• Dearth	• Re-admission • Need to switch antibiotic therapy ≤ 48 hrs of hospitalization • > 3 days of hospitalization	• Discharged well • No need to switch antibiotic therapy during admission • < 3 days of hospitaliza-

CONTINUOUS AND CATEGORICAL POOR PROGNOSIS OUTCOMES (N = 160)
• Area under 5-day SaO_2
• Need to switch antibiotic therapy ≤ 48 hrs of hospitalization
• Length of hospital stay

Figure 2 Exploratory primary (A) and secondary (B) prognosis outcomes.

Recruitment and follow-up

On arrival at any of the study hospitals, patients aged 2–59 months will be screened for the presence of respiratory symptoms and undernutrition by a senior research nurse with extensive experience in research in paediatric pneumonia (figure 1). If they pass this first-level assessment, they will be referred to a second research nurse for a detailed medical history and physical examination. Those patients who are considered eligible will undergo informed consent procedures.

After informed consent, patients will be referred for clinical examination and admission procedures by a research clinician and then for sample collection and chest X-ray. NPS will be performed and 5 mL of venous blood will be drawn (never more than 3 mL/kg per WHO standard[109] with 3 mL immediately placed into one paediatric blood culture bottle (Pedibact; BD, Franklin Lakes, New Jersey, USA) and the remaining 2 mL placed in an EDTA vacutainer. From the EDTA tube, a drop of blood will be used for haemoglobin rapid test and mRDT, both part of standard of care in the Gambia. Other tests based on standard of care may be conducted, including urinalyses, glucose, examination of cerebrospinal fluid (CSF) and pleural fluid, or lung aspirate. We will follow local clinical practice and HIV testing will not be routinely conducted. HIV prevalence is approximately 1% in antenatal surveys in The Gambia. Blood in EDTA will be processed and used within 1 hour from collection for FBC, malaria microscopy, and will be centrifuged and aliquoted to store plasma at −80C in the MRCG biobank for the multiplex bead-based immunoassay. An aliquot of 150 µL will be sent to the Centre of Translational Immunology in Utrecht for quantification of proteins with temperature controlled dry ice shipment.

Follow-up

After admission, patients will be followed up daily for 5 days to record pulse oximetry, antibiotic therapy, feeding, clinical deterioration with need for additional tests. If patients are receiving oxygen during the follow-up visits, they will have their lowest recorded pulse oximetry imputed. At hospital separation, a research nurse will fill a case report form (CRF) to record diagnosis and outcome (eg, death, absconding or discharge alive and well). Moreover, for patients who are discharged, a phone call will be placed to the parent/guardian 30 days after admission.

Study measurements and procedures

► Bacterial culture: approximately 3 mL will be incubated in an automatic BACTEC 9050 system (BD) for a minimum of 4 days. Positive blood cultures will be examined by Gram stain and subcultured onto blood agar, chocolate agar or MacConkey agar plates based on Gram stain. Bacterial growth will be defined as contaminant when bacteria generally considered skin flora are isolated from the positive blood culture (eg, coagulase-negative Staphylococci, *Bacillus* species or Micrococci).[110] Lung fluids will be cultured according to procedures previously described.[111]

► Culture of CSF, lung aspirate and pleural fluid aspirate: cell count and direct gram reaction will be performed as well as bacterial antigen latex agglutination test. A 50–100 µL aliquot of the specimen will be inoculated onto blood agar, chocolate agar and MacConkey Agar plates for isolation of bacterial growth.

► Viral PCR of NPS: NPS will be placed in media that will be aliquoted and stored at −80°C for real time polymerase chain reaction (rtPCR). Assays will use primers and probes for influenza virus types A and B, respiratory syncytial virus types A and B, adenovirus, parainfluenza viruses 1, 2, 3 and 4, human metapneumovirus, coronavirus, bocavirus, swine influenza virus and rhinovirus. *P. jiroveci* will also be screened in this PCR. Viral RNA and DNA will be extracted from supernatant of NPS samples using the QIAamp cador Pathogen Mini Kit or the Qiacube HT manufactured by Qiagen and PCR reactions will be ran as described elsewhere.[112] PCR for SARS-CoV-2 is not planned because paediatric pneumonia caused by SARS-CoV-2 is generally rare and clinical cases of COVID-19 in adults have been rare in the region. However, if future evidence suggests that SARS-CoV-2 shall be considered as a pathogen of paediatric clinical pneumonia, SARS-CoV-2 PCR may be performed.

► Malarial microscopy: thick blood smear slides will be read blindly by two microscopists and, if discrepant, by a third.

► FBC: FBC will be performed on a Medonic M51 five-part haematology analyser (Boule Diagnostics AB, Domnarvsgatan, Spanga, Sweden). The parameters measured will be white blood cell count, absolute and per cent count of lymphocytes, monocytes, neutrophils, eosinophils, basophils, red blood cell count, platelet count, haemoglobin, mean cell volume, haematocrit, red cell distribution width, mean cell haemoglobin, mean corpuscular haemoglobin concentration and mean platelet volume.

► Multiplex Bead-Based Immunoassay: using Luminex technology will quantify proteins in plasma and will be conducted at the Multiplex Core Facility of the Centre of Translational Immunology ISO9001:2015 laboratory using FlexMAP3D systems.[113] A total of 50 proteins will be included in the assay: (1) proteins identified in our previous study to constitute accurate signatures (eg, HP, interleukin (IL)-10 and IL-18 and tissue inhibitor of metalloproteinases 1, tumour necrosis factor receptor 2 and resistin)[114]; (2) proteins that discriminated the three groups in high-throughput scans conducted previously (eg, bactericidal/permeability-increasing protein, lactotransferrin, granulocyte colony-stimulating factor and matrix metallopeptidase 9); (3) proteins specified in previous studies and listed previously (eg, NGAL, PCT, TRAIL and IP-10), and proteins identified in previous studies listed earlier with prognostic value (eg, angiopoietin).

► Chest radiographs: frontal radiographs will be obtained, and images will be recorded in a computer file. Radiographical findings will be classified as proposed by Cherian *et al*[51] into the presence of consolidation, pleural fluid, other infiltrates or normal. Additionally, the presence of reticular/reticulonodular patterns, cavities, pneumatoceles, lymphadenopathy, abscess, focal fibrosis, pneumothorax, cardiomegaly and foreign body will be noted. Chest X-rays will be read by two clinicians and, if discrepant for consolidation, they will be adjudicated by a third reader, all with training in radiographical interpretation for pneumonia and calibration against WHO standard images. Also, a pulmonologist will read a sample of 10% of X-rays for quality assurance and control.

Data management

All data will be captured through laptops/desktop computers using an electronic medical record system (EMRS) developed at the MRCG. The only exception will be for follow-up of admitted patients who will be captured on paper forms and later entered in the EMRS. During power outages, paper forms will be completed and, when electricity is back, data will be entered into the EMRS. To ensure data quality, data validation rules have been established in the EMRS that generates alert messages. Additionally, a complete process flow has been established to sequentially activate forms for entry, and a report listing missing variables is generated during entry of forms to assure form completeness.

Data analysis

Descriptive analysis will be conducted to seek outliers and to assess assumptions for tests and appropriate transformations. Patient severity scores at baseline will be derived based on principal component analysis, including baseline clinical and laboratory variables. These scores will be included in the analysis and used to explore stratification of patients in the study of diagnostic biomarker signatures. Adjustments for potential confounders (eg, overall and cell-specific leucocyte counts, breast feeding and undernutrition) and assessment of heterogeneity across subgroups will be done by incorporating the factor in analysis.

In analysis of objective 1, we will first test previous diagnostic biomarker signatures[5 105] through estimating AUC-ROC, sensitivity and specificity with 95% CI. Using levels of biomarkers included in Valim *et al*[105] and Oved *et al*,[98] we will reproduce signatures using the original models. To discover alternative biomarker signatures, we will start by conducting single biomarker analysis in which we will rank individual markers by pairwise comparisons of the three definitive diagnosis groups (primary analysis) and the two bacterial versus non-bacterial probable diagnosis groups (secondary analysis) and estimation of AUC-ROC with 95% CI.[115] We will also estimate sensitivity and specificity for an optimal cut-off (weighting towards high sensitivity to bacteria) based on comparison of the BA and NBA

definitive diagnosis groups through resampling. Next, to identify composite biomarkers, all markers (including clinical and laboratory tests and severity scores) will be combined and selected through comparison of the three definitive diagnosis groups using logistic regression with elastic net; classification or regression trees depending on the specific outcome; and support vector machine (SVM). In the SVM, non-linear radial basis function kernels will be used and features will be selected through a penalisation and recursive feature elimination.[116 117] We will train models and identify signatures based on the three diagnosis groups, although the focus is in classification of BA and NBA (defined as MA or VI diagnosis) because using three groups, we found that signatures are likely to be more accurate when markers are not monotonically increasing or decreasing across groups. For instance, if a marker is high in BA, low in VI and high in MA, levels of VI with MA are suboptimally averaged when VI and MA are collapsed for training signatures. Within each class of models, tuning parameters will be obtained through a metric of misclassification (eg, deviance or error) estimated in fivefold cross validation. Additionally, models will be selected based on a biased choice of tuning parameters (those determining ≤5 markers). Penalties for missing a bacterial infection will be used. Models will be tested by estimating accuracy to diagnose BA versus non-BA in the definitive diagnosis classification using resampling. Moreover, accuracy will be estimated by comparing BA with NBA probable diagnosis groups.

Analyses of objective 2 will use categorical or continuous surrogates of poor prognosis as outcomes and markers as predictors. Markers will be measured at admission. Signatures will be sought and tested using the same models of objective 1.

To estimate precision, we will report the coefficient of variation (CV) of selected biomarkers in healthy controls. Contaminated cultures will be considered negative. In immunoassays, samples <lowest limit of quantification (LLOQ) or > highest limit of quantification (HLOQ) will be assigned a value below and above the limits of quantification, respectively, through a uniform distribution.

An α-significance level of 0.05 will be used. In single-marker analysis, p values and CIs will be adjusted for family-wise error rate through permutation or false discovery rate through Benjamin-Rocheberg, depending on the analysis.[118]

Sample size justification
We plan to enrol 900 patients to obtain 30 patients with confirmed bacterial disease based on expectations that ~3% of patients will have a positive bacterial culture for objective 1. Also, we anticipate that 30 patients will meet definitive viral or malarial aetiology. If necessary, we will randomly sample patients from the viral and malarial groups. The choice of 30 patients with and 60 patients without bacterial disease was based on feasibility considerations, as was the choice of 20 healthy controls and 160 patients in the analysis of the 'probable' diagnosis

Table 1 Precision (exact 95% CIs) of sensitivity (sample size of 30) and specificity (sample size of 60) when evaluating diagnostic biomarker signatures in analysis of aim 1

| Proportion (%) | 95% CI | |
	Sensitivity (%)	Specificity (%)
85	65 to 94	73 to 95
90	73 to 98	79 to 90
95	78 to 99	86 to 99

categories (a random stratified sample from all available subjects). However, with ~90 patients and only 23 patients with bacterial infections, we could previously identify accurate signatures.[105] We anticipate that in objective 1, accuracy of previously identified signatures will be estimated with relatively narrow CIs (table 1). For healthy controls, studying 20 patients will yield a 0.18 CI range for a CV of 0.20 with 0.80 assurance.[119] Analysis of objective 2 will include a larger number of patients than objective 1 and shall yield precise CIs and more stable cross validation-based estimates.

ETHICS AND DISSEMINATION
Ethics approval has been obtained from the Gambia Government/MRCG Joint Ethics Committee (protocol 1616), and the institutional review board of Boston University Medical Centre (STUDY00000958). We plan to share the findings in peer-reviewed journals.

PROTOCOL REGISTRATION
The study is registered in clinicaltrials.gov.

Author affiliations
[1]Department of Global Health, Boston University School of Public Health, Boston, Massachusetts, USA
[2]Medical Research Council Unit, The Gambia at the London School of Hygiene & Tropical Medicine, Fajara, The Gambia
[3]Center of Translational Immunology, Department of Rheumatology and Clinical Immunology, University Medical Center Utrecht, Utrecht, The Netherlands
[4]Bansang Hospital, Bansang, The Gambia
[5]Basse Hospital, Basse, The Gambia
[6]Hospital Clínic, Universitat de Barcelona, ISGlobal, Barcelona, Spain
[7]Centro de Investigação em Saúde de Manhiça (CISM), Maputo, Mozambique
[8]Luminex Corp, Austin, Texas, USA
[9]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA
[10]Division of Pulmonary and Critical Care Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA
[11]Independent Researcher, Boston, Massachusetts, USA
[12]Division of Infectious Diseases, University of Louisville, Louisville, Kentucky, USA
[13]Vascular Biology Program, Children's Hospital Boston, Boston, Massachusetts, USA
[14]Department of Surgery, Harvard Medical School, Boston, Massachusetts, USA
[15]Disease Elimination and Control, Medical Research Council Unit, Fajara, Gambia
[16]London School of Hygiene & Tropical Medicine, London, UK
[17]Boston University School of Public Health, Boston, Massachusetts, USA
[18]Department of Disease Control, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, UK

ORCID iDs
Clarissa Valim http://orcid.org/0000-0001-9621-203X
Grant A Mackenzie http://orcid.org/0000-0002-4994-2627

REFERENCES

1. Nair H, Simões EA, Rudan I, *et al*. Global and regional burden of hospital admissions for severe acute lower respiratory infections in young children in 2010: a systematic analysis. *Lancet* 2013;381:1380–90.
2. Taylor WRJ, Hanson J, Turner GDH, *et al*. Respiratory manifestations of malaria. *Chest* 2012;142:492–505.
3. Tsao T-C, Tsai H-C, Chang S-C. Clinical usefulness of urinary fatty acid binding proteins in assessing the severity and predicting treatment response of pneumonia in critically ill patients: a cross-sectional study. *Medicine* 2016;95:e3682.
4. Engelmann I, Dubos F, Lobert P-E, *et al*. Diagnosis of viral infections using myxovirus resistance protein A (MxA). *Pediatrics* 2015;135:e985–93.
5. 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.
6. Spoorenberg SMC, Vestjens SMT, Rijkers GT, *et al*. YKL-40, CCL18 and SP-D predict mortality in patients hospitalized with community-acquired pneumonia. *Respirology* 2017;22:542–50.
7. Eden E, Srugo I, Gottlieb T, *et al*. Diagnostic accuracy of a TRAIL, IP-10 and CRP combination for discriminating bacterial and viral etiologies at the emergency department. *J Infect* 2016;73:177–80.
8. Relster MM, Holm A, Pedersen C. Plasma cytokines eotaxin, MIP-1α, MCP-4, and vascular endothelial growth factor in acute lower respiratory tract infection. *APMIS* 2017;125:148–56.
9. Higdon MM, Le T, O'Brien KL, *et al*. Association of C-Reactive Protein With Bacterial and Respiratory Syncytial Virus-Associated Pneumonia Among Children Aged <5 Years in the PERCH Study. *Clin Infect Dis* 2017;64:S378–86.
10. Yan ST, Sun LC, Jia HB, *et al*. Procalcitonin levels in bloodstream infections caused by different sources and species of bacteria. *Am J Emerg Med* 2017;35:579–83.
11. McCluskey SM, Schuetz P, Abers MS, *et al*. Serial procalcitonin as a predictor of bacteremia and need for intensive care unit care in adults with pneumonia, including those with highest severity: a prospective cohort study. *Open Forum Infect Dis* 2017;4:ofw238.
12. Mahende C, Ngasala B, Lusingu J, *et al*. Profile of C-reactive protein, white cells and neutrophil populations in febrile children from rural north-eastern Tanzania. *Pan Afr Med J* 2017;26:51.
13. Li YL, Zhai LC, Ji JH, *et al*. Detection of combined procalcitonin and C-reactive protein applied in the diagnosis of bacterial infections. *J Biol Regul Homeost Agents* 2017;31:177–81.
14. Pilotto A, Dini S, Daragjati J, *et al*. Combined use of the multidimensional prognostic index (Mpi) and procalcitonin serum levels in predicting 1-month mortality risk in older patients hospitalized with community-acquired pneumonia (CAP): a prospective study. *Aging Clin Exp Res* 2018;30:193-197.
15. Andersen SB, Baunbæk Egelund G, Jensen AV, *et al*. Failure of CRP decline within three days of hospitalization is associated with poor prognosis of community-acquired pneumonia. *Infect Dis* 2017;49:251–60.
16. Chang P-Y, Tsao S-M, Chang J-H, *et al*. Plasma levels of soluble intercellular adhesion molecule-1 as a biomarker for disease severity of patients with community-acquired pneumonia. *Clin Chim Acta* 2016;463:174–80.
17. Banoei MM, Vogel HJ, Weljie AM, *et al*. Plasma metabolomics for the diagnosis and prognosis of H1N1 influenza pneumonia. *Crit Care* 2017;21:97.
18. Frenzen FS, Kutschan U, Meiswinkel N, *et al*. Admission lactate predicts poor prognosis independently of the CRB/CURB-65 scores in community-acquired pneumonia. *Clin Microbiol Infect* 2018;24:306.e1–e6.
19. Chkhaidze I, Kapanadze N. Cytokines as the predictors of severe Mycoplasma pneumoniae pneumonia in children (review). *Georgian Med News* 2017;267:89–95.
20. Zar HJ, Andronikou S, Nicol MP. Advances in the diagnosis of pneumonia in children. *BMJ* 2017;358:j2739.
21. Zhou J-M, Ye Q. Utility of assessing cytokine levels for the differential diagnosis of pneumonia in a pediatric population. *Pediatr Crit Care Med* 2017;18:e162–6.
22. Curbelo J, Luquero Bueno S, Galván-Román JM, *et al*. Inflammation biomarkers in blood as mortality predictors in community-acquired pneumonia admitted patients: importance of comparison with neutrophil count percentage or neutrophil-lymphocyte ratio. *PLoS One* 2017;12:e0173947.
23. Salluh JIF, Souza-Dantas VC, Póvoa P. The current status of biomarkers for the diagnosis of nosocomial pneumonias. *Curr Opin Crit Care* 2017;23:391–7.
24. Bolatkale M, Duger M, Ülfer G, *et al*. A novel biochemical marker for community-acquired pneumonia: Ischemia-modified albumin. *Am J Emerg Med* 2017;35:1121–5.
25. Esposito S, Di Gangi M, Cardinale F, *et al*. Sensitivity and specificity of soluble triggering receptor expressed on myeloid cells-1, midregional proatrial natriuretic peptide and midregional proadrenomedullin for distinguishing etiology and to assess severity in community-acquired pneumonia. *PLoS One* 2016;11:e0163262.
26. Li J, Sun L, Xu F, *et al*. Screening and identification of APOC1 as a novel potential biomarker for differentiate of *Mycoplasma pneumoniae* in children. *Front Microbiol* 2016;7:7.
27. Saghafian-Hedengren S, Mathew JL, Hagel E, *et al*. Assessment of cytokine and chemokine signatures as potential biomarkers of childhood community-acquired pneumonia severity: a nested cohort study in India. *Pediatr Infect Dis J* 2017;36:102–8.
28. Vögeli A, Ottiger M, Meier MA, *et al*. Admission levels of asymmetric and symmetric dimethylarginine predict long-term outcome in patients with community-acquired pneumonia. *Respir Res* 2017;18:25.
29. Vestjens SMT, Spoorenberg SMC, Rijkers GT, *et al*. High-sensitivity cardiac troponin T predicts mortality after hospitalization for community-acquired pneumonia. *Respirology* 2017;22:1000–6.
30. Esposito S, Bianchini S, Gambino M, *et al*. Measurement of lipocalin-2 and syndecan-4 levels to differentiate bacterial from viral infection in children with community-acquired pneumonia. *BMC Pulm Med* 2016;16:103.
31. Loonen AJM, Kesarsing C, Kusters R, *et al*. High pneumococcal DNA load, procalcitonin and suPAR levels correlate to severe

disease development in patients with pneumococcal pneumonia. *Eur J Clin Microbiol Infect Dis* 2017;36:1541–7.

32 Meier MA, Ottiger M, Vögeli A, *et al*. Activation of the tryptophan/serotonin pathway is associated with severity and predicts outcomes in pneumonia: results of a long-term cohort study. *Clin Chem Lab Med* 2017;55:1060–9.

33 Zhang Y, Zhou Y, Li S, *et al*. The clinical characteristics and predictors of refractory Mycoplasma pneumoniae pneumonia in children. *PLoS One* 2016;11:e0156465.

34 Levy C, Biscardi S, Dommergues MA, *et al*. Impact of PCV13 on community-acquired pneumonia by C-reactive protein and procalcitonin levels in children. *Vaccine* 2017;35:5058–64.

35 Kim MW, Lim JY, Oh SH. Mortality prediction using serum biomarkers and various clinical risk scales in community-acquired pneumonia. *Scand J Clin Lab Invest* 2017;77:486–92.

36 Stockmann C, Ampofo K, Killpack J. Procalcitonin accurately identifies hospitalized children with low risk of bacterial community-acquired pneumonia. *J Pediat Infect Dis Soc* 2017.

37 Irwin AD, Grant A, Williams R, *et al*. Predicting risk of serious bacterial infections in febrile children in the emergency department. *Pediatrics* 2017;140 doi:10.1542/peds.2016-2853

38 Michelow IC, Olsen K, Lozano J, *et al*. Epidemiology and clinical characteristics of community-acquired pneumonia in hospitalized children. *Pediatrics* 2004;113:701–7.

39 English M, Esamai F, Wasunna A, *et al*. Assessment of inpatient paediatric care in first referral level hospitals in 13 districts in Kenya. *Lancet* 2004;363:1948–53.

40 Reyburn H, Mwakasungula E, Chonya S, *et al*. Clinical assessment and treatment in paediatric wards in the north-east of the United Republic of Tanzania. *Bull World Health Organ* 2008;86:132–9.

41 Mackenzie G. The definition and classification of pneumonia. *Pneumonia* 2016;8:14.

42 Donnelly JP, Baddley JW, Wang HE. Antibiotic utilization for acute respiratory tract infections in U.S. emergency departments. *Antimicrob Agents Chemother* 2014;58:1451–7.

43 Johansson EW, Selling KE, Nsona H, *et al*. Integrated paediatric fever management and antibiotic over-treatment in Malawi health facilities: data mining a national facility census. *Malar J* 2016;15:396.

44 Ashu EE, Jarju S, Dione M, *et al*. Population structure, epidemiology and antibiotic resistance patterns of Streptococcus pneumoniae serotype 5: prior to PCV-13 vaccine introduction in eastern Gambia. *BMC Infect Dis* 2016;16:33.

45 Makoka MH, Miller WC, Hoffman IF, *et al*. Bacterial infections in Lilongwe, Malawi: aetiology and antibiotic resistance. *BMC Infect Dis* 2012;12:67.

46 Feasey NA, Masesa C, Jassi C, *et al*. Three epidemics of invasive multidrug-resistant Salmonella bloodstream infection in Blantyre, Malawi, 1998-2014. *Clin Infect Dis* 2015;61:S363–71.

47 Gwer S, Newton CRJC, Berkley JA. Over-diagnosis and co-morbidity of severe malaria in African children: a guide for clinicians. *Am J Trop Med Hyg* 2007;77:6–13.

48 Bassat Q, Machevo S, O'Callaghan-Gordo C, *et al*. Distinguishing malaria from severe pneumonia among hospitalized children who fulfilled integrated management of childhood illness criteria for both diseases: a hospital-based study in Mozambique. *Am J Trop Med Hyg* 2011;85:626–34.

49 Ramirez M, Melo-Cristino J. Expanding the diagnosis of pediatric bacteremic pneumococcal pneumonia from blood cultures to molecular methods: advantages and caveats. *Clin Infect Dis* 2010;51:1050–2.

50 Bradley JS, Byington CL, Shah SS, *et al*. The management of community-acquired pneumonia in infants and children older than 3 months of age: clinical practice guidelines by the pediatric infectious diseases Society and the infectious diseases Society of America. *Clin Infect Dis* 2011;53:e25–76.

51 Cherian T, Mulholland EK, Carlin JB, *et al*. Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies. *Bull World Health Organ* 2005;83:353–9.

52 Edwards M, Lawson Z, Morris S, *et al*. The presence of radiological features on chest radiographs: how well do clinicians agree? *Clin Radiol* 2012;67:664–8.

53 Davies HD, Wang EE, Manson D, *et al*. Reliability of the chest radiograph in the diagnosis of lower respiratory infections in young children. *Pediatr Infect Dis J* 1996;15:600–4.

54 Albaum MN, Hill LC, Murphy M, *et al*. Interobserver reliability of the chest radiograph in community-acquired pneumonia. Port Investigators. *Chest* 1996;110:343–50.

55 Hagaman JT, Rouan GW, Shipley RT, *et al*. Admission chest radiograph lacks sensitivity in the diagnosis of community-acquired pneumonia. *Am J Med Sci* 2009;337:236–40.

56 Johnson J, Kline JA. Intraobserver and interobserver agreement of the interpretation of pediatric chest radiographs. *Emerg Radiol* 2010;17:285–90.

57 Loeb MB, Carusone SBC, Marrie TJ, *et al*. Interobserver reliability of radiologists' interpretations of mobile chest radiographs for nursing home-acquired pneumonia. *J Am Med Dir Assoc* 2006;7:416–9.

58 O'Grady K-AF, Torzillo PJ, Ruben AR, *et al*. Identification of radiological alveolar pneumonia in children with high rates of hospitalized respiratory infections: comparison of WHO-defined and pediatric pulmonologist diagnosis in the clinical context. *Pediatr Pulmonol* 2012;47:386–92.

59 Murdoch DR, O'Brien KL, Scott JAG, *et al*. Breathing new life into pneumonia diagnostics. *J Clin Microbiol* 2009;47:3405–8.

60 Gonsalves WI, Cornish N, Moore M, *et al*. Effects of volume and site of blood draw on blood culture results. *J Clin Microbiol* 2009;47:3482–5.

61 Avni T, Mansur N, Leibovici L, *et al*. PCR using blood for diagnosis of invasive pneumococcal disease: systematic review and meta-analysis. *J Clin Microbiol* 2010;48:489–96.

62 Resti M, Moriondo M, Cortimiglia M, *et al*. Community-acquired bacteremic pneumococcal pneumonia in children: diagnosis and serotyping by real-time polymerase chain reaction using blood samples. *Clin Infect Dis* 2010;51:1042–9.

63 Morpeth SC, Deloria Knoll M, Scott JAG, *et al*. Detection of pneumococcal DNA in blood by polymerase chain reaction for diagnosing pneumococcal pneumonia in young children from low- and middle-income countries. *Clin Infect Dis* 2017;64:S347–56.

64 Anjay MA, Anoop P. Diagnostic utility of rapid immunochromatographic urine antigen testing in suspected pneumococcal infections. *Arch Dis Child* 2008;93:628–31.

65 Charkaluk M-L, Kalach N, Mvogo H, *et al*. Assessment of a rapid urinary antigen detection by an immunochromatographic test for diagnosis of pneumococcal infection in children. *Diagn Microbiol Infect Dis* 2006;55:89–94.

66 Thea DM, Seidenberg P, Park DE, *et al*. Limited utility of polymerase chain reaction in induced sputum specimens for determining the causes of childhood pneumonia in resource-poor settings: findings from the pneumonia etiology research for child health (PERCH) study. *Clin Infect Dis* 2017;64:S289–300.

67 Murdoch DR, Morpeth SC, Hammitt LL, *et al*. The diagnostic utility of induced sputum microscopy and culture in childhood pneumonia. *Clin Infect Dis* 2017;64:S280–8.

68 Koram KA, Molyneux ME. When is 'malaria' malaria? The different burdens of malaria infection, malaria disease, and malaria-like illnesses. *Am J Trop Med Hyg* 2007;77:1–5.

69 D'Acremont V, Kilowoko M, Kyungu E, *et al*. Beyond malaria-causes of fever in outpatient Tanzanian children. *N Engl J Med* 2014;370:809–17.

70 Maniaci V, Dauber A, Weiss S, *et al*. Procalcitonin in young febrile infants for the detection of serious bacterial infections. *Pediatrics* 2008;122:701–10.

71 Nascimento-Carvalho CM, Cardoso M-RA, Barral A, *et al*. Procalcitonin is useful in identifying bacteraemia among children with pneumonia. *Scand J Infect Dis* 2010;42:644–9.

72 Zhydkov A, Christ-Crain M, Thomann R, *et al*. Utility of procalcitonin, C-reactive protein and white blood cells alone and in combination for the prediction of clinical outcomes in community-acquired pneumonia. *Clin Chem Lab Med* 2015;53:559–66.

73 Tang BMP, Eslick GD, Craig JC, *et al*. Accuracy of procalcitonin for sepsis diagnosis in critically ill patients: systematic review and meta-analysis. *Lancet Infect Dis* 2007;7:210–7.

74 Thayyil S, Shenoy M, Hamaluba M, *et al*. Is procalcitonin useful in early diagnosis of serious bacterial infections in children? *Acta Paediatr* 2005;94:155–8.

75 Krüger S, Ewig S, Marre R, *et al*. Procalcitonin predicts patients at low risk of death from community-acquired pneumonia across all CRB-65 classes. *Eur Respir J* 2008;31:349–55.

76 Hirakata Y, Yanagihara K, Kurihara S, *et al*. Comparison of usefulness of plasma procalcitonin and C-reactive protein measurements for estimation of severity in adults with community-acquired pneumonia. *Diagn Microbiol Infect Dis* 2008;61:170–4.

77 Meili M, Kutz A, Briel M, *et al*. Infection biomarkers in primary care patients with acute respiratory tract infections-comparison of procalcitonin and C-reactive protein. *BMC Pulm Med* 2016;16:43.

78 Lindstrom ST, Wong EKC. Procalcitonin, a valuable biomarker assisting clinical decision-making in the management of community-acquired pneumonia. *Intern Med J* 2014;44:390–7.

79 Heiskanen-Kosma T, Korppi M. Serum C-reactive protein cannot differentiate bacterial and viral aetiology of community-acquired pneumonia in children in primary healthcare settings. *Scand J Infect Dis* 2000;32:399–402.

80 Korppi M, Remes S, Heiskanen-Kosma T. Serum procalcitonin concentrations in bacterial pneumonia in children: a negative result in primary healthcare settings. *Pediatr Pulmonol* 2003;35:56–61.

81 Rodríguez AH, Avilés-Jurado FX, Díaz E, *et al*. Procalcitonin (PCT) levels for ruling-out bacterial coinfection in ICU patients with influenza: a CHAID decision-tree analysis. *J Infect* 2016;72:143–51.

82 Haran JP, Beaudoin FL, Suner S, *et al*. C-Reactive protein as predictor of bacterial infection among patients with an influenza-like illness. *Am J Emerg Med* 2013;31:137–44.

83 Johansson N, Kalin M, Backman-Johansson C, *et al*. Procalcitonin levels in community-acquired pneumonia - correlation with aetiology and severity. *Scand J Infect Dis* 2014;46:787–91.

84 Schuetz P, Batschwaroff M, Dusemund F, *et al*. Effectiveness of a procalcitonin algorithm to guide antibiotic therapy in respiratory tract infections outside of study conditions: a post-study survey. *Eur J Clin Microbiol Infect Dis* 2010;29:269–77.

85 Long W, Deng X, Zhang Y, *et al*. Procalcitonin guidance for reduction of antibiotic use in low-risk outpatients with community-acquired pneumonia. *Respirology* 2011;16:819–24.

86 Christ-Crain M, Stolz D, Bingisser R, *et al*. Procalcitonin guidance of antibiotic therapy in community-acquired pneumonia: a randomized trial. *Am J Respir Crit Care Med* 2006;174:84–93.

87 Huang DT, Weissfeld LA, Kellum JA, *et al*. Risk prediction with procalcitonin and clinical rules in community-acquired pneumonia. *Ann Emerg Med* 2008;52:48–58.

88 Don M, Valent F, Korppi M, *et al*. Efficacy of serum procalcitonin in evaluating severity of community-acquired pneumonia in childhood. *Scand J Infect Dis* 2007;39:129–37.

89 Michelow IC, Katz K, McCracken GH, *et al*. Systemic cytokine profile in children with community-acquired pneumonia. *Pediatr Pulmonol* 2007;42:640–5.

90 Florin TA, Ambroggio L. Biomarkers for community-acquired pneumonia in the emergency department. *Curr Infect Dis Rep* 2014;16:451.

91 Müller B, Harbarth S, Stolz D, *et al*. Diagnostic and prognostic accuracy of clinical and laboratory parameters in community-acquired pneumonia. *BMC Infect Dis* 2007;7:10.

92 Bafadhel M, Clark TW, Reid C, *et al*. Procalcitonin and C-reactive protein in hospitalized adult patients with community-acquired pneumonia or exacerbation of asthma or COPD. *Chest* 2011;139:1410–8.

93 Principi N, Esposito S. Biomarkers in pediatric community-acquired pneumonia. *Int J Mol Sci* 2017;18.

94 Bordon J, Aliberti S, Fernandez-Botran R, *et al*. Understanding the roles of cytokines and neutrophil activity and neutrophil apoptosis in the protective versus deleterious inflammatory response in pneumonia. *Int J Infect Dis* 2013;17:e76–83.

95 Mahajan P, Kuppermann N, Mejias A, *et al*. Association of RNA Biosignatures with bacterial infections in febrile infants aged 60 days or younger. *JAMA* 2016;316:846–57.

96 Herberg JA, Kaforou M, Wright VJ, *et al*. Diagnostic test accuracy of a 2-Transcript host RNA signature for discriminating bacterial vs viral infection in febrile children. *JAMA* 2016;316:835–45.

97 Huang H, Ideh RC, Gitau E, *et al*. Discovery and validation of biomarkers to guide clinical management of pneumonia in African children. *Clin Infect Dis* 2014;58:1707-15.

98 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.

99 Srugo I, Klein A, Stein M, *et al*. Validation of a novel assay to distinguish bacterial and viral infections. *Pediatrics* 2017;140. doi:10.1542/peds.2016-3453. [Epub ahead of print: 13 09 2017].

100 Stein M, Lipman-Arens S, Oved K, *et al*. A novel host-protein assay outperforms routine parameters for distinguishing between bacterial and viral lower respiratory tract infections. *Diagn Microbiol Infect Dis* 2018;90:206–13.

101 Ashkenazi-Hoffnung L, Oved K, Navon R, *et al*. A host-protein signature is superior to other biomarkers for differentiating between bacterial and viral disease in patients with respiratory infection and fever without source: a prospective observational study. *Eur J Clin Microbiol Infect Dis* 2018;37:1361–71.

102 Self WH, Rosen J, Sharp SC, *et al*. Diagnostic accuracy of FebriDx: a rapid test to detect immune responses to viral and bacterial upper respiratory infections. *J Clin Med* 2017;6:94. doi:10.3390/jcm6100094

103 Shapiro NI, Self WH, Rosen J, *et al*. A prospective, multi-centre US clinical trial to determine accuracy of FebriDx point-of-care testing for acute upper respiratory infections with and without a confirmed fever. *Ann Med* 2018;50:420–9.

104 Dittrich S, Tadesse BT, Moussy F, *et al*. Target product profile for a diagnostic assay to differentiate between bacterial and non-bacterial infections and reduce antimicrobial overuse in resource-limited settings: an expert consensus. *PLoS One* 2016;11:e0161721.

105 Valim C, Ahmad R, Lanaspa M, *et al*. Responses to bacteria, virus, and malaria distinguish the etiology of pediatric clinical pneumonia. *Am J Respir Crit Care Med* 2016;193:448–59.

106 Irwin AD, Marriage F, Mankhambo LA, *et al*. Novel biomarker combination improves the diagnosis of serious bacterial infections in Malawian children. *BMC Med Genomics* 2012;5:13.

107 Berlutti F, Pantanella F, Natalizi T, *et al*. Antiviral properties of lactoferrin-a natural immunity molecule. *Molecules* 2011;16:6992–7018.

108 World Health Organization. *Integrated management of childhood illness*. Geneva: World Health Organization, 2014.

109 Howie SRC. Blood sample volumes in child health research: review of safe limits. *Bull World Health Organ* 2011;89:46–53.

110 Bates DW, Lee TH. Rapid classification of positive blood cultures. Prospective validation of a multivariate algorithm. *JAMA* 1992;267:1962–6.

111 Adegbola RA, Falade AG, Sam BE, *et al*. The etiology of pneumonia in malnourished and well-nourished Gambian children. *Pediatr Infect Dis J* 1994;13:975–82.

112 Gunson RN, Collins TC, Carman WF. Real-time RT-PCR detection of 12 respiratory viral infections in four triplex reactions. *J Clin Virol* 2005;33:341–4.

113 Scholman RC, Giovannone B, Hiddingh S, *et al*. Effect of anticoagulants on 162 circulating immune related proteins in healthy subjects. *Cytokine* 2018;106:114–24.

114 Valim C, Ahmad R, Lanaspa M. Responses to bacteria, virus and malaria distinguish the etiology of pediatric clinical pneumonia. *Am J Respir Crit Care Med* 2015.

115 DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837–45.

116 Guyon I, Weston J, Barnhill S, *et al*. Gene selection for cancer classification using support vector machines. *Mach Learn* 2002;46:389–422.

117 Becker N, Toedt G, Lichter P, *et al*. Elastic SCAD as a novel penalization method for SVM classification tasks in high-dimensional data. *BMC Bioinformatics* 2011;12:138.

118 Goeman JJ, Solari A. Multiple hypothesis testing in genomics. *Stat Med* 2014;33:1946–78.

119 Kelley K. Sample size planning for the coefficient of variation from the accuracy in parameter estimation approach. *Behav Res Methods* 2007;39:755–66.