12-1-2021

Comparing the Diagnostic Accuracy of Clinician Judgment to a Novel Host Response Diagnostic for Acute Respiratory Illness

Ian S. Jaffe

Anja K. Jaehne
Henry Ford Health, ajaehne2@hfhs.org

Eugenia Quackenbush

Emily R. Ko

Emanuel P. Rivers

See next page for additional authors

Recommended Citation

Jaffe IS, Jaehne AK, Quackenbush E, Ko ER, Rivers EP, McClain MT, Ginsburg GS, Woods CW, and Tsalik EL. Comparing the Diagnostic Accuracy of Clinician Judgment to a Novel Host Response Diagnostic for Acute Respiratory Illness. Open Forum Infect Dis 2021; 8(12):ofab564.

Authors

Ian S. Jaffe, Anja K. Jaehne, Eugenia Quackenbush, Emily R. Ko, Emanuel P. Rivers, Micah T. McClain, Geoffrey S. Ginsburg, Christopher W. Woods, and Ephraim L. Tsalik

Open Forum Infectious Diseases

MAJOR ARTICLE

Comparing the Diagnostic Accuracy of Clinician Judgment to a Novel Host Response Diagnostic for Acute Respiratory Illness

Ian S. Jaffe,[1] Anja K. Jaehne,[2] Eugenia Quackenbush,[3] Emily R. Ko,[1] Emanuel P. Rivers,[2] Micah T. McClain,[1,4] Geoffrey S. Ginsburg,[1] Christopher W. Woods,[1,4] and Ephraim L. Tsalik[1,5]

[1]Duke Center for Applied Genomics & Precision Medicine, Department of Medicine, Duke University School of Medicine, Durham, North Carolina, USA, [2]Department of Emergency Medicine, Henry Ford Hospital, Wayne State University, Detroit, Michigan, USA, [3]Department of Emergency Medicine, University of North Carolina Medical Center, Chapel Hill, North Carolina, USA, [4]Medical Service, Durham Veterans Affairs Health Care System, Durham, North Carolina, USA, and [5]Emergency Medicine Service, Durham Veterans Affairs Health Care System, Durham, North Carolina, USA

Background. Difficulty discriminating bacterial from viral infections drives antibacterial misuse. Host gene expression tests discriminate bacterial and viral etiologies, but their clinical utility has not been evaluated.

Methods. Host gene expression and procalcitonin levels were measured in 582 emergency department participants with suspected infection. We also recorded clinician diagnosis and clinician-recommended treatment. These 4 diagnostic strategies were compared with clinical adjudication as the reference. To estimate the clinical impact of host gene expression, we calculated the change in overall Net Benefit (ΔNB; the difference in Net Benefit comparing 1 diagnostic strategy with a reference) across a range of prevalence estimates while factoring in the clinical significance of false-positive and -negative errors.

Results. Gene expression correctly classified bacterial, viral, or noninfectious illness in 74.1% of subjects, similar to the other strategies. Clinical diagnosis and clinician-recommended treatment revealed a bias toward overdiagnosis of bacterial infection resulting in high sensitivity (92.6% and 94.5%, respectively) but poor specificity (67.2% and 58.8%, respectively), resulting in a 33.3% rate of inappropriate antibacterial use. Gene expression offered a more balanced sensitivity (79.0%) and specificity (80.7%), which corresponded to a statistically significant improvement in average weighted accuracy (79.9% vs 71.5% for procalcitonin and 76.3% for clinician-recommended treatment; $P < .0001$ for both). Consequently, host gene expression had greater Net Benefit in diagnosing bacterial infection than clinician-recommended treatment (ΔNB = 6.4%) and procalcitonin (ΔNB = 17.4%).

Conclusions. Host gene expression–based tests to distinguish bacterial and viral infection can facilitate appropriate treatment, improving patient outcomes and mitigating the antibacterial resistance crisis.

Keywords. bacterial infection; clinical decision-making; gene expression; procalcitonin; diagnostic test.

Acute respiratory illness (ARI) is one of the most common complaints seen in emergency department (ED) and outpatient settings but remains difficult to treat accurately [1–4]. Possible causes of ARI include bacterial infection, viral infection, bacterial/viral co-infection, and noninfectious etiologies, such as asthma, allergic rhinitis, postinfection cough, and chronic obstructive pulmonary disease (COPD) [5]. Diagnostic uncertainty leads to prolonged ED visits and improper antibiotic use due to concerns about missing bacterial infection, which has the unintended consequence of driving antibacterial resistance, worsening health outcomes, and increasing costs [6–11]. Even with recent improvements in antibiotic stewardship, 30% of prescribed antibacterials in the United States are considered inappropriate [12]. Conversely, knowing when to appropriately give antibacterials and antivirals like oseltamivir is important as early intervention can significantly reduce morbidity and mortality [13, 14].

In recent years, emerging technologies have made host response–based diagnostics a viable option to discriminate bacterial, viral, and noninfectious illness [15]. Initially, procalcitonin was examined as a diagnostic biomarker for bacterial infections, but multigene transcriptional panels have proven superior at discriminating bacterial from viral etiologies and even predicting disease before clinical symptoms [16–21]. We previously discovered a host gene expression signature that differentiates bacterial ARI, viral ARI, and noninfectious illness with an overall accuracy of 87% when compared with expert adjudications [22]. This host response signature has also been developed into a rapid research use–only test (45 minutes) providing accurate discrimination of bacterial and viral etiologies directly from a blood sample [23].

Received 8 October 2021; editorial decision 25 October 2021; accepted 2 November 2021; published online 10 November 2021.

Correspondence: Ephraim Tsalik, MD, MHS, PhD, DUMC 102359, Durham, NC 27710 (e.t@duke.edu).

Open Forum Infectious Diseases®2021

https://doi.org/10.1093/ofid/ofab564

Whereas clinical utility is best measured using prospective clinical trials, a retrospective approach affords a readily available estimate of potential clinical utility. We present evidence from a large population of ED patients showing that a host gene expression test can augment current clinical practice to improve diagnostic accuracy and clinical care. This is the first such analysis to retrospectively define the potential clinical impact of a transcriptional biomarker for bacterial/viral discrimination in ARI.

METHODS

Patient Consent

All studies were approved by relevant institutional review boards and in accordance with the Declaration of Helsinki. All subjects or their legally authorized representatives provided written informed consent.

Study Design

Subjects analyzed in this study derived from a prior publication evaluating host gene expression measured using quantitative reverse transcription polymerase chain reaction (qRT-PCR) on a research use–only BioFire system [23]. That study reported results for 623 participants. However, this study focused on 582 after excluding those with inadequate medical record information, fungal infection, or incomplete data. Details regarding enrollment criteria and study sites can be found in the Supplementary Methods.

Reference Standard

As previously described, expert adjudications served as the reference standard [23]. Two independent adjudications were conducted by emergency medicine, infectious disease, pulmonary/critical care, or hospital medicine specialists using patient charts and additional microbiological test data that were reviewed retrospectively. This included clinical data (notes, laboratory testing, subject-reported symptoms, and radiology, as available) from the enrollment and follow-up visits [22]. Supplemental etiology testing included the BinaxNOW *S. pneumoniae* urinary antigen test (Alere) and multiplex respiratory pathogen testing (ResPlex V2.0, Qiagen; Respiratory Viral Panel, Luminex; or Respiratory Pathogen Panel, Luminex). As such, adjudicators had access to more information than would be available to clinicians in real time. Subjects were classified as having a bacterial infection, viral infection, bacterial–viral co-infection, or a noninfectious process. Disagreements were reconciled by panel review consisting of at least 3 adjudicators. Subjects adjudicated as bacterial–viral co-infection were treated as bacterial, as both would lead to antibacterial therapy [24]. Adjudications were blind to any host gene expression data and procalcitonin concentrations, which were not available at the enrolling institutions during the study period.

Biosignature Creation and Probability Determination

No new gene expression or procalcitonin data were generated for this analysis. Instead, we relied upon data generated in previously published research. Details regarding measurement of the host gene expression signature and procalcitonin can be found in the Supplementary Methods. We used the established procalcitonin threshold of >0.25 μg/L to denote a bacterial infection [25]. Probabilities for bacterial, viral, and noninfectious illness using the gene expression test were determined by the predictive algorithm as previously published [23]. Host gene expression samples were classified as bacterial, viral, co-infection, or noninfectious illness based on whether the predicted probabilities of bacterial and viral infection exceeded the specified thresholds [23].

Medical Record Review

Medical records were reviewed to extract the variables listed in Supplementary Table 1. This included the clinician's documented diagnosis (bacterial, viral, noninfectious, inconclusive, or unknown). The clinician-diagnosed etiology was considered inconclusive if the clinician diagnosis was a symptom (eg, cough) or included multiple unresolved diagnoses of distinct etiologies (eg, bacterial pneumonia vs viral upper respiratory infection). Visit outcome (treated and released, admitted, or deceased), length of hospitalization, and length of intensive care unit stay were also recorded, if applicable. Clinician treatment was categorized as antiviral or antibacterial. The same information was collected for any other health care visits 30 days before and after enrollment.

Statistical Analysis

Statistical analysis was conducted in R, version 3.6 (Foundation for Statistical Computing, Vienna, Austria). The 4 diagnostic strategies were (1) clinician diagnosis, (2) clinician-recommended treatment, (3) procalcitonin-defined etiology (bacterial vs nonbacterial), and (4) host gene expression–based diagnosis (bacterial, viral, co-infection, or no infection). These were all compared with the expert panel adjudication as the reference standard (Figure 1). "Inconclusive" clinician-diagnosed etiologies (ie, cases in which the clinician only provided a differential diagnosis) were assigned an etiology based on the clinician-recommended treatment (eg, prescription for antibacterials would classify the subject as having a bacterial infection diagnosis).

Sensitivity and specificity were calculated using adjudication as the reference standard and compared using McNemar's test. Average weighted accuracy (AWA)—a metric of test utility accounting for the relative importance of sensitivity and specificity in a range of disease prevalence—was calculated for bacterial and viral discrimination as previously described [24]. Based on a previously reported clinician survey, a relative importance of 0.25 (indicating that sensitivity is 4 times more important

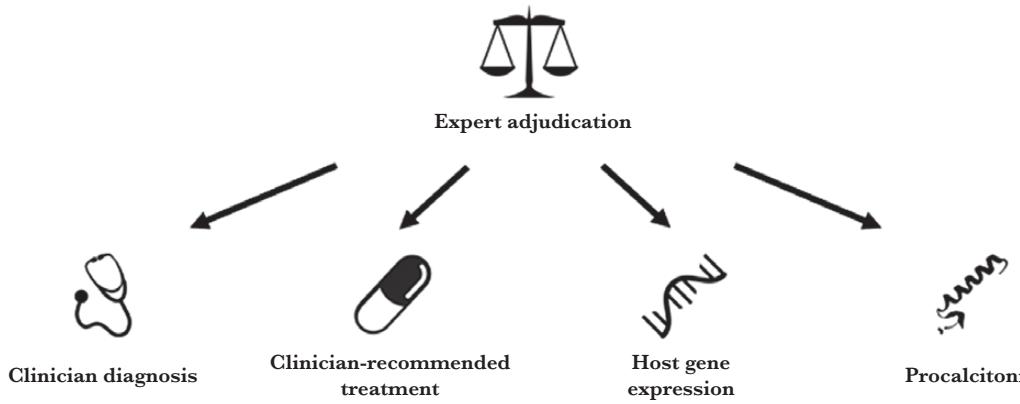

Figure 1. Summary of major diagnostic approaches. Blinded expert adjudications were used as the reference standard to compare performance characteristics of 4 strategies. The arrows denote the 4 diagnostic strategies where expert adjudication served as the reference standard. The clinician diagnosis was based on the clinician's assessment and discharge diagnosis documented in clinical notes. Clinician-recommended treatment was defined as bacterial if antibacterial drugs were prescribed, viral if antiviral drugs or supportive care was prescribed, or noninfectious if a noninfectious disease process was treated. The host gene expression test results were based on probability thresholds for bacterial vs nonbacterial infection and viral vs nonviral infection. Procalcitonin concentrations >0.25 μg/L were considered to indicate a bacterial infection.

than specificity) and a prevalence estimate of 10%–30% were used to compute bacterial AWA [24]. For viral AWA, a relative importance of 2 (indicating that specificity is twice as important as sensitivity) and a prevalence estimate of 50%–80% were used. Immunocompetent and immunocompromised subjects were also compared with respect to overall accuracy, sensitivity, and specificity using Barnard's exact test. Change in Net Benefit (ΔNB) was calculated as previously described, using the same relative importance and prevalence estimates as for AWA [26, 27]. Net Benefit (NB) incorporates a relative weighting of severity for false negatives and false positives to estimate the percentage of patients benefiting from using a new diagnostic as the sole means of diagnosing an illness. NB is defined as True Positives + True Negatives – (False Negatives × Harm_{FN}) – (False Positives × Harm_{FP}), where the Harm_{FN} is the harm of an individual going untreated (Profit/Loss ratio) and Harm_{FP} is its reciprocal [27]. In short, $\text{NB} = \text{TP} + \text{TN} - \text{FN}\left(\frac{\text{P}}{\text{L}}\right) - \text{FP}\left(\frac{\text{L}}{\text{P}}\right)$. The ΔNB is the difference in Net Benefit when comparing 1 diagnostic strategy to a reference.

RESULTS

Cohort Characteristics

The study cohort of 582 subjects was 52.4% female, 51.2% White, 45.0% Black/African American, and 3.8% other self-reported race, with a mean age (range) of 47.5 (14–94) years. Demographics of subjects grouped by expert-adjudicated etiology are shown in Supplementary Table 2. There were 115 subjects (19.8%) who had additional visits with a health care provider for the same illness within 30 days of enrollment. A medical history of lung disease was present in 226 (38.8%) subjects (most commonly asthma [17.4%] and COPD [11.2%]). Clinical adjudication assigned 271 (46.6%) subjects as having a bacterial infection (including 62 bacterial–viral co-infections), 208 (35.7%) with a viral infection, and 103 (17.7%) with a noninfectious illness. Viral infections had the lowest hospital admission rates (20.7%) and shortest mean length of stay (1.1 days). The majority (60.5%) of subjects admitted with viral etiologies had a history of noninfectious respiratory illness such as asthma or COPD and were older than subjects who were treated and released (mean, 52 vs 40 years; $P < .0001$). Conversely, subjects with bacterial infections and noninfectious illness had high admission rates (69.4% and 66.0%, respectively) and longer mean lengths of stay (5.8 days and 4.6 days, respectively). Of the subjects admitted, 6 died during their hospital admission (5 with bacterial etiologies and 1 with a noninfectious etiology).

Diagnostic Performance

Clinician diagnosis, treatment plan, host gene expression, and procalcitonin were compared with adjudication as the reference standard (Table 1). As adjudicators had access to more clinical information including supplemental microbiological testing, results from clinically ordered cultures, and follow-up assessments, it was felt that adjudicator assignments were more accurate than clinical diagnoses made in real time. Clinicians diagnosed 75.4% of subjects with the correct etiology but reduced their diagnostic accuracy to 72.2% as reflected by their treatment plans. This treatment incongruent with the diagnosis was largely due to clinicians prescribing antibacterials despite diagnosing a viral or noninfectious condition. Clinicians showed a strong sensitivity bias toward diagnosing and treating bacterial infections, at the expense of bacterial specificity and viral sensitivity. Using clinician-recommended treatment as the source of diagnostic information, clinicians had 94.5% sensitivity (95% CI, 91.0%–96.9%) and 58.8% specificity (95% CI, 53.1%–64.4%) for bacterial etiologies. The converse was true with respect to viral infection diagnosis: 52.4% sensitivity (95% CI, 45.4%–59.4%) and 96.3% specificity (95% CI, 93.8%–98.0%).

Table 1. Comparison of Diagnostic Test Performance

Test or Approach (n = 582)	Overall Accuracy, %	Bacterial Sensitivity, %	Bacterial Specificity, %	Bacterial AWA, %	Viral Sensitivity, %	Viral Specificity, %	Viral AWA, %
Host gene expression	74.1	79.0[a] (73.6–83.7)	80.7[b,c] (75.9–85.0)	79.9[a,b] (76.6–83.1)	76.0[b,d] (69.6–81.6)	86.4 (82.5–89.7)	81.3[b,c] (77.9–84.6)
Procalcitonin	72.2	56.8 (50.7–62.8)	85.5 (81.1–89.3)	71.5 (68.0–75.0)			
Clinician diagnosis	75.4	92.6[e] (88.8–95.4)	67.2 (61.7–72.4)	79.7 (76.6–82.7)	60.6 (53.6–67.3)	94.9[e] (92.2–96.9)	78.2 (74.7–81.6)
Clinician-recommended treatment	72.2	94.5[e] (91.0–96.9)	58.8 (53.1–64.4)	76.3 (73.2–79.4)	52.4 (45.4–59.4)	96.3[e] (93.8–98.0)	74.9 (71.4–78.3)

All values are shown with 95% CIs.

Abbreviation: AWA, average weighted accuracy.

[a] $P < .0001$ compared with procalcitonin.

[b] $P < .0001$ compared with clinician-recommended treatment.

[c] $P < .0001$ compared with clinician diagnosis.

[d] $P < .001$ compared with clinician diagnosis.

[e] $P < .0001$ compared with host gene expression.

As a result of this propensity to overdiagnose and overtreat bacterial infections, clinicians prescribed antibacterials inappropriately to 128 subjects adjudicated as having viral etiologies (22.0% of all subjects and 33.3% of subjects who received antibiotics).

In contrast, host gene expression had more balanced performance characteristics for the diagnosis of bacterial infection: 79.0% sensitivity (95% CI, 73.6%–83.7%) and 80.7% specificity (95% CI, 75.9%–85.0%). Sensitivity and specificity were also more balanced using host gene expression to diagnose viral infections as compared with clinician diagnosis and treatment: 76.0% sensitivity (95% CI, 69.6%–81.6%) and 86.4% specificity (95% CI, 82.5%–89.7%). In terms of bacterial AWA, host gene expression had the best diagnostic performance of all the approaches, with a bacterial AWA of 79.9% (95% CI, 76.6%–83.1%) compared with clinician diagnosis (79.7%; 95% CI 76.6%–82.7%; $P = .70$) and clinician treatment (76.3%; 95% CI, 73.2%–79.4%; $P < .00001$). Similar results were seen for viral illness, with host gene expression having the highest viral AWA of all the approaches at 81.3% (95% CI, 77.9%–84.6%). This was significantly better than clinician diagnosis (viral AWA, 78.2%; 95% CI, 74.7%–81.6%; $P < .00001$) and clinician-recommended treatment (viral AWA, 74.9%; 95% CI, 71.4%–78.3%; $P < .00001$).

Meanwhile, procalcitonin performed worse than all other diagnostic approaches on every measure. Notably, procalcitonin had the lowest bacterial sensitivity of all the assays, 56.8% (95% CI, 50.7%–62.8%), with a corresponding specificity of 85.5% (95% CI, 81.1%–89.3%), both significantly worse than host gene expression ($P < .0001$). The AWA for bacterial infection was 71.5% (95% CI, 68.0%–75.0%), which was significantly worse than clinician diagnosis, clinician-recommended treatment, and host gene expression ($P < .00001$ for all 3. Overall, of the 4 diagnostic approaches (clinician diagnosis, clinician-recommended treatment, host gene expression, and procalcitonin), host gene expression had the highest AWA for discriminating viral and bacterial infections.

We then evaluated these diagnostic strategies in subjects with (n = 57) or without (n = 525) immunocompromising conditions (Table 2). The overall accuracies were similar for the 2 groups using both host gene expression and procalcitonin. Although accuracies were lower among immunocompromised subjects using clinician diagnosis (64.9% vs 76.6%) and clinician-recommended treatment (64.9% vs 73.0%), the difference was not statistically significant. Differences were also observed for sensitivity and specificity, which in some cases were statistically significant, as noted in Table 2.

Change in Net Benefit

In order to estimate the clinical impact of host response strategies, we calculated the change in overall Net Benefit (ΔNB). At a basic level, this value represents the percentage of patients who would benefit from a new diagnostic strategy as compared with current practice. For classifying bacterial etiologies, the base case scenario assumed a 20% prevalence of bacterial etiologies and a relative importance value (r) of 0.25 based on the same assumptions made when calculating AWA. In this scenario, host gene expression provided an improvement over clinician-recommended treatment with a ΔNB of 6.4%. In this scenario, host gene expression also provided an overall Net Benefit improvement over procalcitonin (ΔNB = 17.4%). With respect to viral ARI, the base case assumed a viral prevalence of 65% and an r value of 2. Host gene expression showed an improvement over clinician-recommended treatment with a ΔNB 12.6%. We did not calculate a ΔNB for procalcitonin pertaining to viral infection as procalcitonin is unable to differentiate viral from noninfectious illness.

We performed sensitivity analyses to evaluate the impact of prevalence as well as relative importance, r, on ΔNB for the 3 main comparisons (bacterial ΔNB of gene expression vs

Table 2. Comparison of Diagnostic Performance in Immunocompromised and Immunocompetent Subjects

Approach	Immune Status	Overall Accuracy, %	Bacterial Sensitivity, %	Bacterial Specificity, %	Bacterial AWA, %	Viral Sensitivity, %	Viral Specificity, %	Viral AWA, %
Host gene expression	Immunocompetent (n = 525)	74.5	78.0 (72.2–83.0)	81.8 (76.8–86.1)	79.9 (76.4–83.3)	76.9 (70.4–82.6)	85.8 (81.5–89.3)	81.4 (78.0–84.9)
	Immunocompromised (n = 57)	70.2	88.5 (69.9–97.6)	71.0 (52.0–85.8)	79.5 69.4–89.7)	61.5 (31.6–86.1)	90.9 (78.3–97.5)	76.6 (63.0–90.2)
Procalcitonin	Immunocompetent (n = 525)	72.4	55.9 (49.5–62.2)	86.8 (82.3–90.5)	71.7 (68.0–75.3)			
	Immunocompromised (n = 57)	70.2	65.4 (44.3–82.8)	74.2 (55.4–88.1)	69.9 (58.0–81.8)			
Clinician diagnosis	Immunocompetent (n = 525)	76.6	92.7 (88.6–95.6)	69.3[a] (63.5–74.6)	80.7 (77.5–83.9)	63.1[a] (55.9–69.9)	94.9 (91.9–97.0)	79.4 (75.8–82.9)
	Immunocompromised (n = 57)	64.9	92.3 (74.9–99.1)	48.4 (30.2–66.9)	69.9 59.6–80.2)	23.1 (5.0–53.8)	95.5 (84.5–99.4)	60.2 (48.6–71.8)
Clinician-recommended treatment	Immunocompetent (n = 525)	73.0	94.3 (90.6–96.8)	60.7[a] (54.7–66.5)	77.2 (73.9–80.4)	54.4 (4 7.1–61.5)	98.4 (93.7–98.1)	75.9 (72.3–79.4)
	Immunocompromised (n = 57)	64.9	96.2 (80.4–99.9)	41.9 (24.6–60.9)	68.5 (58.9–78.1)	23.1 (5.0–53.8)	95.5 (84.5–99.4)	60.2 (48.6–71.8)

All values are shown with 95% CIs.

Abbreviation: AWA, average weighted accuracy.

[a]$P < .05$ between immunocompetent and immunocompromised groups after comparing all test metrics between groups.

clinician-recommended treatment, bacterial ΔNB of gene expression vs procalcitonin, and viral ΔNB of gene expression vs clinician-recommended treatment) (Figures 2 and 3). In general, host gene expression ΔNB increases when the prevalence of bacterial infections declines or when viral infection prevalence rises. Examining the impact of relative importance, *r*, we found that host gene expression provided a greater Net Benefit than did procalcitonin at all values of *r*. However, if r decreased to 0.125 (willing to accept 8 false-positive bacterial diagnoses to avoid 1 false-negative), then gene expression would no longer provide a Net Benefit over clinician-recommended treatment plans. In this scenario, the perceived harm of antibiotic overuse was considered less impactful than the harm of undertreating a bacterial infection.

Impact of Diagnostic Errors

Inappropriate treatment (both under- and overtreatment) could result in repeated visits for clinical worsening or treatment complications. We therefore examined the impact of diagnostic errors and whether alternative diagnostic testing could have averted them. There were 115 subjects (19.8%) who had ARI-related visits within 30 days before or after enrollment. For

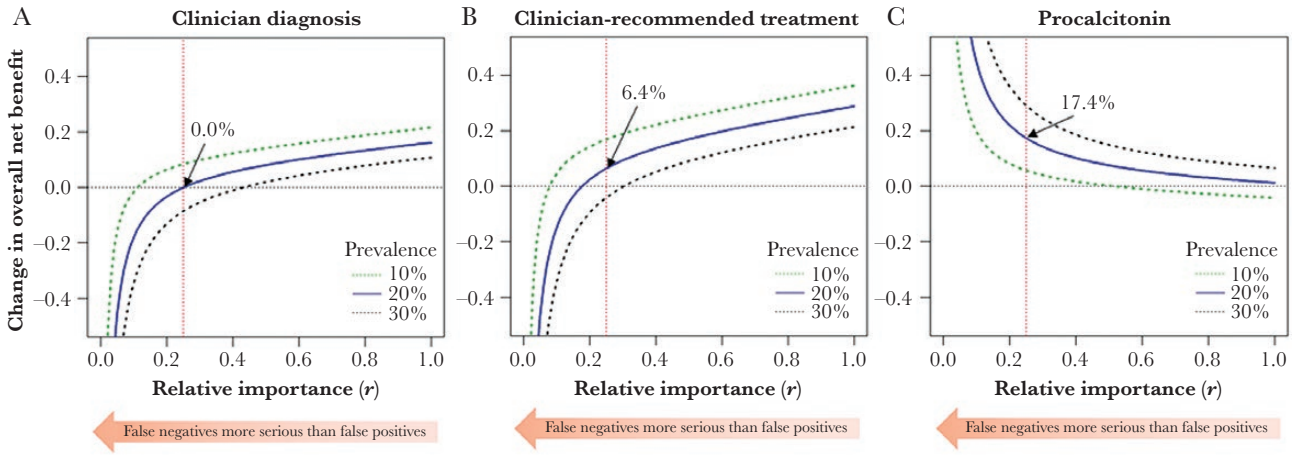

Figure 2. Change in overall Net Benefit (ΔNB) for bacterial infection. This figure shows change in overall Net Benefit curves for diagnosis of bacterial vs nonbacterial etiologies. As relative importance (*r*) decreases, there is a greater emphasis on avoiding false-negative bacterial diagnoses (failing to diagnose a bacterial infection when one is present) at the expense of false-positive bacterial diagnoses (diagnosing a bacterial infection when none is present). The base case assumes an *r* value of 0.25 (vertical red dashed line) and a bacterial prevalence of 20% (solid blue line) with +/-10% margins. A positive ΔNB indicates that the host gene expression test provides a net clinical benefit over the comparator diagnostic strategy. A, Comparison of host gene expression with clinician diagnosis. B, Comparison of host gene expression with clinician-recommended treatment. C, Comparison of host gene expression with procalcitonin.

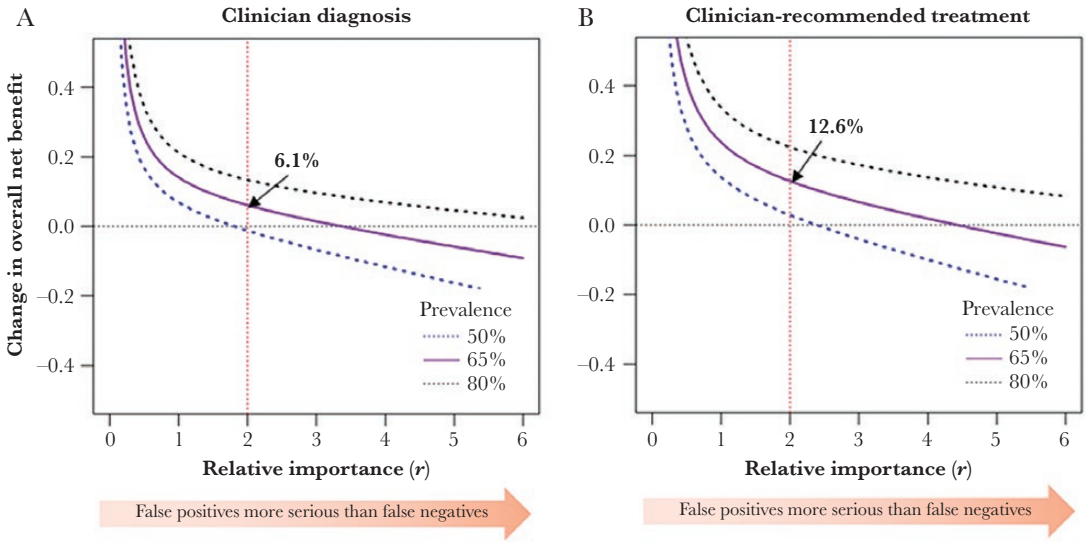

Figure 3. Viral change in overall Net Benefit (ΔNB) curves. This figure shows change in overall Net Benefit curves for diagnosis of viral vs nonviral etiologies. As relative importance (*r*) increases, there is a greater emphasis on avoiding false positives (diagnosing a viral infection when none is present) at the expense of false negatives (failing to diagnose a viral infection when one is present). The base case assumes an *r* value of 2 (vertical red dashed line) and a viral prevalence of 65% (solid purple line) with +/-15% margins. A positive ΔNB indicates that the host gene expression test provides a net clinical benefit over the comparator diagnostic strategy. A, Comparison of host gene expression with clinician diagnosis. B, Comparison of host gene expression with clinician-recommended treatment. A comparison with procalcitonin is not feasible as this biomarker does not discriminate viral from noninfectious etiologies.

31 subjects, the physician diagnosis changed with subsequent visits. In 15 of these cases, a bacterial infection was initially missed. Ten of these subjects were hospitalized due to delays in initiation of antibacterial therapy. The host gene expression test would have identified 9 as bacterial, highlighting an opportunity to start appropriate antibacterial therapy more quickly in those who need it. The 84 cases in which the diagnosis was unchanged through 30 days of follow-up were distributed among all 3 adjudicated etiologies (n = 32 bacterial/coinfection, n = 34 viral, and n = 18 noninfectious illness) and were largely characterized by return visits for persistent symptoms.

In addition to assessing the impact of missed bacterial infections, we examined inappropriate antibacterial prescription rates. In our cohort, 128 subjects adjudicated as having a nonbacterial illness were prescribed antibacterial drugs, resulting in a 33.3% rate of inappropriate prescribing. A hypothetical strategy in which host gene expression was the sole determinant of antibacterial therapy would have decreased inappropriate antibacterial administration to 21.7% of antibacterials prescribed, slightly better than procalcitonin algorithms, which would have reduced it to 22.6%.

DISCUSSION

Antimicrobial resistance poses a global threat that has worsened over time. There are many strategies aimed at combatting this problem. Among them is the more prudent use of antimicrobials so as to reduce the selective pressures driving resistance. ARI is the most frequent reason for infectious disease–related acute care visits and the most common condition in which antibacterials are inappropriately prescribed, especially as diagnostic testing remains unavailable in many outpatient settings. Therefore, easy and rapid diagnostics to guide appropriate antibiotic use are paramount. Specifically, tests to differentiate viral and bacterial infections can play an important role in improving individual patient care as well as public health. Host response–based diagnostics are particularly compelling as the immune response differentiates bacterial and viral disease. New and emerging host response strategies have thus far focused on differentiating bacterial from viral infection as the primary end point [23, 28–31]. However, this is only useful if it translates into clinical action. Procalcitonin has demonstrated mixed results regarding the ability to limit unnecessary antibiotic use, demonstrating that clinical utility is paramount [32, 33]. In this study, we model the potential clinical utility of a host gene expression test that discriminates bacterial from viral infection as compared with current clinical decision-making and procalcitonin. We show that current practice is heavily weighted to overtreat with antibacterials and that host gene expression can mitigate this practice even more so than procalcitonin.

The overall accuracy of the host gene expression–based test provides an indication that the current iteration of the host gene expression diagnostic can perform well in a clinical setting. And the relatively well-balanced sensitivity and specificity of host gene expression provides evidence that this performance is relatively unbiased, especially when compared with clinicians—who have a sizeable sensitivity bias toward finding bacterial infections, to the detriment of specificity in identifying these infections. This sensitivity bias is likely driven by a desire

not to miss bacterial infections. Indeed, we observed that 10 of 15 subjects whose bacterial infections were clinically missed and for whom we had prior visit or follow-up visit data later required hospitalization. Furthermore, we found that this clinician sensitivity bias was exacerbated in immunocompromised subjects. The host response diagnostic remained superior in this group, albeit with lower accuracy than in immunocompetent subjects, consistent with prior findings [34]. However, antibacterial overuse is not the solution, as demonstrated by our current antimicrobial resistance crisis. This underscores the important role diagnostics can play. Tests, such as the host gene expression test described here, can provide clinicians with the diagnostic confidence needed to treat or withhold antibacterials more appropriately.

This clinician bias and the overall strong accuracy of the host gene expression test are likely the 2 main contributors to host gene expression having a significantly higher AWA than the clinician-recommended treatment for both bacterial and viral infections. This improved accuracy over clinician treatment was also reflected by the overall change in Net Benefit (ΔNB), which showed that use of the host gene expression diagnostic would have benefited the subjects in our cohort when using an r value of 0.25 (willing to accept 4 false-positive bacterial diagnoses to avoid 1 false-negative). Interestingly, use of the host gene expression diagnostic would have more than halved the number of patients receiving inappropriate antibacterial therapy in our cohort—a potentially sizeable source of savings in terms of both health care costs and adverse medication events. Considering that true bacterial infections were overrepresented in this cohort (and viral infections underrepresented), this reduction is an underestimate. This study also did not account for antibiotic-associated adverse events due to the retrospective nature of the data collection. Accounting for such adverse events would further support use of diagnostic testing to limit unnecessary antibacterial use.

Unlike the host gene expression diagnostic, procalcitonin did not perform well in this cohort, giving the lowest overall accuracy and bacterial AWA of all 4 diagnostic approaches. Unsurprisingly, host gene expression's improvement over procalcitonin was significant in terms of bacterial AWA and provided a Net Benefit improvement over a wide range of parameters. Although our findings are based purely on modeling, they are consistent with recent observations that a procalcitonin-based algorithm did not reduce antibacterial use in ED patients with lower respiratory tract infections [33]. Recent work also supports that procalcitonin-directed care in the ED does not reduce in-hospital mortality in patients with a low sepsis risk [35]. In contrast, modeling of host response–based diagnostics has estimated that early accurate diagnosis of ARI using a host response assay may result in significant cost savings, shorter lengths of hospital stay, reduced length of antibiotic use, and lower 30-day mortality [36]. Our findings support the notion that host gene expression testing should be used instead of procalcitonin to improve the initial treatment of ARI, which is likely to improve patient outcomes and health care costs, while avoiding unnecessary antibacterial usage.

It is important to note that the distribution of disease etiology seen in this sample is not representative of the actual distribution within the population [15]. Rather, samples were selected to have a roughly equal distribution of etiologies for the purposes of the initial biosignature development. Bacterial infections were overrepresented while viral etiologies were underrepresented in this study compared with actual disease prevalence [15]. As a result of this skewed distribution and the disparate clinical impacts of bacterial, viral, and noninfectious etiologies, standard measures of diagnostic performance (accuracy, sensitivity, specificity, etc.) are inadequate for estimating the potential diagnostic value of different methodologies. Thus, AWA and ΔNB (which correct for these issues) were utilized in our analyses to account for these limitations.

While this study was inherently limited by its retrospective nature, we showed that this host gene expression diagnostic has the potential to significantly improve the clinical diagnosis of acute respiratory illness etiologies, potentially guiding more accurate clinical interventions. We note that our findings reflect an idealized scenario and did not account for the complexities motivating clinicians to prescribe antibacterials (eg, patient requests, patient satisfaction scores, time pressures, etc.). This improved diagnostic accuracy can reduce antibacterial misuse, averting financial and health costs, while also averting missed diagnoses of potentially critical infections. However, the behavior of clinicians in the context of a novel diagnostic cannot be retrospectively assessed, especially in the context of a growing emphasis on antibacterial stewardship. Thus, confirmation of these findings with prospective testing, including the use of stewardship programs and a more heterogeneous patient population, is necessary to demonstrate clinical utility and support the adoption of novel host-based diagnostics.

Acknowledgments

Financial support. None.

Potential conflicts of interest. E.L.T., M.T.M., G.S.G., and C.W.W. have filed for a patent pertaining to the signatures discussed in this study (WO 2017/004390 A1). E.L.T., G.S.G., and C.W.W. hold equity in Biomeme, Inc. The authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Craig JC, Williams GJ, Jones M, et al. The accuracy of clinical symptoms and signs for the diagnosis of serious bacterial infection in young febrile children: prospective cohort study of 15 781 febrile illnesses. BMJ **2010**; 340:c1594.
2. Van den Bruel A, Thompson MJ, Haj-Hassan T, et al. Diagnostic value of laboratory tests in identifying serious infections in febrile children: systematic review. BMJ **2011**; 342:d3082.
3. Espino JU, Wagner MM. Accuracy of ICD-9-coded chief complaints and diagnoses for the detection of acute respiratory illness. Paper presented at: Proceedings AMIA Symposium, Washington, DC, 3–7 November **2001**.

4. Hong CY, Lin RT, Tan ES, et al. Acute respiratory symptoms in adults in general practice. Fam Pract **2004**; 21:317–23.
5. Monto AS. Epidemiology of viral respiratory infections. Am J Med **2002**; 112(Suppl 6A):4S–12S.
6. Aminov RI. The role of antibiotics and antibiotic resistance in nature. Environ Microbiol **2009**; 11:2970–88.
7. Alanis AJ. Resistance to antibiotics: are we in the post-antibiotic era? Arch Med Res **2005**; 36:697–705.
8. Doan Q, Enarson P, Kissoon N, et al. Rapid viral diagnosis for acute febrile respiratory illness in children in the emergency department. Cochrane Database Syst Rev **2014**; 2014:CD006452.
9. Acar J, Kaplan E, O'Brien T. Monitoring and management of bacterial resistance to antimicrobial agents: a World Health Organization symposium. Clin Infect Dis **1997**; 24(Suppl 1):S1–176.
10. Boucher HW, Talbot GH, Bradley JS, et al. Bad bugs, no drugs: no ESKAPE! An update from the Infectious Diseases Society of America. Clin Infect Dis **2009**; 48:1–12.
11. Spellberg B, Guidos R, Gilbert D, et al; Infectious Diseases Society of America. The epidemic of antibiotic-resistant infections: a call to action for the medical community from the Infectious Diseases Society of America. Clin Infect Dis **2008**; 46:155–64.
12. Fleming-Dutra KE, Hersh AL, Shapiro DJ, et al. Prevalence of inappropriate antibiotic prescriptions among US ambulatory care visits, 2010-2011. JAMA **2016**; 315:1864–73.
13. Cooper NJ, Sutton AJ, Abrams KR, et al. Effectiveness of neuraminidase inhibitors in treatment and prevention of influenza A and B: systematic review and meta-analyses of randomised controlled trials. BMJ **2003**; 326:1235.
14. Shapiro DJ, Hicks LA, Pavia AT, Hersh AL. Antibiotic prescribing for adults in ambulatory care in the USA, 2007–09. J Antimicrob Chemother **2013**; 69:234–40.
15. Zaas AK, Garner BH, Tsalik EL, et al. The current epidemiology and clinical decisions surrounding acute respiratory infections. Trends Mol Med **2014**; 20:579–88.
16. Sweeney TE, Wong HR, Khatri P. Robust classification of bacterial and viral infections via integrated host gene expression diagnostics. Sci Transl Med **2016**; 8:346ra91.
17. Herberg JA, Kaforou M, Wright VJ, et al; IRIS Consortium. Diagnostic test accuracy of a 2-transcript host RNA signature for discriminating bacterial vs viral infection in febrile children. JAMA **2016**; 316:835–45.
18. 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.
19. Ramilo O, Allman W, Chung W, et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. Blood **2007**; 109:2066–77.
20. Zaas AK, Burke T, Chen M, et al. A host-based RT-PCR gene expression signature to identify acute respiratory viral infection. Sci Transl Med **2013**; 5:203ra126.
21. Woods CW, McClain MT, Chen M, et al. A host transcriptional signature for presymptomatic detection of infection in humans exposed to influenza H1N1 or H3N2. PLoS One **2013**; 8:e52198.
22. Tsalik EL, Henao R, Nichols M, et al. Host gene expression classifiers diagnose acute respiratory illness etiology. Sci Transl Med **2016**; 8:322ra11.
23. Tsalik EL, Henao R, Montgomery JL, et al; Antibacterial Resistance Leadership Group. Discriminating bacterial and viral infection using a rapid host gene expression test. Crit Care Med **2021**; 49:1651–63.
24. Liu Y, Tsalik EL, Jiang Y, et al. Average weighted accuracy (AWA): pragmatic analysis for a radical study. Clin Infect Dis **2020**; 70:2736–42.
25. Briel M, Christ-Crain M, Young J, et al. Procalcitonin-guided antibiotic use versus a standard approach for acute respiratory tract infections in primary care: study protocol for a randomised controlled trial and baseline characteristics of participating general practitioners [ISRCTN73182671]. BMC Fam Pract **2005**; 6:34.
26. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making **2006**; 26:565–74.
27. Rousson V, Zumbrunn T. Decision curve analysis revisited: overall Net Benefit, relationships to ROC curve analysis, and application to case-control studies. BMC Med Inform Decis Mak **2011**; 11:45.
28. Ramilo O, Mejias A, Mahajan P, Kuppermann N. RNA signature test to distinguish bacterial from viral infection. J Pediatr **2017**; 182:401–4.
29. Mayhew MB, Buturovic L, Luethy R, et al. A generalizable 29-mRNA neural-network classifier for acute bacterial and viral infections. Nat Commun **2020**; 11:1177.
30. 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.
31. 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.
32. Schuetz P, Müller B, Christ-Crain M, et al. Procalcitonin to initiate or discontinue antibiotics in acute respiratory tract infections. Evid Based Child Health **2013**; 8:1297–371.
33. Huang DT, Yealy DM, Filbin MR, et al; ProACT Investigators. Procalcitonin-guided use of antibiotics for lower respiratory tract infection. N Engl J Med **2018**; 379:236–49.
34. Mahle RE, Suchindran S, Henao R, et al. Validation of a host gene expression test for bacterial/viral discrimination in immunocompromised hosts. Clin Infect Dis **2021**; 73:605–13.
35. Covino M, Gallo A, Montalto M, et al. The role of early procalcitonin determination in the emergency department in adults hospitalized with fever. Medicina **2021**; 57:179.
36. Schneider JE, Romanowsky J, Schuetz P, et al. Cost impact model of a novel multi-mRNA host response assay for diagnosis and risk assessment of acute respiratory tract infections and sepsis in the emergency department. J Health Econ Outcomes Res **2020**; 7:24–34.