

Review article

Host-Response Biomarkers for Distinguishing Bacterial and Viral Infections: A Review of Diagnostic Performance and Antibiotic Stewardship Program

Arzoo Nazir ¹, Muhammad Salman ², Muhammad Noman ^{2,*}

¹ College of Veterinary Medicine, Institute of Comparative Medicine, Yangzhou University, 225009 Yangzhou, China

² Department of Medical Laboratory Technology, Faculty of Allied Health Sciences, Superior University, Lahore, Punjab, Pakistan

*Correspondence: muhammadnomanmushtaq787@gmail.com

Article information:

Received: 13-09-2026

Revised: 17-10-2026

Accepted: 17-10-2026

Published: 18-10-2026

Academic Editor:

Dr. Naveed Ahmed

Keywords:

Host response

Biomarkers

Bacterial infections

Viral infections

Transcriptomics

Procalcitonin

Antibiotic stewardship

Abstract: Distinguishing bacterial from viral infection remains difficult when symptoms overlap, pathogen testing is incomplete, or microbial detection does not establish causation. Host-response biomarkers offer complementary information by measuring the patient's inflammatory and immune response rather than identifying the organism alone. This narrative review examines conventional inflammatory markers, multianalyte protein signatures and blood transcriptomic classifiers, with particular attention to acute respiratory illness and their implications for antibiotic stewardship. Evidence from diagnostic studies, randomized trials, economic evaluations and clinical guidelines was considered separately to distinguish biological discrimination from demonstrated clinical benefit. C-reactive protein and procalcitonin can support selected prescribing pathways, but neither provides a universal bacterial-viral classification rule. Multianalyte combinations incorporating TRAIL, IP-10 and C-reactive protein, and transcriptomic assays incorporating interferon-associated and bacterial-response genes, have shown promising discrimination in validation cohorts. However, performance varies with patient selection, reference standards, disease severity, immune status and the handling of indeterminate results. Procalcitonin-guided strategies have reduced antibiotic exposure in some trials, whereas the ProACT trial found no reduction, illustrating the importance of clinical context and implementation. Emerging protein-panel intervention studies remain insufficient to establish benefit across populations. Safe adoption requires timely results, explicit intended use, clinical override mechanisms, follow-up and evaluation of both antibiotic use and missed bacterial infection. Host-response testing should complement clinical assessment and pathogen-based diagnostics within validated stewardship pathways, rather than replace them.

Article citation: Nazir, A., Salman, M. & Noman, M. Host-Response Biomarkers for Distinguishing Bacterial and Viral Infections: A Review of Diagnostic Performance and Antibiotic Stewardship Program. Electron J Med Res. 2026. 2(4): 43-53. <https://doi.org/10.64813/ejmr.2026.143>

Introduction

Acute infectious syndromes commonly present before their cause can be established. Fever, cough, constitutional symptoms and abnormal inflammatory indices overlap between bacterial infection, viral infection and noninfectious disease. The clinical problem is therefore not simply identifying an

organism: it is deciding whether an antibiotic is justified, how urgently treatment is needed and when therapy can be stopped. Pathogen-focused tests remain essential, but negative cultures do not exclude infection, and positive molecular results may reflect carriage, prolonged shedding or a pathogen unrelated to the current illness. Host-response testing addresses

a different question by measuring how the patient is responding to the suspected disease process [1].

These distinctions matter because etiological uncertainty and treatment urgency are not interchangeable. In a stable outpatient, additional evidence may support observation without antibiotics. In a patient with shock or a strongly supported bacterial syndrome, waiting for a biomarker could create harm. The adult community-acquired pneumonia guideline and the 2026 Surviving Sepsis Campaign illustrate this separation: biomarker information should not displace appropriate initial treatment when the clinical indication is established, although selected serial-marker pathways may assist later discontinuation [2, 3]. A diagnostic classifier therefore needs a defined clinical role, not an assumption that any improved discrimination will improve every treatment decision.

The field has expanded from individual inflammatory markers to combinations of host proteins and patterns of gene expression. These approaches aim to exploit partially complementary biological signals, potentially reducing the ambiguity of a single concentration [4, 5]. Nevertheless, discrimination against an adjudicated diagnosis is an intermediate endpoint. The stronger translational question is whether reporting a result change prescribing appropriately while maintaining patient safety. This review examines that evidence chain, emphasizing acute respiratory presentations while drawing on pediatric and critical-care studies where they clarify generalizability or stewardship. Its central argument is that host-response assays should be evaluated as components of care pathways, with diagnostic performance, clinical utility and implementation assessed separately.

Review approach and scope

This article is a focused narrative review rather than a systematic review or meta-analysis. We conducted targeted literature searches on 18 June 2026 using PubMed-indexed records, Europe PMC bibliographic records, and relevant journal and professional-society websites. Searches combined terms including “host response,” “bacterial viral,” “biomarker,” “TRAIL,” “IP-10,” “C-reactive protein,” “procalcitonin,” “gene expression,” “transcriptomic,” “antibiotic stewardship” and named trials or assays. Foundational studies were considered alongside recent validation, intervention, and implementation publications. Bibliographic details were checked against indexed records; quantitative findings were drawn from

accessible abstracts or full text. No exhaustive retrieval, independent duplicate screening, formal risk-of-bias scoring, or pooled effect calculation was performed.

Selection prioritized studies directly informing bacterial–viral discrimination or antibiotic decisions: prospective diagnostic cohorts, external validations, randomized prescribing trials, evidence syntheses and clinical guidelines. Studies of sepsis diagnosis or severity were included only when they clarified the distinction between infection classification and prognostic use. Findings were organized by biological modality and evidence type. Study-specific accuracy estimates were not treated as directly comparable rankings because populations, thresholds, and reference standards differed. The interpretation and proposed implementation framework represent a critical synthesis, not a validated treatment algorithm. This approach permits examination of the field's clinical logic while acknowledging selection bias and the absence of a comprehensive systematic search.

Biological rationale and clinical interpretation

Host-response assays measure consequences of infection, not an intrinsic microbial label. The biological response depends on pathogen characteristics, anatomical site, time since onset, illness severity, and the patient's immune state. A circulating inflammatory protein may be increased in bacterial illness without being specific to bacteria; an interferon-associated signal may support viral infection without excluding simultaneous bacterial disease. The original TRAIL/IP-10/CRP work deliberately combined responses with different associations rather than assuming that one biomarker could capture all infection biology [4]. This provides a plausible basis for multianalyte testing, but biological complementarity must still be demonstrated clinically.

Gene-expression strategies extend the same principle across multiple immune pathways. Instead of measuring only downstream circulating proteins, they quantify patterns of host messenger RNA, usually from blood. Early work developed separate classifiers for bacterial respiratory illness, viral respiratory illness, and noninfectious conditions, acknowledging that the diagnostic space is broader than two mutually exclusive categories [5]. Separate bacterial and viral scores can also accommodate simultaneous signals more naturally than a single forced label. However, the existence of such scores does not itself validate detection of coinfection; sufficient mixed-infection

cases and an appropriate reference standard are needed.

Etiological and severity information must also remain distinct. A strong inflammatory response can accompany severe viral disease, while early or localized bacterial infection may produce a less pronounced systemic response. An assay trained to

identify infection cannot automatically be assumed to predict deterioration, and a prognostic score cannot identify which antibiotic is appropriate. Clinical validation of newer integrated platforms explicitly evaluates infection status and the need for critical-care interventions as separate outcomes [6]. **Figure 1** summarizes these complementary diagnostic roles and the modifiers that can change interpretation.

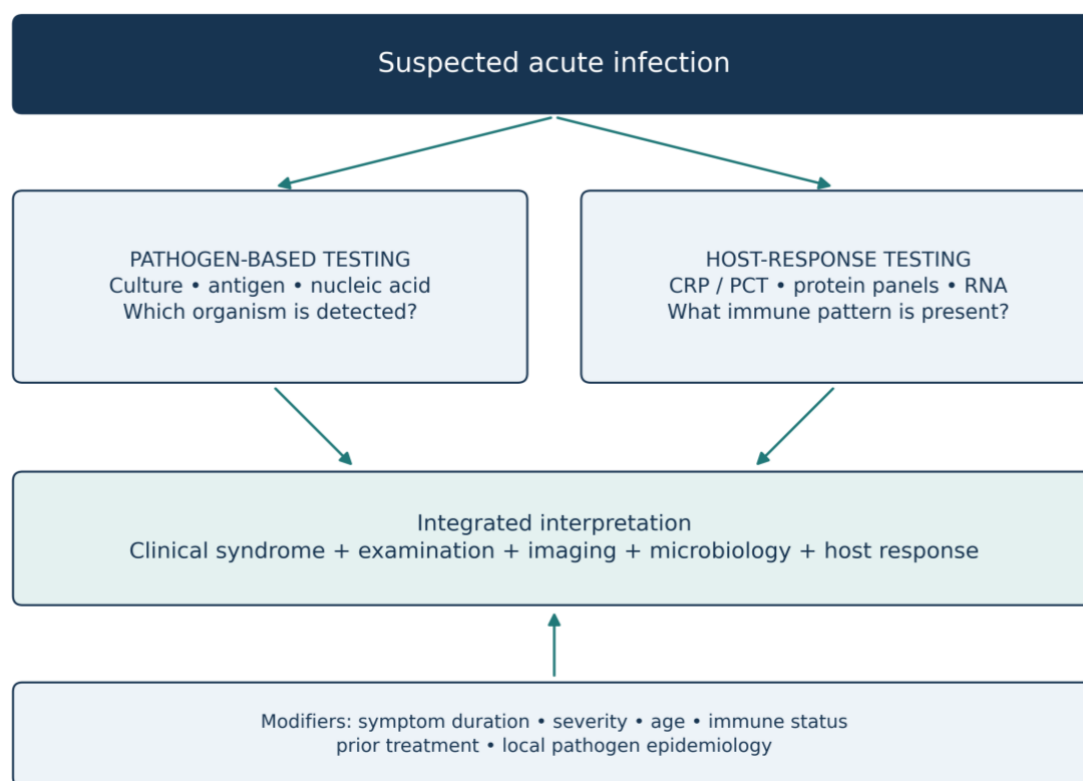


Figure 1. Complementary roles of pathogen detection and host-response biomarkers. The schematic illustrates integration rather than a causal model. Biomarker patterns are modified by the patient and clinical context. CRP, C-reactive protein; PCT, procalcitonin; RNA, ribonucleic acid.

Conventional host-response biomarkers

C-reactive protein

C-reactive protein (CRP) is useful because it offers an accessible measure of inflammation in clinical assessment, rather than because it perfectly separates bacterial and viral etiologies. Its stewardship value has been tested in particular syndromes. In a Vietnamese primary-care trial involving nonsevere respiratory infections, reported antibiotic use within 14 days was 64% with CRP-guided care and 78% with routine care among participants with primary-outcome data. Recovery was not compromised in that study, but the exclusion of severe illness and missing outcome data are relevant to interpretation [7]. These findings should not be generalized to unstable patients.

The PACE trial provides a second example of syndrome-specific utility. Among 653 patients with acute exacerbations of chronic obstructive pulmonary disease in UK primary care, patient-reported antibiotic use was 57.0% with CRP-guided care versus 77.4% with usual care, without evidence of worse COPD-related health status [8]. This is evidence for a prescribing strategy in exacerbations, not proof that a particular CRP concentration establishes a bacterial cause. The important outcome was whether the decision support reduced unnecessary treatment without worsening the measured clinical outcomes.

A Cochrane review found that CRP point-of-care testing likely reduced antibiotic prescribing at the index consultation, with a risk ratio of 0.77 across 12 trials involving 10,218 participants; heterogeneity was substantial [9]. That variability supports interpretation of CRP as a contextual intervention whose

effectiveness depends on population, accompanying guidance and usual prescribing practice. It also cautions against assuming that thresholds developed in one setting should transfer unchanged elsewhere. A practical review should distinguish test availability from the complete intervention, including explanation of results, reassessment and communication with patients.

Procalcitonin

Procalcitonin (PCT) has a substantial intervention evidence base, particularly when used serially to support treatment duration decisions. An individual-patient meta-analysis of 26 respiratory-infection trials involving 6,708 participants found lower antibiotic exposure with PCT guidance: 5.7 versus 8.1 days. Thirty-day mortality was also lower in the pooled analysis, although this should not be converted into a universal survival claim for all syndromes or healthcare systems [10]. The intervention included algorithms and clinical interpretation; the result cannot be attributed to measurement alone.

The ProACT trial illustrates the limits of transferability. Across 14 US hospitals, providing PCT results and interpretation guidance did not significantly reduce antibiotic-days compared with usual care: 4.2 versus 4.3 days in 1,656 patients [11]. Usual-care prescribing and the opportunity for an additional test to alter decisions differ between settings. Therefore, a negative implementation trial does not necessarily invalidate the underlying biomarker, just as a positive trial does not guarantee benefit wherever the assay is introduced. Institutions need to establish whether the proposed test addresses an actual local decision gap. In ADAPT-Sepsis, 2,760 critically ill adults were randomized across 41 UK intensive-care units. Daily PCT-guided care reduced mean antibiotic duration from 10.7 to 9.8 days compared with standard care; daily CRP guidance did not significantly reduce duration. PCT met the prespecified mortality noninferiority criterion, whereas the CRP safety comparison was inconclusive [12]. These findings support a specific discontinuation strategy and do not establish PCT as an adequate reason to withhold initial treatment for sepsis.

Routine blood counts and single-marker interpretation

Routine blood counts and differential indices can contribute to clinical assessment, but convenience should not be confused with etiological specificity. Their most defensible role is as part of the baseline

information against which a new assay is evaluated. If a multianalyte classifier is compared only with one weak marker, its apparent improvement may exaggerate the gain over real clinical care. Head-to-head work comparing PCT, a protein panel and host RNA illustrates the value of evaluating several strategies in the same patients, although the analytical design still limits conclusions about deployed commercial tests [13]. **Table 1** summarizes the principal biomarker approaches and their limitations.

Multianalyte protein signatures

Combining bacterial-associated and viral-associated signals

The TRAIL/IP-10/CRP approach is an important example of integrating biologically different signals. Oved and colleagues screened candidate proteins in prospectively recruited patients and developed a signature with an area under the receiver-operating-characteristic curve (AUROC) of 0.94 for distinguishing bacterial and viral etiologies [4]. This established the potential of a combined response, but discovery and split-sample validation do not demonstrate prescribing benefit. The commercial and inventor relationships disclosed in development studies also make independent replication important, without implying that industry involvement invalidates the findings.

Later clinical validation should be interpreted according to its actual population. The Apollo study enrolled patients from emergency departments and urgent-care centres; 476 adults and children comprised the study population. Sensitivity and specificity for bacterial infection were 90.0% and 92.8%, respectively, with a negative predictive value of 98.8%; 7.2% yielded equivocal scores [14]. These figures indicate promising performance in the evaluated setting. They do not represent a universal probability of safety for every patient with a viral result, particularly when bacterial prevalence, eligibility criteria, or management of equivocal cases differ. Reporting an equivocal interval is clinically meaningful because some immune profiles do not support confident classification. However, intermediate results also reduce the proportion of encounters in which a test provides direct decision support. Studies should therefore report diagnostic yield, exclusions and operational failures alongside sensitivity and specificity. For a health service, an excellent estimate among classifiable patients may be less useful if many real-world patients fall outside validated indications or generate unresolved results.

Table 1. Host-response approaches and their clinical interpretation.

Approach	Information provided	Potential clinical role	Main limitations / evidence
CRP	Acute-phase inflammation	Adjunct to syndrome-specific primary-care prescribing	Not pathogen-specific; trial benefit does not establish bacterial etiology [7, 8, 9].
PCT	Systemic inflammatory response; serial trajectory	Selected antibiotic discontinuation pathways	Not a stand-alone initiation rule; context-dependent intervention results [2, 3, 10, 11, 12].
Routine blood counts	Leukocyte abundance and differential	Baseline clinical information and comparator	Limited etiological specificity; compare new tests with complete clinical care [13].
TRAIL / IP-10 / CRP	Combined immune protein pattern	Bacterial–viral discrimination in validated populations	Equivocal results, reference-standard uncertainty and limited definitive intervention evidence [4, 14, 15].
Host RNA classifiers	Multigene immune expression patterns	Bacterial, viral and sometimes noninfectious classification	Model/platform differences; external validation and immune-status effects [16, 17, 18, 19, 13, 20, 23, 24, 25].
Integrated host–microbe sequencing	Host response plus broad microbial detection	Research approach to difficult sepsis diagnosis	Different workflow and target from rapid outpatient testing; requires further clinical evaluation [26].

CRP, C-reactive protein; PCT, procalcitonin; TRAIL, tumor necrosis factor-related apoptosis-inducing ligand; IP-10, interferon gamma-induced protein 10. Clinical roles are contextual and do not constitute prescribing instructions.

From protein classification to prescribing impact

JUNO supplies early randomized evidence beyond diagnostic accuracy. This pilot enrolled adults with lower respiratory symptoms and consideration of antibiotics; 214 participants were included in the reported analysis. Overall antibiotic prescribing was 30% with standard care and 24% with test-informed care, an absolute difference of –6 percentage points with a confidence interval spanning –18 to 6. Prescribing shifted toward bacterial scores and away from viral scores, but the pilot was not definitive evidence of an overall reduction or broad safety [15]. Its main contribution is feasibility and an estimate for designing larger trials.

The appropriate next question is whether test-guided treatment improves the match between bacterial need and antibiotic exposure. A strategy that decreases prescribing in low-likelihood cases while increasing appropriate treatment in high-likelihood cases might leave the total prescription rate

unchanged. Conversely, a lower total rate could conceal undertreatment. Future protein-panel trials should therefore prespecify appropriateness measures, patient outcomes and the handling of intermediate results. Their reports should avoid treating hypothetical avoided prescriptions, based on blinded results, as equivalent to observed treatment changes.

Transcriptomic classifiers

Small gene signatures and early proof of concept

A two-transcript signature involving FAM89A and IFI44L demonstrated that a compact host RNA measurement could distinguish bacterial and viral infection in febrile children. In the definite-diagnosis validation subset, all 23 bacterial cases and 27 of 28 viral cases were classified correctly [16]. These are encouraging proof-of-concept results, but the denominator is crucial: the validation cohort also

included many indeterminate cases. Small, microbiologically clear subsets may not represent the uncertainty encountered in routine practice, and the absence of observed false negatives does not mean the true false-negative risk is zero.

Compact signatures offer potential advantages in assay simplicity, but parsimony creates trade-offs. A small number of genes may be easier to measure and standardize while capturing less biological variation. The key question is not whether a signature is the smallest, but whether it preserves discrimination across age, severity, pathogens and noninfectious mimics. Development datasets should therefore be separated from external validation, and thresholds should be fixed before evaluation. Clinical implementation also requires demonstration that the laboratory platform measures the relevant signal reproducibly in fresh patient samples.

Multigene classifiers and external validation

Mayhew and colleagues developed a 29-mRNA classifier using 1,069 samples from 18 retrospective transcriptomic studies. In an independent cohort of 163 patients, bacterial-versus-other and viral-versus-other AUROCs were 0.86 and 0.85, respectively, lower than the development estimates [17]. The importance of this work is not a single headline accuracy figure, but evaluation without retraining in an independent cohort. Differences between development and validation also demonstrate why pooled retrospective datasets cannot substitute for prospective testing in the intended clinical workflow.

Bauer and colleagues subsequently assessed a related 29-mRNA model in 312 adults with suspected acute infection or sepsis. The bacterial score AUROC was 0.90 and the viral score AUROC was 0.83 [18]. Interpretation must account for adjudication: PCT was available in the reference assessment, making comparisons with that biomarker potentially dependent on how diagnoses were assigned. In a separate multisite study of 602 adults with suspected influenza, the 29-mRNA approach yielded AUROCs of 0.81 for bacterial and 0.87 for viral illness [19]. Together, these studies support feasibility while showing that performance is population-dependent.

A direct comparison in 286 acute respiratory illness presentations reported greater discrimination for an mRNA panel than PCT or a three-protein panel [13]. However, the protein and RNA approaches were evaluated with statistical modeling and cross-validation rather than as a simple contest between independently locked, commercially implemented

assays. It would therefore be inappropriate to conclude that every RNA platform outperforms every protein test. Useful comparative studies need matched specimens, equivalent blinding, prespecified thresholds and the same reference standard, with no overlap between training and test participants.

A systematic comparison of 28 published signatures across 4,589 subjects from 51 datasets found wide variation in performance, with differences related to signature size and validation-population characteristics [20]. This argues against treating “host transcriptomics” as one homogeneous diagnostic technology. Different signatures, platforms and analytical versions represent different tests. Reuse of public datasets also requires careful accounting to prevent the same patients contributing to both development and apparently independent validation. Transparent dataset provenance is therefore part of clinical credibility, not merely a computational detail.

Translation to rapid instruments

Recent studies have moved from laboratory expression measurements toward sample-to-answer instruments. SEPSIS-SHIELD evaluated a device generating bacterial, viral and severity information in 1,222 patients, advancing the evidence beyond discovery signatures [6]. Analytical evaluation of the corresponding rapid RT-LAMP system examined reproducibility and other technical performance characteristics [21]. These steps are important because model accuracy on extracted RNA does not establish reliability after changes in specimen handling, amplification chemistry, instrument calibration or assay manufacturing. Analytical validation and clinical validation answer complementary questions.

A 2026 Acutelines report adds real-world evaluation of the same general platform [22]. Such studies need to be read alongside the exact model version, thresholds, and recruitment conditions rather than merged indiscriminately with earlier 29-mRNA research. A device combining etiological and prognostic outputs may influence several decisions, but this also complicates evaluation of which output produced benefit or harm. Prospective studies should specify whether results are visible, whether clinicians must act on them, and how diagnostic information changes care beyond existing assessment.

Special populations and diagnostic uncertainty

Immunocompromised patients are a particularly important test of generalizability. In a validation study

of 134 immunocompromised participants, an 81-gene test showed lower bacterial-classification accuracy than in its immunocompetent training cohort [23]. Immune-modifying treatment and disease can alter the very signals the assay measures. A result that appears reassuring in a general emergency population may therefore require different interpretation in transplantation, chemotherapy or advanced immune dysfunction. Tests should be used within validated populations, with uncertainty clearly displayed rather than concealed behind a universal score.

Geographical validation is equally important. A Sri Lankan study applied previously developed classifiers to 79 patients with febrile respiratory presentations, including influenza, dengue, leptospirosis and scrub typhus, and reported high classification accuracy [24]. This demonstrates the value of testing beyond the original geographical setting, but the small selected cohort cannot establish transportability across all endemic diseases. A separate community study of the Fever-7 signature reported an AUROC of 0.81 for a positive respiratory viral test [25]. That endpoint is informative but differs from ruling out bacterial infection; a virus-positive patient can still need antibiotics.

Coinfection, prior treatment and uncertain diagnoses should be retained as explicit clinical categories. Excluding these encounters can produce attractive accuracy estimates while removing the cases in which clinicians most need help. Integrated host-microbe sequencing research suggests that combining immune-response information with broad pathogen detection can help characterize difficult sepsis presentations [26]. Nevertheless, the technique's performance, workflow and target question differ from rapid outpatient bacterial-viral discrimination. Its conceptual lesson is complementarity: host signals can contextualize microbial findings, while organism detection can clarify an immune response that is otherwise nonspecific.

Evaluating diagnostic performance responsibly

The reference standard is a major limitation because many acute infections lack definitive microbiological confirmation. Expert adjudication can incorporate follow-up and multiple tests, but it remains vulnerable to misclassification. If the adjudicators see the index assay, incorporation bias can inflate apparent accuracy; if they see a comparator such as PCT, head-to-head interpretation can also be distorted. Independent adjudication, explicit diagnostic

definitions and sensitivity analyses using stricter or broader certainty categories help expose this problem. STARD provides a reporting framework, while QUADAS-2 identifies bias and applicability concerns in participant selection, index testing, reference assessment and flow [27, 28].

Spectrum and prevalence require separate consideration. Sensitivity and specificity can vary with the severity and mix of illness, while predictive values additionally depend on the frequency of bacterial disease in the tested population. As an illustrative calculation, not a published assay estimate, a test with 90% sensitivity and 90% specificity has a negative predictive value of approximately 98.8% when bacterial prevalence is 10%, but 90.0% when prevalence is 50%. The same negative result therefore supports different residual risks. Pretest assessment and intended clinical consequences must accompany accuracy statistics.

AUROC summarizes ranking across possible thresholds but does not identify the best prescribing threshold. Clinical adoption also requires calibration, likelihood ratios, uncertainty intervals and performance in the range where treatment decisions actually change. Intermediate outputs, invalid cartridges and missing samples should be counted transparently. Investigators should report both accuracy among evaluable results and the proportion of all eligible patients obtaining an actionable result. **Table 2** distinguishes representative diagnostic and intervention studies; their estimates should not be pooled or compared as if their populations and outcomes were interchangeable.

From statistical accuracy to clinical usefulness

Threshold selection should begin with the consequences of a wrong decision. A threshold intended to support withholding antibiotics needs a different balance of sensitivity and specificity from one intended to increase confidence that treatment is warranted [26]. The acceptable residual risk also depends on follow-up access, comorbidity, and the seriousness of the syndrome. It is therefore misleading to select a threshold only because it maximizes a statistical index and then present it as an optimal treatment rule. Trial protocols should explain which clinical action is expected at each score range and why that trade-off is acceptable [26]. The incremental contribution of a test should be evaluated against information already available when the result arrives. A biomarker may appear valuable against a sparse baseline model but add little after examination, imaging and rapid pathogen tests. Conversely, it may

be useful when conventional testing is delayed or does not address the suspected pathogen. Analyses should distinguish improvement in discrimination from changes in calibrated risk and from the proportion of

patients whose management would actually differ. Decision-curve analysis can explore assumptions about treatment thresholds, but it cannot replace observed clinical outcomes in a test-guided trial.

Table 2. Selected diagnostic and stewardship studies and limits of interpretation.

Study / design	Population	Selected finding	Interpretation
[16] Diagnostic validation	Febrile children; definite validation subset: 23 bacterial, 28 viral	23/23 bacterial and 27/28 viral cases correctly classified by two-transcript signature	Small definite subset; many indeterminate cases. Not an implementation trial.
[17] External validation	Independent cohort, n = 163	Bacterial AUROC 0.86; viral AUROC 0.85	Locked classifier evaluation; retrospective development and platform translation remain relevant.
[18] Prospective observational	Adults with suspected infection, n = 312	Bacterial AUROC 0.90; viral AUROC 0.83	PCT contributed to adjudication; consider reference-standard dependence.
[14] Prospective blinded accuracy	Adults and children; study population n = 476	Sensitivity 90.0%; specificity 92.8%; NPV 98.8%; equivocal 7.2%	Diagnostic accuracy, not demonstrated prescribing benefit; NPV is prevalence dependent.
[15] Pilot randomized trial	Adult ED/urgent-care LRTI; analyzed n = 214	Prescribing 30% standard care vs 24% test-informed; difference -6 points (95% CI -18 to 6)	Overall difference uncertain; pilot cannot establish rare-event safety.
[10] Individual-patient meta-analysis	26 respiratory-infection trials; n = 6,708	Antibiotic exposure 5.7 vs 8.1 days with PCT guidance vs control	Evidence for algorithms across studied settings; not universal initiation guidance.
[11] Randomized trial	US suspected LRTI; n = 1,656	Antibiotic-days 4.2 vs 4.3; no significant reduction	Shows context-dependent incremental utility.
[12] Randomized trial	41 UK ICUs; n = 2,760	PCT: 9.8 vs 10.7 antibiotic-days; CRP did not significantly reduce duration	PCT met prespecified mortality noninferiority; CRP safety comparison inconclusive.
[6] Device clinical validation	Suspected acute infection; n = 1,222	Bacterial AUROC 0.83; viral AUROC 0.91; severity AUROC 0.78	Diagnostic and prognostic endpoints differ; projected prescribing changes are not observed trial effects.

AUROC, area under the receiver-operating-characteristic curve; CI, confidence interval; ED, emergency department; ICU, intensive-care unit; LRTI, lower respiratory tract infection; NPV, negative predictive value. Estimates are study-specific, not pooled or adjusted for cross-study differences. “Analyzed” and “study population” denominators may differ from total enrollment.

Antibiotic stewardship applications

Initiation and discontinuation are different decisions

The safest application begins with the decision the assay is intended to support. In stable patients whose antibiotic need is genuinely uncertain, a validated low bacterial-likelihood result may add evidence to an observation strategy. It should not override a compelling bacterial focus, physiological instability or a population outside the assay's validation. The 2019 adult pneumonia guideline recommends initial

antibacterial treatment in clinically suspected, radiographically confirmed community-acquired pneumonia regardless of initial PCT [2]. This is a defined syndrome recommendation, not a rejection of all biomarker-supported stewardship.

For sepsis, the 2026 Surviving Sepsis Campaign favors clinical evaluation over adding PCT to decide when to start antimicrobials. It conditionally supports PCT together with clinical evaluation when deciding discontinuation after adequate source control and when optimal duration remains unclear [3]. The asymmetry is logical: a low early marker cannot safely

erase a high-consequence diagnosis, whereas serial information in an improving patient may help resolve whether treatment remains necessary. No host-response test should delay urgent care while clinicians wait for etiological certainty.

Making results actionable without creating automation bias

A proposed pathway should specify who orders the test, who receives it, what action it supports and when reassessment occurs. It places severity assessment before testing, preserves organism-based investigation when indicated and treats discordant or intermediate results as unresolved. Clinical override should remain possible, but its reason should be documented. Otherwise, a pathway can alternate between excessive confidence in the test and routine disregard of it, making both clinical effectiveness and audit difficult to interpret.

Communication is part of the intervention. Patients need an explanation of why antibiotics are or are not advised, what symptoms require reassessment and how results relate to uncertainty. Clinicians need practical training that distinguishes “viral response detected” from “bacterial infection impossible.” Local stewardship teams should review prescribing patterns across high, low and intermediate scores, not just total use. A recent process evaluation linked to ADAPT-Sepsis identified the importance of organizational support, microbiology involvement and communication of evidence [29]. These observations suggest that successful implementation is a service-design task as well as a laboratory purchase.

Implementation priorities and future research

Cost-effectiveness cannot be inferred from fewer prescriptions alone. Services must include instrument and reagent costs, quality control, staffing, repeat tests, laboratory availability and downstream investigations. The ADAPT-Sepsis economic evaluation showed that conclusions depended strongly on assumptions about mortality benefit; within-trial and meta-analysis-based estimates did not produce the same economic interpretation [30]. This is particularly important when antibiotic courses are inexpensive compared with assay costs. Evaluation should measure clinically relevant benefits and harms and report uncertainty rather than assume that reduced exposure automatically generates savings.

Future research should prioritize adequately powered pragmatic trials with consecutive recruitment and clearly defined intended use. Outcomes should include antibiotic initiation, days of therapy, spectrum, appropriate treatment of bacterial disease, return visits, hospitalization and adverse events. Rare harms require sufficient follow-up and sample size; failure to observe a safety signal in a pilot is not equivalent to demonstrating safety. Trials should distinguish the assay from co-interventions such as prescribing education or audit. They should also assess whether benefits persist after the initial implementation period and whether repeat testing adds useful information.

For Saudi Arabia and comparable health systems, a reasonable research programme would begin with local validation across primary care, emergency and inpatient settings, followed by an implementation trial where genuine diagnostic uncertainty remains. Recruitment should represent age, immune status, comorbidity, symptom duration and prior antibiotic exposure. Arabic-language communication materials, access outside major centres and reliable follow-up should be evaluated as components of the pathway. These are proposed design priorities, not findings about current Saudi test performance. A locally useful study should identify a clinical decision gap before selecting a technology to address it.

Finally, independent external evaluation and transparent reporting should accompany commercial development. Assay version, thresholds, exclusions, funding and investigator relationships should be explicit. New biomarkers and machine-learning models should demonstrate improvement over a credible clinical comparator, not simply statistical separation between unusually clear cases. The most valuable advance may be a modestly more informative result that arrives in time and changes care safely, rather than the highest AUROC obtained from stored specimens. Research agendas should therefore connect analytical reliability, diagnostic validity, clinician behavior and patient outcomes within the same evidence program.

Equity and surveillance after implementation

An implementation program should avoid restricting useful diagnostics to patients able to pay or to facilities with the greatest existing resources. Economic evaluation can compare alternative placement strategies, such as selected emergency-department use versus centralized laboratory testing, without assuming that more testing is always beneficial. The relevant denominator is the group for whom a result

can change management safely. Testing patients whose treatment would be unchanged regardless of the result adds cost and may generate confusion rather than clinical value.

Post-implementation surveillance should monitor valid-result rates, actual time to clinician review, prescribing responses, and unexpected deterioration after a reassuring score. Seasonal shifts in circulating pathogens or changes in recruitment can alter the population in which a model operates. A service should therefore review performance periodically and specify how software or assay updates are evaluated. Version changes should be documented so that local outcomes are not attributed to an unspecified evolving test. These are proposed governance measures, rather than evidence that any particular platform has failed after deployment.

Conclusion

Host-response biomarkers offer a complementary approach to bacterial–viral discrimination, particularly when clinical presentations overlap, and pathogen testing leaves uncertainty. CRP and PCT have evidence for selected stewardship pathways, but their usefulness is syndrome- and context-dependent. Multianalyte protein signatures and transcriptomic classifiers broaden the biological information available, with promising validation studies and emerging rapid platforms. Their clinical value cannot be established by diagnostic accuracy alone. Safe translation requires appropriate patient selection, interpretable results, timely reporting, clinical reassessment, and prospective evidence that treatment improves without increasing missed bacterial disease. Integration with clinical judgment, microbiology, and stewardship infrastructure is therefore more defensible than replacement of those systems.

Author Contributions: Conceptualization, M.N.; methodology, A.N. and M.S.; validation, M.N.; resources, A.N.; writing—original draft preparation, A.N. and M.S.; writing—review and editing, M.N.; supervision, M.N.; project administration, A.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: This review article is based on information obtained from previously published literature and publicly available sources. No new experimental or primary datasets were generated or analyzed in this study. All relevant information and sources are appropriately cited within the manuscript and reference list. The authors are responsible for the accuracy and integrity of the information presented.

Acknowledgments: The authors gratefully acknowledge the authors and researchers whose published scientific contributions provided the foundation for this review. The authors also acknowledge the scientific databases and publicly accessible literature resources used during the preparation of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Woodhouse, E.W., McClain, M.T., and Woods, C.W. Harnessing the host response for precision infectious disease diagnosis. *Clinical Microbiology Reviews*. **2024**. 37(4): e00078–24.
2. Metlay, J.P., Waterer, G.W., et al. Diagnosis and treatment of adults with community-acquired pneumonia. An official clinical practice guideline of the American Thoracic Society and Infectious Diseases Society of America. *American journal of respiratory and critical care medicine*. **2019**. 200(7): e45–e67.
3. Prescott, H.C., Antonelli, M., et al. Surviving Sepsis Campaign: international guidelines for management of sepsis and septic shock 2026. *Intensive care medicine*. **2026**. 52(5): 863–936.
4. Oved, K., Cohen, A., et al. A novel host-proteome signature for distinguishing between acute bacterial and viral infections. *PloS one*. **2015**. 10(3): e0120012.
5. Tsalik, E.L., Henao, R., et al. Host gene expression classifiers diagnose acute respiratory illness etiology. *Science translational medicine*. **2016**. 8(322): 322ra11–322ra11.
6. Liesenfeld, O., Arora, S., et al. Clinical validation of an AI-based blood testing device for diagnosis and prognosis of acute infection and sepsis. *Nature medicine*. **2025**. 31(12): 4044–4054.
7. Do, N.T., Ta, N.T., 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–e641.
8. Butler, C.C., Gillespie, D., et al. C-reactive protein testing to guide antibiotic prescribing for COPD exacerbations. *New England Journal of Medicine*. **2019**. 381(2): 111–120.
9. Smedemark, S.A., Aabenhus, R., et al. Biomarkers as point-of-care tests to guide prescription of antibiotics in people with acute respiratory infections in primary care. *Cochrane Database of Systematic Reviews*. **2022**(10).
10. Schuetz, P., Wirz, Y., et al. Effect of procalcitonin-guided antibiotic treatment on mortality in acute respiratory infections: a patient level meta-analysis.

- The Lancet infectious diseases. **2018**. 18(1): 95–107.
11. Huang, D.T., Yealy, D.M., et al. Procalcitonin-guided use of antibiotics for lower respiratory tract infection. *New England Journal of Medicine*. **2018**. 379(3): 236–249.
12. Dark, P., Hossain, A., et al. Biomarker-guided antibiotic duration for hospitalized patients with suspected sepsis: the ADAPT-sepsis randomized clinical trial. *Jama*. **2025**. 333(8): 682–693.
13. Ross, M., Henao, R., et al. A comparison of host response strategies to distinguish bacterial and viral infection. *PLoS One*. **2021**. 16(12): e0261385.
14. Bachur, R.G., Kaplan, S.L., et al. A rapid host–protein test for differentiating bacterial from viral infection: Apollo diagnostic accuracy study. *Journal of the American College of Emergency Physicians Open*. **2024**. 5(3): e13167.
15. Singer, A.J., Hollander, J.E., et al. Effect of host-protein test (TRAIL/IP-10/CRP) on antibiotic prescription and emergency department or urgent care center return visits: The JUNO pilot randomized controlled trial. *Academic Emergency Medicine*. **2025**. 32(9): 975–984.
16. Herberg, J.A., Kaforou, M., 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): 835–845.
17. Mayhew, M.B., Buturovic, L., et al. A generalizable 29-mRNA neural-network classifier for acute bacterial and viral infections. *Nature communications*. **2020**. 11(1): 1177.
18. Bauer, W., Kappert, K., et al. A novel 29-messenger RNA host-response assay from whole blood accurately identifies bacterial and viral infections in patients presenting to the emergency department with suspected infections: a prospective observational study. *Critical care medicine*. **2021**. 49(10): 1664.
19. Shojaei, M., Chen, U.I., et al. Multisite validation of a host response signature for predicting likelihood of bacterial and viral infections in patients with suspected influenza. *European Journal of Clinical Investigation*. **2023**. 53(5): e13957.
20. Bodkin, N., Ross, M., et al. Systematic comparison of published host gene expression signatures for bacterial/viral discrimination. *Genome medicine*. **2022**. 14(1): 18.
21. Figueiredo-Pereira, C., Fleming, P., et al. Analytical evaluation of TriVerity, a rapid diagnostic and prognostic host gene expression test performed on the Myrna instrument using RT-LAMP. *Journal of clinical microbiology*. **2025**. 63(10): e00352–25.
22. Ter Voert, M.A., Lu, J., et al. Real-world clinical evaluation of the Inflammatrix TriVerity test for sepsis at the emergency department: an Acutelines study. *Internal and Emergency Medicine*. **2026**: 1–12.
23. Mahle, R.E., Suchindran, S., et al. Validation of a host gene expression test for bacterial/viral discrimination in immunocompromised hosts. *Clinical Infectious Diseases*. **2021**. 73(4): 605–613.
24. Tillekeratne, L.G., Suchindran, S., et al. Previously derived host gene expression classifiers identify bacterial and viral etiologies of acute febrile respiratory illness in a South Asian population. in *Open forum infectious diseases*. 2020. Oxford University Press US.
25. Almansa, R., Herrero-Rodríguez, C., et al. A host transcriptomic signature for identification of respiratory viral infections in the community. *European Journal of Clinical Investigation*. **2021**. 51(12): e13626.
26. Kalantar, K.L., Neyton, L., et al. Integrated host-microbe plasma metagenomics for sepsis diagnosis in a prospective cohort of critically ill adults. *Nature microbiology*. **2022**. 7(11): 1805–1816.
27. Bossuyt, P.M., Reitsma, J.B., et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *Radiology*. **2015**. 277(3): 826–832.
28. Whiting, P.F., Rutjes, A.W., et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine*. **2011**. 155(8): 529–536.
29. Poole, R.L., Lloydwin, C., et al. Implementation of procalcitonin-guided antibiotic stopping advice in sepsis: A mixed-methods process evaluation in a clinical trial context. *Journal of the Intensive Care Society*. **2026**: 17511437261458968.
30. Stevenson, M., Forsyth, J.E., et al. Cost-effectiveness of procalcitonin-guided antibiotic duration for hospitalized patients with sepsis. *Critical Care*. **2025**. 29(1): 508.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of RES Publishers and/or the editor(s). RES Publishers and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.