

Host biomarkers to predict the severity of acute febrile illness: A scoping review

Matthew L. Robinson,¹ Meklit Workneh,¹ Sabine Dittrich,^{2,3} Spruha Kurlekar,⁴ Rebecca Yee,⁵ Maya C. Nirmalraj,⁵ Karen A. Robinson,⁶ Yukari C. Manabe¹

1. Johns Hopkins School of Medicine, Division of Infectious Diseases, Baltimore, MD, USA
2. FIND, Geneva, Switzerland
3. University of Oxford, Nuffield School of Medicine, Faculty of Tropical Medicine and Global Health, Oxford, UK
4. Seth G. S. Medical College and KEM Hospital, Mumbai, India
5. Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA
6. Johns Hopkins School of Medicine, Division of General Internal Medicine, Baltimore, MD, USA

Keywords: biomarker, acute febrile illness, fever, sepsis, severity

Running title: Biomarkers for febrile illness severity

Corresponding Author

Matthew L Robinson

mrobin85@jhmi.edu

600 N. Wolfe St., Phipps Building, Room 528

Baltimore, MD 21287

Phone: 410-287-2443

Alternate Corresponding Author

Yukari C. Manabe

ymanabe@jhmi.edu

1830 East Monument Street, Rm 443

Baltimore, MD, 21205

Phone: 410-955-8571

Summary: A systematic review and limited metanalysis of 89 host biomarkers revealed that most individual biomarkers offer modest performance in predicting the severity of acute febrile illness; several however have performance characteristics which have shown promise in multiple studies.

Abstract

Background: Acute febrile illness (AFI) ranges from mild to mortal, yet severity is difficult to assess.

Host biomarkers may identify patients with AFI who require a higher level of care; choosing appropriate biomarkers for this role among an expanding pool of candidates is challenging. We performed a scoping review to evaluate the performance of novel host biomarkers to predict AFI severity.

Methods: PubMed was systematically searched for manuscripts published January 1, 2013 to February 10, 2018 for studies reporting the association of host biomarker levels and a measure of disease severity among patients with a suspected or diagnosed cause of AFI. Identified abstracts and full text manuscripts were reviewed for eligibility by 2 reviewers. Biomarker performance was evaluated primarily by the area under the curve (AUC) of the receiver operator characteristic to distinguish severe disease. We aggregated data describing biomarker performance by AUC using weighted mean, fixed effects meta-analyses, and random effects meta-analyses.

Results: Among 2,303 manuscripts identified, 281 manuscripts met criteria for analysis. Data was extracted for 278 biomarkers evaluated in 45,737 participants. Among 89 biomarkers evaluated by ≥ 2 studies, there were 6 biomarkers (proadrenomedullin, copeptin, pro-atrial natriuretic peptide, serum triggering receptor expressed on myeloid cells-1, chitinase-3-like protein-1, and the pediatric sepsis biomarker risk model), that showed a weighted mean AUC >0.75 (range 0.75-0.84) in >500 patients over >2 studies.

Conclusions: Although several biomarkers show promise in predicting AFI severity across multiple studies, their test characteristics do not suggest that they may be used alone to determine AFI prognosis.

Introduction

Acute febrile illness (AFI) is among the most common medical problems in low- and middle-income countries (LMICs).[1, 2] Causes of AFI vary by geographical location, season, and year and can range from minor ailments to mortal illness.[3-9] Provision of disease-specific diagnostic tests to determine AFI etiology may not always be feasible.[10] Many patients with AFI can be treated in a community outpatient setting; others require referral to a higher level of care to prevent further complications and death.[11, 12]

Accurately predicting the prognosis of patients with AFI, even in high-income countries is challenging; clinical signs and symptoms may be difficult to recognize and do not consistently predict poor disease outcome.[13-19] In LMICs, limited epidemiologic data on disease incidence, lack of diagnostics, limited healthcare worker training, and overburdened healthcare systems pose additional challenges.[20, 21] Determining the severity of AFI to rapidly identify the patients who require a higher level of care is important particularly in settings where human resources and diagnostic tests are limited.[22-26]

Host biomarkers have shown promise as prognostic indicators of disease severity in conditions ranging from cancer to inflammatory bowel disease and tuberculosis.[27-29] Novel biomarkers have also been able to discriminate bacterial from viral infection in certain cohorts.[10, 30] C-reactive protein (CRP) has been used as a measure of disease severity, including AFI, for almost 90 years.[31] Other laboratory tests in common clinical use such as procalcitonin (PCT), white blood cell count, and liver function tests have been evaluated for their ability to predict the severity of AFI.[32-38]

There is now growing interest in using biomarkers to guide referral and management of patients with AFI in LMICs.[39, 40] CRP and PCT are already available for routine clinical use, thus the evidence base describing their ability to predict illness severity is already well described.[41-43] However, advances in highly multiplex immunoassay platforms, next-generation sequencing, proteomics, and metabolomics have sharply increased the availability of data describing the host response to AFI and the ability to predict its severity.[44-47] We performed a scoping review of recently published literature

on the performance of novel host biomarkers to predict the severity of AFI with the aim to inform development of diagnostics for clinical use.

Methods

Search strategy

Search terms were selected to identify studies which evaluated the performance of a biomarker to classify patients at risk for severe disease or death in an analyzable population consisting entirely of patients with a syndrome known to cause AFI. A non-exhaustive list of causes of AFI in LMICs was determined by review of AFI etiology in primary papers and systematic reviews.[48-50] Tuberculosis was intentionally excluded from this list as there is already a separate evidence base describing the performance of biomarkers for tuberculosis diagnosis and prognosis.[51] Causes of AFI explicitly included in the process to select search terms included sepsis, septic shock, pneumonia, diarrhea, malaria, dengue, leptospirosis, meningitis, typhoid fever and scrub typhus. Multiple manual searches were performed to identify 32 manuscripts that met our inclusion criteria and 16 near-misses.

Controlled vocabulary and text words were identified from these manuscripts. A custom-built Python script was written to maximize the ratio of correctly identified manuscripts to incorrectly identified manuscripts using the most parsimonious search terms. Briefly, individual phrases were extracted from the titles, abstracts, and Medical Subject Heading terms for all 48 manuscripts. Each phrase was ranked for its sensitivity and specificity to identify manuscripts meeting inclusion criteria. Phrases were chosen that identified all 32 manuscripts warranting inclusion, and then sequentially eliminated to create the most parsimonious search query while not missing any manuscripts warranting inclusion. Using the selected terms, a search was conducted on February 10, 2018 in PubMed as shown in Table 1.

Study Selection

Study population

Studies analyzing populations with one or multiple, suspected or diagnosed causes of AFI were included. If the analyzed study population included patients without a suspected or diagnosed cause of

AFI, the study was excluded (i.e., patients with systemic inflammatory response syndrome without sepsis). Hospital-acquired infections were not considered cases of AFI and their presence in the study population was cause for study exclusion. Studies in which all participants had malignancy, neutropenia, or organ transplantation were excluded. Studies that included only participants under 6 months of age were not considered. Interventional trials were excluded. Papers included only English language papers published from January 1, 2013 to February 10, 2018.

Biomarkers

The studies included in this review tested one or more biomarkers produced by human hosts and used the tests to assess the severity of AFI. Studies were restricted to those reporting biomarker measurement in a routinely available body fluid - blood or its components, saliva, and urine. Biomarker measurement of bodily fluids requiring invasive sampling (e.g. cerebrospinal fluid, pleural fluid) were not considered. Biomarkers already available for routine clinical use were excluded including: complete blood count (CBC) and its components, CRP, PCT, liver function tests, kidney function tests, lactate, routine endocrine tests like thyroid function tests and cardiac biomarkers such as troponin and pro-BNP. Biomarkers reported from large-scale discovery data sets such as generated in transcriptomic or proteomic studies without subsequent validation were excluded.

Measure of severity

Only studies reporting disease severity in terms of association with mortality, organ dysfunction, established severe disease classification system, or clinical severity score were included. Studies that reported kidney injury as the only measure of disease severity were excluded. Studies reporting only biomarker levels collected after 48 hours of presentation to healthcare were excluded. If the investigators performed more than one biomarker measurement, only the earliest measurement was considered. If more than one measure of disease severity was reported, preference was given to mortality as an outcome for further analysis.

Screening Process

Each abstract was reviewed for inclusion by 2 independent reviewers. When the 2 reviewers did not reach agreement, a third reviewer was sought. After abstract review, provisionally included abstracts were subject to full text review again by 2 independent reviewers. Inter-reviewer disagreement on full-text reviews was handled similarly to abstracts. Abstract and full text review were managed using Covidence.[52]

Quality Assessment

We adapted the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2)[53] into a simplified scoring system to report the risk of bias and concern for applicability using the QUADAS-2 domains of patient selection, index test, reference standard, and flow and timing. One point each was given if the study design was prospective, participants were enrolled consecutively, and biomarker cutoff values were defined prospectively. One point each was subtracted if the study employed a case control design, participants with hospital-acquired infections were not explicitly excluded, or ambiguous methods or reporting did not explicitly exclude patients without infections.

Data Extraction

Information was extracted from individual studies into a custom Microsoft Access database. Key information extracted from each study included the level of care at the time of patient enrollment which was categorized nonexclusively as outpatient, emergency department, inpatient ward, and intensive care unit (ICU). The age groups of participants were categorized as including children under age 5 years, children between age 5 to 18 years, and adults. Non-exclusive diagnoses of the study population were recorded. Disease severity or patient outcome was recorded either categorically or in correlation with existing severity scores. Categorical patient outcome measures included death, septic shock, severe sepsis, respiratory failure, subsequent requirement for ICU admission, and severe disease according to disease-specific standards. The p-value results of significance testing of an association of a biomarker with an unfavorable patient outcome or severe disease was extracted. The strength of the association as reported by area under the curve (AUC) of the receiver operator

characteristic (ROC), logistic regression odds ratio, and Cox proportional hazard ratio. Unadjusted measures of association were recorded preferentially over adjusted measures if both were reported as adjustment was performed infrequently and heterogeneously. Correlation with severity scores including Acute Physiology and Chronic Health Evaluation-II (APACHE-II),[54] Sequential Organ Failure Assessment (SOFA), CURB-65, Mortality in Emergency Department Sepsis (MEDS), Multiple Organ Dysfunction Score (MODS), Pneumonia Severity Index (PSI) was recorded by Spearman r values, p values, and by severity score above or below a defined biomarker cutoff where reported.

Biomarker Categorization

Each reported protein biomarker was queried for its inclusion in the Uniprot database.[55] Biomarkers found to be reported by multiple names were assigned one name for further analysis. Additionally, biomarkers and their prologues or breakdown products (e.g. adrenomedullin and pro-adrenomedullin) were reported together. The function of each biomarker reported by more than one study was determined by query of Uniprot and Cytoscape as well as additional literature review.[56] A flowchart was created placing the biomarkers in their primary immunological pathway. These immunological pathways were grouped into eight immune functions to simplify analysis and interpretation. Biomarkers that had been analyzed in more than one study were grouped parsimoniously according to their immune function, where applicable: vasoactivity, proinflammatory cytokines, pathogen-associated molecular pattern (PAMP), and danger associated molecular pattern (DAMP) recognition, macrophage differentiation, leukocyte migration, innate effector mechanism, and immune regulation.

Data Analysis

For each biomarker, summary statistics describing the available evidence base to predict the severity of AFI were compiled including number of studies, combined number of participants, diagnoses, and level of care of participants. A weighted median and mean AUC was assigned for each biomarker by calculating the median and mean AUC across studies weighted by the number of participants in each study. The most promising biomarkers were shortlisted by the following criteria: inclusion of > 500 participants in ≥ 2 studies with a weighted mean AUC of > 0.70 where $\geq 75\%$ of studies showed an

association with the outcome of interest. The most promising biomarkers were further shortlisted by restricting to biomarkers with a weighted mean >0.75 . Additional sub-analyses were performed by restricting to studies performed only in LMICs that did not include participants admitted to the ICU at the time of enrollment.

Results

Search results

The initial search yielded 2,303 results (Figure 1). Abstract review identified 444 publications that provisionally fit inclusion criteria. Review of full text manuscripts eliminated 163 publications due to reasons identified in Figure 1, leaving 281 manuscripts for data review.

Study and patient characteristics

The included manuscripts reported results for 45,737 participants in 49 countries (Figure 2). The median number of participants per study was 104 (interquartile range [IQR] 53-186). Participants were classified as having 21 distinct diagnoses or clinical syndromes (Figure 3); sepsis was the most common. Among the included manuscripts, 146 (52%) studies were conducted in 24 LMICs and included 19,513 (43%) participants. Further restricting to studies conducted in LMICs in non-ICU settings, there were 11,630 participants included in 20 countries.

Diversity of evaluated biomarkers.

Data for 278 unique biomarkers were extracted. Individual publications reported a median of 1 (IQR 1-3) biomarkers. The most frequently evaluated host biomarkers were IL-6 ($n = 50$), IL-10 ($n = 33$), and TNF- α ($n = 26$). Most biomarkers ($n = 189$) were only evaluated in a single publication. There were 89 biomarkers evaluated in at least 2 studies (Supplementary Table); 52 biomarkers were evaluated in ≥ 3 studies. The most commonly studied biomarker categories were pro-inflammatory cytokines (19 biomarkers) followed by biomarkers involved in innate effector mechanisms (13 biomarkers), immune regulation (12 biomarkers), leukocyte migration (12 biomarkers) and vasoactivity (12 biomarkers) (Supplementary Figure 1).

Biomarker performance

Mortality was the reported outcome in 193 (65%) studies. Disease-specific measures of severity were the second most commonly reported outcome ($n = 53$, 18%). A measure of strength of association was reported by 168 (60%) studies. The most commonly reported measure of strength of association was AUC-ROC, which was reported in 151 (54%) of studies (Table 2). Among the biomarkers studied, there were 15 biomarkers that showed a weighted mean AUC > 0.7 studied in > 500 patients over ≥ 2 studies in which $\geq 75\%$ of studies showed an association of the biomarker with the severity outcome; 6 of which showed a weighted mean AUC > 0.75 (Figure 4). Among studies conducted in LMICs without ICU patients, there were 4 biomarkers which showed a weighted mean AUC > 0.7 studied in > 500 patients over ≥ 2 studies (Supplementary Figure 2).

Discussion

In this scoping review, we report data for 278 novel biomarkers studied in 45,737 patients with 21 causes of acute febrile illness in 281 studies from 50 countries, yielding an understanding of the current evidence base for biomarkers to predict the severity of AFI. Among the biomarkers studied, there were 6 biomarkers that showed a weighted mean AUC greater than 0.75 in more than 500 patients over more than 2 studies. The current evidence base reports mostly the performance of biomarkers predominantly in high-income intensive care settings; a smaller, yet still robust evidence base describes the performance of biomarkers outside of ICUs in LMICs. The overall performance of biomarkers suggests that although they may aid in determining the prognosis of AFI; they are insufficient as the sole predictors of severity of febrile illness.

Biomarkers in 8 categories demonstrated potentially useful performance in multiple studies exceeding 600 total study patients, suggesting their promise for future study. The number of studies exploring a biomarker, the number of participants in each study, as well as the variation in ages, disease states, and settings in which the patient population has been studied, contributes to the breadth of evidence supporting biomarker performance. As AUC is a single number that accounts for both sensitivity and specificity independent of a cutoff value, we chose to use AUC as the measure to aggregate biomarker performance across studies. We found several biomarkers including proadrenomedullin and copeptin

that had sufficiently large evidence bases and favorable performance to suggest that they may warrant further exploration as fever triage candidates. Proadrenomedullin has been studied in >4,800 patients in 18 studies across varied geographical settings, age groups, and disease states. Almost all of the studies of proadrenomedullin reported an AUC, with a weighted mean AUC of 0.78. Similarly, copeptin has been studied in >3,000 patients with a weighted mean AUC of 0.81.

Current research reporting the performance of biomarkers to predict AFI severity has not been primarily conducted in the settings where their potential impact is greatest. The highest burden of AFI is in tropical LMICs¹ where inpatient and critical care beds are limited. The most likely intended use for a fever triage test would be among patients under assessment for the appropriate level care where the capacity for clinical assessment is limited. Most of the studies in this review were conducted in ICU settings where patients already likely show frank clinical signs of the requirement for a higher level of care. Almost half of the included studies were conducted in the US and Europe which may not represent the population and healthcare systems that would most benefit from an AFI severity biomarker.

Readily available clinical signs such as respiratory rate and simple scoring systems such as qSOFA are already used routinely to assess AFI clinical severity. The evidence compiled in this scoping review does not suggest that single biomarkers, even the most promising ones, perform sufficiently well enough to clearly supercede existing clinical severity assessment tools such as qSOFA, which has an AUC of 0.66-0.85 to predict severe or mortal AFI.[57, 58] However, unlike subjective clinical assessment, biomarkers may be standardized and not rely on the skill of a clinical assessor. Earlier experience with combining qSOFA and biomarkers have shown promising results, with a higher AUC for the combination than any individual method of prognosis assessment.[59-61]

There are several limitations to this scoping review. The available data are highly heterogeneous in patient population, disease state, and study design. The majority of observations in this scoping review only reported the presence or absence of an association with an unfavorable outcome; only 48% of observations reported the strength of the association. Only studies published in the last 5 years were analyzed which may miss earlier and important contributions. The quality assessment that we

performed was basic. Because we required only that patients be suspected or diagnosed to have a syndrome that is known to cause AFI, a significant proportion of participants may not have been febrile at the time of study enrollment which may limit the generalizability of the findings to the intended use case of evaluating febrile patients.

Biomarkers which demonstrated promising performance but have a limited evidence base such as serum Triggering Receptor Expressed on Myeloid cells-1 may show further promise as triage candidates when studied in more patients. A formal meta-analysis of pro-inflammatory cytokines and vasoactive biomarkers, or a prospective study in larger and more varied cohorts would provide higher confidence evidence than obtained in this scoping review. Additionally, it would allow the formulation and testing of algorithms with overall improved performance characteristics. Finally, prospective evaluation of biomarkers in clinical practice will be necessary to show that determination of AFI prognosis impacts meaningful treatment outcomes.

AFI impacts a significant portion of the global population every year. This scoping review synthesized a large amount of data to shortlist biomarkers that may be useful in the assessment of AFI severity as triage test candidates. Further research is necessary to determine whether biomarkers can be integrated with existing fever assessment guidelines to improve clinical practice and patient outcome.

Funding

This work was supported by the National Institute of Allergy and Infectious Diseases of the National Institutes of Health [Grant UM1AI104681 for M. L. R] . FIND receives funding for improving AFI diagnostics from the Government of the United Kingdom; FIND was supported in this work by the Doctors Without Borders/Médecins Sans Frontières Febrile Illness Diagnostic Program. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Acknowledgments

We are grateful to Rahul Rajkumar for assistance in data abstraction. We further would like to thank Tomas Jansen, Teri Roberts and Arlene Chua from Médecins Sans Frontières for their input in study design. There are no reported conflicts of interest.

References

1. Vasanthan A, Dilip Mathai KM, Anand M, et al. Acute undifferentiated fever in India: a multicentre study of aetiology and diagnostic accuracy. *BMC Infectious Diseases* **2017**; 17(1): 665.
2. Acestor N, Cooksey R, Newton PN, et al. Mapping the aetiology of non-malarial febrile illness in Southeast Asia through a systematic review--terra incognita impairing treatment policies. *PLoS One* **2012**; 7(9): e44269.
3. Prasad N, Sharples KJ, Murdoch DR, Crump JA. Community Prevalence of Fever and Relationship with Malaria Among Infants and Children in Low-Resource Areas. *The American Journal of Tropical Medicine and Hygiene* **2015**; 93(1): 178-80.
4. Black RE, Cousens S, Johnson HL, et al. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *The Lancet* **2010**; 375(9730): 1969-87.
5. Chipwaza B, Mhamphi GG, Ngatunga SD, et al. Prevalence of Bacterial Febrile Illnesses in Children in Kilosa District, Tanzania. *PLOS Neglected Tropical Diseases* **2015**; 9(5): e0003750.
6. Crump JA, Morrissey AB, Nicholson WL, et al. Etiology of Severe Non-malaria Febrile Illness in Northern Tanzania: A Prospective Cohort Study. *PLOS Neglected Tropical Diseases* **2013**; 7(7): e2324.
7. Capeding MR, Chua MN, Hadinegoro SR, et al. Dengue and Other Common Causes of Acute Febrile Illness in Asia: An Active Surveillance Study in Children. *PLOS Neglected Tropical Diseases* **2013**; 7(7): e2331.
8. Punjabi NH, Taylor WRJ, Murphy GS, et al. Etiology of Acute, Non-Malaria, Febrile Illnesses in Jayapura, Northeastern Papua, Indonesia. *The American Journal of Tropical Medicine and Hygiene* **2012**; 86(1): 46-51.
9. Singhi S, Rungta N, Nallasamy K, et al. Tropical fevers in Indian intensive care units: A prospective multicenter study. *Indian Journal of Critical Care Medicine* **2017**; 21(12): 811-8.
10. Lubell Y, Althaus T, Blacksell SD, et al. Modelling the Impact and Cost-Effectiveness of Biomarker Tests as Compared with Pathogen-Specific Diagnostics in the Management of Undifferentiated Fever in Remote Tropical Settings. *PLOS ONE* **2016**; 11(3): e0152420.
11. Parry JAC, Sandy G, Christopher M. Management of adolescents and adults with febrile illness in resource limited areas. **2011**.
12. WHO. IMAI acute care: guidelines for first-level facility health workers at health centre and district outpatient clinic. **2009**.
13. Kapasi AJ, Dittrich S, Gonzalez IJ, Rodwell TC. Host Biomarkers for Distinguishing Bacterial from Non-Bacterial Causes of Acute Febrile Illness: A Comprehensive Review. *PLoS One* **2016**; 11(8): e0160278.
14. Nadjm B, Amos B, Mtove G, et al. WHO guidelines for antimicrobial treatment in children admitted to hospital in an area of intense *Plasmodium falciparum* transmission: prospective study. *BMJ* **2010**; 340.
15. Drain PK, Hyle EP, Noubary F, et al. Evaluating Diagnostic Point-of-Care Tests in Resource-Limited Settings. *The Lancet infectious diseases* **2014**; 14(3): 239-49.
16. Bell D. FIND Acute Febrile Syndrome Strategy. **2012**.
17. Van den Bruel A, Haj-Hassan T, Thompson M, Buntinx F, Mant D. Diagnostic value of clinical features at presentation to identify serious infection in children in developed countries: a systematic review. *The Lancet* **2010**; 375(9717): 834-45.

18. 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.
19. Lee ACC, Chandran A, Herbert HK, et al. Treatment of Infections in Young Infants in Low- and Middle-Income Countries: A Systematic Review and Meta-analysis of Frontline Health Worker Diagnosis and Antibiotic Access. *PLOS Medicine* **2014**; 11(10): e1001741.
20. Hopkins H, Bruxvoort KJ, Cairns ME, et al. Impact of introduction of rapid diagnostic tests for malaria on antibiotic prescribing: analysis of observational and randomised studies in public and private healthcare settings. *BMJ* **2017**; 356.
21. Bhandari L DS. Health Infrastructure in Rural India. **2007**.
22. Croft P, Altman DG, Deeks JJ, et al. The science of clinical practice: disease diagnosis or patient prognosis? Evidence about “what is likely to happen” should shape clinical practice. *BMC Medicine* **2015**; 13: 20.
23. Silva MM, Gil LH, Marques ET, Jr., Calzavara-Silva CE. Potential biomarkers for the clinical prognosis of severe dengue. *Mem Inst Oswaldo Cruz* **2013**; 108(6): 755-62.
24. 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(8): e0161721.
25. Kosack CS, Page AL, Klatser PR. A guide to aid the selection of diagnostic tests. *Bull World Health Organ* **2017**; 95(9): 639-45.
26. Escadafal C, Nsanzabana C, Archer J, Chihota V, Rodriguez W, Dittrich S. New Biomarkers and Diagnostic Tools for the Management of Fever in Low- and Middle-Income Countries: An Overview of the Challenges. *Diagnostics* **2017**; 7(3): 44.
27. Lee S, Eo W, Jeon H, Park S, Chae J. Prognostic Significance of Host-related Biomarkers for Survival in Patients with Advanced Non-Small Cell Lung Cancer. *Journal of Cancer* **2017**; 8(15): 2974-83.
28. Rachow A, Heinrich N, Geldmacher C. Early Identification of Progressive TB Disease Using Host Biomarkers. *EBioMedicine* **2015**; 2(2): 107-8.
29. Barnes EL, Burakoff R. New Biomarkers for Diagnosing Inflammatory Bowel Disease and Assessing Treatment Outcomes. *Inflammatory bowel diseases* **2016**; 22(12): 2956-65.
30. Tsalik EL, Henao R, Nichols M, et al. Host gene expression classifiers diagnose acute respiratory illness etiology. *Science Translational Medicine* **2016**; 8(322): 322ra11-ra11.
31. Tillett WS, Francis T. SEROLOGICAL REACTIONS IN PNEUMONIA WITH A NON-PROTEIN SOMATIC FRACTION OF PNEUMOCOCCUS. *The Journal of experimental medicine* **1930**; 52(4): 561-71.
32. Erdman LK, D’Acromont V, Hayford K, et al. Biomarkers of Host Response Predict Primary End-Point Radiological Pneumonia in Tanzanian Children with Clinical Pneumonia: A Prospective Cohort Study. *PLOS ONE* **2015**; 10(9): e0137592.
33. Chin CD, Laksanasopin T, Cheung YK, et al. Microfluidics-based diagnostics of infectious diseases in the developing world. *Nature Medicine* **2011**; 17: 1015.
34. Lubell Y, Blacksell SD, Dunachie S, et al. Performance of C-reactive protein and procalcitonin to distinguish viral from bacterial and malarial causes of fever in Southeast Asia. *BMC Infectious Diseases* **2015**; 15(1): 511.
35. Page A-L, de Rekeneire N, Sayadi S, et al. Diagnostic and Prognostic Value of Procalcitonin and C-Reactive Protein in Malnourished Children. *Pediatrics* **2014**; 133(2): e363-e70.
36. Sambursky R, Shapiro N. Evaluation of a combined MxA and CRP point-of-care immunoassay to identify viral and/or bacterial immune response in patients with acute febrile respiratory infection. *European Clinical Respiratory Journal* **2015**; 2: 10.3402/ecrj.v2.28245.
37. Wangrangsimakul T, Althaus T, Mukaka M, et al. Causes of acute undifferentiated fever and the utility of biomarkers in Chiangrai, northern Thailand. *PLOS Neglected Tropical Diseases* **2018**; 12(5): e0006477.
38. Do NTT, Ta NTD, Tran NTH, et al. Point-of-care C-reactive protein testing to reduce inappropriate use of antibiotics for non-severe acute respiratory infections in Vietnamese

- primary health care: a randomised controlled trial. *The Lancet Global Health* **2016**; 4(9): e633-e41.
39. Escadafal C, Nsanzabana C, Archer J, Chihota V, Rodriguez W, Dittrich S. New Biomarkers and Diagnostic Tools for the Management of Fever in Low- and Middle-Income Countries: An Overview of the Challenges. *Diagnostics* **2017**; 7(3): 9.
 40. World Health Organization. WHO informal consultation on fever management in peripheral health care settings: A global review of evidence and practice. Available at: <https://www.who.int/malaria/publications/atoz/9789241506489/en/>. Accessed Dec.
 41. Liu D, Su L, Han G, Yan P, Xie L. Prognostic Value of Procalcitonin in Adult Patients with Sepsis: A Systematic Review and Meta-Analysis. *PLOS ONE* **2015**; 10(6): e0129450.
 42. Lobo SMA, Lobo FRM, Bota DP, et al. C-reactive protein levels correlate with mortality and organ failure in critically ill patients. *Chest* **2003**; 123(6): 2043-9.
 43. Chalmers JD, Singanayagam A, Hill AT. C-reactive protein is an independent predictor of severity in community-acquired pneumonia. *Am J Med* **2008**; 121(3): 219-25.
 44. 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. *The Lancet Infectious Diseases* **2017**; 17(4): 431-40.
 45. Sweeney TE, Khatri P. Generalizable Biomarkers in Critical Care: Toward Precision Medicine. *Critical Care Medicine* **2017**; 45(6): 934-9.
 46. 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(8): 835-45.
 47. Gordo-Remartínez S, Calderón-Moreno M, Fernández-Herranz J, et al. Usefulness of midregional proadrenomedullin to predict poor outcome in patients with community acquired pneumonia. *PloS one* **2015**; 10(6): e0125212.
 48. Shrestha P, Roberts T, Homsana A, et al. Febrile illness in Asia: gaps in epidemiology, diagnosis, management for informing health policy. *Clin Microbiol Infect* **2018**.
 49. Maze MJ, Bassat Q, Feasey NA, Mandomando I, Musicha P, Crump JA. The epidemiology of febrile illness in sub-Saharan Africa: implications for diagnosis and management. *Clin Microbiol Infect* **2018**.
 50. Infectious Diseases Data Observatory. Non-malaria febrile illness (NMFI) surveyor. Available at: www.wwarn.org/surveyor/NMFIv3/. Accessed 1 Jun.
 51. Walzl G, McNerney R, du Plessis N, et al. Tuberculosis: advances and challenges in development of new diagnostics and biomarkers. *The Lancet Infectious Diseases* **2018**; 18(7): e199-e210.
 52. Covidence systematic review software. Melbourne, Australia: Veritas Health Innovation,.
 53. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies. *Annals of Internal Medicine* **2011**; 155(8): 529-U104.
 54. Knaus WAM, Draper EAM, Wagner DPP, Zimmerman JEM. APACHE II: A severity of disease classification system. *Crit Care Med* **1985**; 13(10): 818-29.
 55. The UniProt Consortium. UniProt: the universal protein knowledgebase. *Nucleic Acids Research* **2017**; 45(D1): D158-D69.
 56. Su G, Morris JH, Demchak B, Bader GD. Biological network exploration with Cytoscape 3. *Current protocols in bioinformatics* **2014**; 47: 8.13.1-8.24.
 57. Boillat-Blanco N, Mbarack Z, Samaka J, et al. Prognostic value of quickSOFA as a predictor of 28-day mortality among febrile adult patients presenting to emergency departments in Dar es Salaam, Tanzania. *PLOS ONE* **2018**; 13(6): e0197982.
 58. Churpek MM, Snyder A, Han X, et al. Quick Sepsis-related Organ Failure Assessment, Systemic Inflammatory Response Syndrome, and Early Warning Scores for Detecting Clinical Deterioration in Infected Patients outside the Intensive Care Unit. *American Journal of Respiratory and Critical Care Medicine* **2017**; 195(7): 906-11.
 59. Kim MW, Lim JY, Oh SH. Mortality prediction using serum biomarkers and various clinical risk scales in community-acquired pneumonia. *Scandinavian Journal of Clinical and Laboratory Investigation* **2017**; 77(7): 486-92.

60. Ho KM, Lan NSH. Combining quick Sequential Organ Failure Assessment with plasma lactate concentration is comparable to standard Sequential Organ Failure Assessment score in predicting mortality of patients with and without suspected infection. *Journal of Critical Care* **2017**; 38: 1-5.
61. Kutz A, Hausfater P, Amin D, et al. The TRIAGE-ProADM Score for an Early Risk Stratification of Medical Patients in the Emergency Department - Development Based on a Multi-National, Prospective, Observational Study. *PLoS ONE* **2016**; 11(12): e0168076.

Table 1: Terms included in PubMed search

Category	Terms
Biomarker	biomarkers[mh] OR protein precursors[mh] OR metabolome[mh] OR biomarker[tiab] OR biomarkers[tiab]
Severity	severity of illness index[mh] OR survival analysis[mh] OR prognosis[mh] OR prognostic[tiab] OR mortality[tiab]
Acute febrile illness	sepsis[mh] OR pneumonia[mh] OR fever[mh] OR malaria[mh] OR dengue OR typhoid fever[mh] OR leptospirosis OR diarrhea[mh] OR influenza[tiab] OR meningitis[tiab]
Other	English[la] NOT review[pt] NOT (animals[mh] NOT humans[mh])
Each category was joined by an AND statement	

Table 2: Evidence base and performance of leading^a biomarkers to predict acute febrile illness severity

Biomarker	Studies, n ^b	Patients studied	Studies reporting positive association with severity, n (%)	Studies reporting			AUC	Number of patients included in AUC	Weighted mean AUC ^c	Random effects meta-analysis	I ²
				Odds ratio	Logistic regression	Hazard ratio					
copeptin	9	3326	9 (100)	1	0	0	5	2953	0.81	0.79	0.67
IL-8	18	2282	17 (94)	1	2	0	6	755	0.73	0.74	0.65
NGAL	9	2770	8 (89)	1	3	0	6	1522	0.71	0.66	0.89
PAI-1	5	925	5 (100)	0	1	0	3	612	0.73	0.75	0.92
PDL-1	2	730	2 (100)	1	1	0	2	730	0.72	0.72	0
PERSEVERE	3	802	3 (100)	0	0	0	3	802	0.80	0.82	0
proADM	18	4724	17 (94)	4	0	1	14	4261	0.78	0.80	0.58
proANP	4	3208	4 (100)	0	0	0	3	3123	0.80	0.80	0
PTX-3	8	1434	7 (88)	0	0	1	4	890	0.72	0.76	0
sTREM-1	8	997	7 (88)	1	1	1	6	569	0.75	0.8	0.77
suPAR	5	529	5 (100)	0	1	0	5	529	0.71	0.68	0
TIMP-1	6	1627	6 (100)	0	2	0	3	1212	0.71	0.76	0.93
TM	4	1647	3 (75)	1	1	0	2	1253	0.81	0.81	0.9
VCAM-1	11	1174	10 (91)	2	0	0	8	878	0.73	0.73	0.9
YKL-40	3	521	3 (100)	0	0	0	3	521	0.84	0.87	0.91

^aReported in >500 participants in ≥2 studies with a weighted mean AUC of >0.70. ^bReferences are available in the supplementary materials. ^cWeighted by number of participants in study. AUC; Area under the curve; IL-8, interleukin-8; NGAL, neutrophil gelatinase-associated lipocalin; PAI-1, plasminogen activator inhibitor-1; PDL-1, programmed cell death ligand-1; PERSEVERE, pediatric sepsis biomarker risk model; proADM, pro-adrenomedullin; proANP, pro-atrial natriuretic peptide; PTX-3, pentraxin- 3; sTREM-1, serum triggering receptor expressed on myeloid cells-1; suPAR, soluble urokinase-type plasminogen activator receptor; TIMP-1, metalloproteinase inhibitor-1; TM, thrombomodulin; vCAM-1, vascular cell adhesion molecule-1; YKL-40, chitinase-3-like protein-1

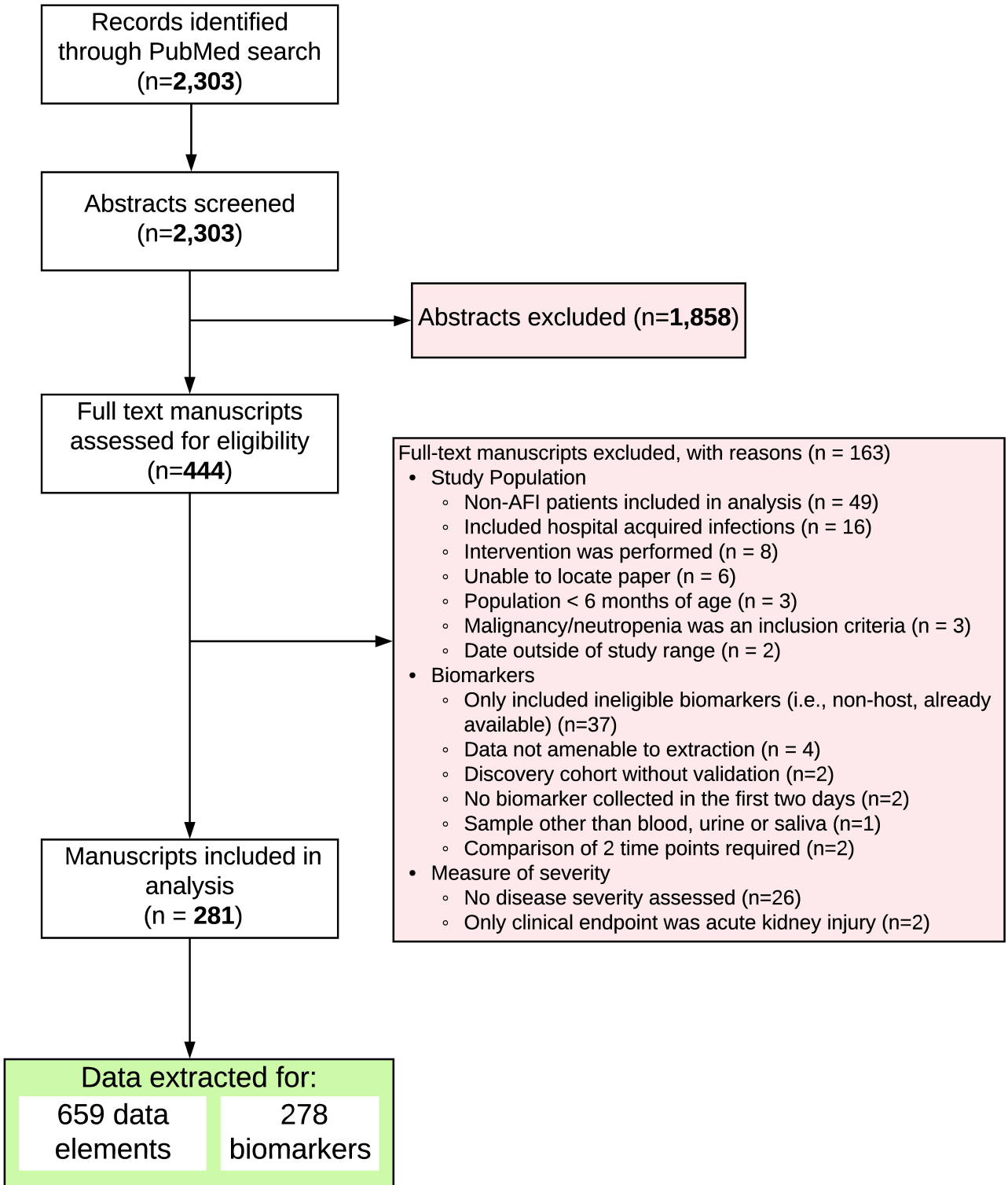
Figure Legends

Figure 1: Identification, screening, eligibility assessment, and data extraction

Figure 2: Number of manuscripts included by country

Figure 3: Diagnoses of participants, N = 45,737

Figure 4: Participants studied, effect size, and quality of leading biomarkers to predict the severity of acute febrile illness. Each shaded tile represents one study. The width of the tile is proportional to the number of subjects enrolled; the height is proportional to the effect size of the association as measured by AUC or odds ratio.



Participant diagnosis

