APS European Society for Paediatric Research SPR Society for Pediatric Research

REVIEW ARTICLE

Check for updates

Transcript host-RNA signatures to discriminate bacterial and viral infections in febrile children

Danilo Buonsenso[1,2,3,5]✉, Giorgio Sodero[4] and Piero Valentini[2,3,4]

Traditional laboratory markers, such as white blood cell count, C-reactive protein, and procalcitonin, failed to discriminate viral and bacterial infections in children. The lack of an accurate diagnostic test has a negative impact on child's care, limiting the ability of early diagnosis and appropriate management of children. This, on the one hand, may lead to delayed recognition of sepsis and severe bacterial infections, which still represent the leading causes of child morbidity and mortality. On the other hand, this may lead to overuse of empiric antibiotic therapies, particularly for specific subgroups of patients, such as infants younger than 90 days of life or neutropenic patients. This approach has an adverse effect on costs, antibiotic resistance, and pediatric microbiota.
Transcript host-RNA signatures are a new tool used to differentiate viral from bacterial infections by analyzing the transcriptional biosignatures of RNA in host leukocytes. In this systematic review, we evaluate the efficacy and the possible application of this new diagnostic method in febrile children, along with challenges in its implementation. Our review support the growing evidence that the application of these new tools can improve the characterization of the spectrum of bacterial and viral infections and optimize the use of antibiotics in children.

Pediatric Research (2022) 91:454–463; https://doi.org/10.1038/s41390-021-01890-z

IMPACT:

- Transcript host RNA signatures may allow to better characterize the spectrum of viral, bacterial, and inflammatory illnesses in febrile children and can be used with traditional diagnostic methods to determine if and when to start antibiotic therapy.
- This is the first review on the use of transcript RNA signatures in febrile children to distinguish viral from bacterial infections.
- Our review identified a wide variability of target populations and gold standards used to define sepsis and SBIs, limiting the generalization of our findings.

INTRODUCTION

Sepsis remains a leading cause of morbidity and mortality in neonatal and pediatric age,[1] in both low-to-middle countries (LMICs) and well-resourced setting.[2] The early diagnosis of sepsis and serious bacterial infections (SBIs) is pivotal to establish early treatments and reduce mortality and morbidity. However, clinical examination is often non-specific and laboratory tests usually used have little sensitivity and specificity; to date, there is no single laboratory test that can accurately identify sepsis and SBIs during the early stages of disease.[3]

Blood culture remains the gold standard for diagnosis of sepsis and SBIs, although false-negative cultures can occur as a result of small volumes of blood drawn (particularly in neonates, infants, and young children), low levels of bacteremia, use of previous antibiotics, or the presence of bacteria that require specific cultures or have high replication time.[4] Moreover, blood culture results require hours to days according to settings, and therefore they cannot guide initial diagnostic and management decisions. The traditional laboratory tests used in differential diagnostics (such as white blood cell count, C-reactive protein, and procalcitonin) failed to distinguish the etiology of fever in children and also have numerous limitations. For example, C-reactive protein (CRP) is a non-specific marker and can increase in all inflammatory situations (including non-infectious ones), while procalcitonin (PCT) showed better accuracy to recognized sepsis and SBIs than other tests, but has higher costs (limiting its widespread use in LMICs) and can also be negative in the early stages of disease.[5]

Since invasive bacterial infections must be treated immediately, the absence of accurate diagnostic tools leads to high use of empirical antibiotic therapies. Although fever in infants less than 90 days old is more often of viral etiology, it is associated with a severe bacterial infection in 8–12.5% of cases.[6] For this reason, antibiotic therapy is usually empirically started pending cultures or clinical evolution and is often continued beyond 48 h (despite negative cultures) when clinical signs and laboratory markers are consistent with infection. This has an adverse effect on costs, antibiotic resistance, and microbiota.

In recent years, new methods have been studied to better and earlier recognize sepsis and SBIs, including RNA biosignatures. Transcript host RNA signatures are a new tool increasingly used to

[1]Department of Woman and Child Health and Public Health, Fondazione Policlinico Universitario A. Gemelli IRCCS, 00168 Rome, Italy. [2]Global Health Research Institute, Istituto di Igiene, Università Cattolica del Sacro Cuore, 00168 Rome, Italy. [3]Dipartimento di Scienze Biotecnologiche di Base, Cliniche Intensivologiche e Perioperatorie, Università Cattolica del Sacro Cuore, 00168 Rome, Italy. [4]Istituto di Pediatria, Università Cattolica del Sacro Cuore, 00168 Rome, Italy. [5]Present address: Danilo Buonsenso, Largo A. Gemelli 8, 00168 Rome, Italy. ✉email: danilobuonsenso@gmail.com

Received: 23 April 2021 Revised: 8 November 2021 Accepted: 30 November 2021
Published online: 15 December 2021

differentiate viral from bacterial infections by analyzing the transcriptional biosignatures of RNA in host leukocytes, in response to a clinically undifferentiated infection.[6]

Studies on the adult population have shown encouraging results regarding the use of these methods in the diagnosis of infectious diseases. Suarez et al.[7] conducted a prospective work in patients aged ≥21 years hospitalized for low respiratory tract infection and identified 10 classifier genes (8/10 related with Interferon pathway) that discriminated between bacterial and viral infection with a 95% sensitivity and 92% specificity, compared with a sensitivity of 38% and a specificity of 91% for procalcitonin.

However, the use RNA biosignatures to recognize children with sepsis and SBIs is not yet completely established; therefore, we performed a systematic literature review to better understand the possible application of transcript host RNA signatures to discriminate bacterial and viral infections in febrile children.

METHODS

Search strategy

We performed a systematic literature review to evaluate the efficacy of transcript host RNA signatures for the discrimination of bacterial and viral infections in febrile children.

Five different searches were conducted on Pubmed using the following keywords: "Transcriptomic & sepsis & children", "RNA signatures & sepsis & children", "transcriptome & infection & children", "RNA signatures & infection & children" "gene expression signature & sepsis & children". All studies published between January 2011 and February 2021 were considered ($n = 1073$). In total, 188 duplicates were removed after comparison of the five searches.

We included all studies of quantitative research, while letters to the editor ($n = 2$), protocols ($n = 4$), reviews ($n = 72$) and off-topic publications ($n = 777$) were excluded. No articles were excluded due to the language. We additionally analyzed reference lists of the included articles to identify further relevant studies ($n = 3$).

Study selection

A first screening process was performed to identify off-topic studies: 777 articles were excluded from the database. The remaining list was screened to identify quantitative research about the efficacy of transcript host RNA signatures for diagnosis of bacterial and viral infections in febrile children. After an initial preselection of 33 articles, 16 publications were fully considered to be included in this review.

Data extraction

The data extraction template was filled by the author coordinated by the two leading researchers, based on a sample of three studies previously screened and assessed by them, in order to adjust the template to make it perfectly suitable for the study's aim. The results were then checked again across the original manuscript by a third researcher.

Data synthesis

Characteristics of the included studies were presented in tabulated form. The data were collected in columns: study reference, analyzed population and mean age, aim, main result and study limitations. We also presented a supplementary table which reported: type of study, country, setting and conflicts of interest (when present and/or reported).

The data were then summarized using a narrative synthesis approach.

RESULTS

Study selection and description

A total of 1076 articles were identified according to the specified selection criteria on Pubmed. After duplicates were removed ($n = 188$), a total of 888 citations were identified from searches of electronic databases and review article references. Based on the title and the abstract, 777 were excluded, with 108 full text articles to be retrieved and assessed for eligibility. We were able to retrieve all the selected articles. Of these, 78 were not considered to be original quantitative research (72 review articles, 2 letters to the editor, 4 protocol studies).

We preselected 33 articles and, after discarding 17 works (because they did not have as their main purpose the distinction between infections), finally, the remaining 16 studies were considered eligible for this review (Fig. 1).

No meta-analysis was done because the published childhood transcript host biomarker studies were highly heterogeneous. The method used to confirm bacterial and viral infections was also heterogeneous, ranging from positive or negative cultures to different ranges of inflammation indices, up to clinical evaluation alone.

Characteristics of sources of evidence

The modules are described in Table 1, together with a description of aim, target group, main results, and study limitations.

A critical appraisal within sources of evidence, including the studies' place of origin, design, setting, and conflicts of interest are presented in the Supplementary Table 1.

Critical appraisal within sources of evidence

Results of sources of evidence. We included 16 studies,[8–23] with a total of 4841 analyzed patients. Of the 16 selected studies, 14 (87,5%) examined the discrimination of bacterial and/or viral infections using transcript host RNA biosignatures, with 2 of these conducted on a population of oncological children.[12,16] Of the remaining studies, one concerned the identification of a marker of severity in hospitalized children with pneumonia[19] and the other studied markers that allowed to distinguish real asthma exacerbations from episodes that resolved spontaneously in the absence of treatment.[20] In both cases, differentiation of the underlying infectious agents was considered a secondary endpoint.

One study[23] analyzed also the correlation between sepsis severity and gene expression, but did not find statistically significant correlation. Most of the included were prospective studies.

Six studies were conducted in a hospital ward, eight in an emergency department, and two using data from the GEO (Gene Expression Omnibus) dataset.

The majority of publications analyzed children from the United States and China. Other countries with less studies included: Mexico, Spain, Sweden, Mozambique, and Brazil. One study analyzed children from different countries (United Kingdom, Spain, the Netherlands, and the United States). About the two studies that used the GEO dataset, one was conducted in the USA analyzing only data from Mexican children.

In 10 studies the authors declared that they have no conflict of interest; in two studies there is no reference about conflict of interest, while in only four studies possible confounding factors are reported: after our careful analysis, both were deemed suitable for our review.

The majority were conducted only among children ($n = 14$; 87.5%); 10 studies included only children under 10 years old and 4 studies included children under 18 years. Only two studies analyzed children and adults.[8,14]

The 16 studies reported diverse and heterogeneous forms of biomarkers as a gold standard for infection diagnosis, which require different types of clinical specimens and utilized diverse techniques and laboratory assays for identifying biomarkers. IFI44L was one of the most used markers for the diagnosis of viral infections[9–11,15] with also genes linked to IFN pathway.[13,14,16,18,20,22,23]

The vast majority of studies[15,16] report that transcriptomics can be used as a diagnostic supplement for pediatric diseases and allow to

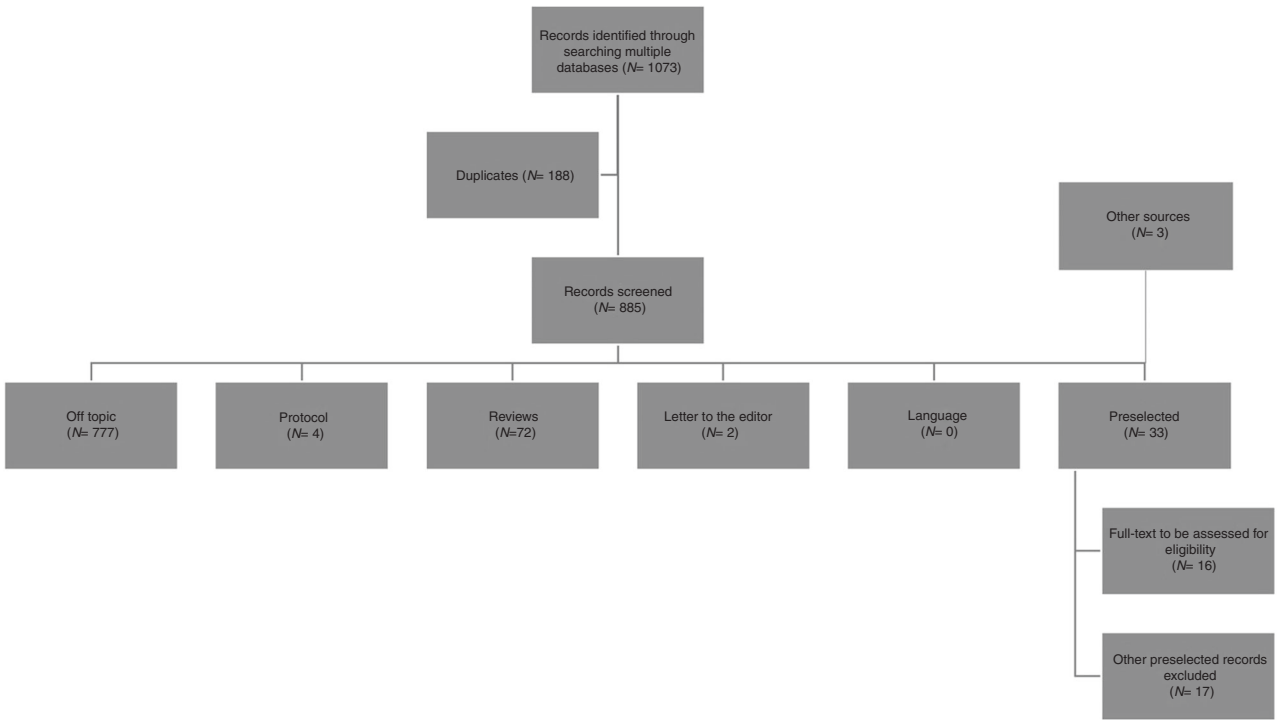

Fig. 1 Study flow: selection and description.

distinguish bacterial from viral infections. We found only one study that reported no benefit between transcript host RNA signatures and the discrimination of pathogens and involved 63 episodes of febrile neutropenia in 45 oncological children.[12] The average numbers of participants (much lower than in others which report an association) and the particular clinical condition of children analyzed in this study could be confounding factors.

DISCUSSION

We conducted a systematic review of host-response and pathogen-derived biomarkers for viral and bacterial infections in children to find a possible new test to optimize diagnosis and antibiotic therapy in children. Most of the studies analyzed in our systematic review highlighted that RNA biosignatures are an effective method to differentiate the different types of infections in children with infectious diseases of various types. The distinction between viral and SBIs is often difficult, with blood cultures (the gold standard for diagnosing sepsis) taking days to positivize and laboratory markers that can increase in different cases of inflammation.[24] The identification of new methods for early diagnosis and the distinction of bacterial from viral infections is one of the priorities of scientific research, which may have a major impact on clinical practice and quality of care with short- and long-term benefits.

Gene-expression signatures derived from whole blood have been reported for several diseases; one of the most important applications is in oncology, where these methods are used to distinguish among the types of tumor, identify stages of cancer, and predict the effectiveness of administered therapies.[25] About infectious diseases, the use of these diagnostics has not yet been fully defined, especially in children. However, it is known that peripheral blood cells share more than 80% of their transcriptome with those of specific organs; for this reason whole-blood expression profiling can be used to identify distinct disease signatures regardless of the site of infection.[26] Sherrianne et al.[27] showed that the transcriptional responses of very preterm infants with confirmed late onset sepsis differed significantly from non-sick children, for a total of 1317 differentially expressed genes. However, infants diagnosed with possible sepsis (with negative blood culture and altered CPR) could not be distinguished from children without or with confirmed infection; this limitation highlights the difficulty in transforming transcriptomics into a rapid and effective diagnostic test.

In this scoping review we identified 16 primary studies addressing the discrimination of bacterial and viral infections through transcript host RNA signatures, published between January 2011 and February 2021. Our findings indicate that these molecular tests could be used in combination with standard exams in children with fever to characterize any underlying infection (viral or bacterial), pending the blood culture report and other ongoing tests.

The differential gene expression taken into consideration in the analyzed studies is very large and allows to accurately distinguish bacterial infections from viral ones, even if the use of a few markers would seem adequate for the identification of certain pathologies. For example, Herberg et al.,[10] using only a 2-transcript signature, found that the expression of IFI44L was increased in patients with viral infection and FAM89A was increased in patients with bacterial infection, relative to healthy children. Similar results, using the same markers, were achieved in other studies.[9] Gomez-Carballa et al.[15] also found that the monogenic IFI44L expression signature yielded better results than those obtained from the 2-transcript test with sensitivity 90.9% and specificity 92.8%. Differences in transcriptomic expression are also reported on the basis of the type of bacterium or virus present, with markers more expressed by specific pathogens (such as Shigella) compared to pathogens belonging to the same category.[11] The use of targeted markers could improve diagnostic accuracy and, considering the smaller amount of transcripts to be analyzed, optimize costs.

The transcriptomic methods could also be compared with classic inflammation markers measured in pediatric age in the suspicion of infection, such as procalcitonin. In comparison to this, Lydon et al.[8] reported that the reverse transcriptase PCR (RT-PCR) test used in his study offered similar performance in distinguishing

Table 1. Characteristics of sources of evidence.

Reference First author Date	Study population	Aim (s)	Main result (s)	Study limitations
31631046 Emily C. Lydon (2019)[8]	151 subjects (children and adults)	Validate a microarray-derived host-response signature to distinguishing bacterial and viral respiratory infections and explore performance in microbiology-negative and coinfection cases. Reference standard was an adjudicated diagnosis of bacterial infection, viral infection, both, or neither. 41 transcripts: 34 for the bacterial model, 15 for the viral model, and 8 for the non-infectious illness model	With a targeted RT-PCR test, the bacterial, viral, and non-infectious classifiers had overall accuracies of 88%, 82%, and 84%, respectively. compared to an imperfect reference standard. In comparison to procalcitonin, the RT-PCR test offered similar performance in distinguishing bacterial versus non-bacterial infections but offered better discrimination in others such as viral vs. non-infectious illness and bacterial/viral coinfection	Lack of a gold standard to diagnose bacterial or viral infection. Discordant classifications could cause errors in adjudication or in the test
29795312 R. Barral-Arca (2018)[9]	174 children (<10 years)	Validate a biomarker signature consisting of 2-transcript host RNAs proposed for discriminating bacterial from viral infections in febrile children	IFI44L and FAM89A transcripts provide a strong signal to differentiate bacterial from viral infections, a signal that is non-population dependent, and useful for discriminating a wide range of pathogens and different levels of severity that the RNA test can discriminate between viral and bacterial causes of infection (sensitivity: 68% specificity: 84%), but the strength of the signal differs substantially depending on the causal pathogen, with the stronger signal being that of Shigella (sensitivity: 70% specificity: 100%)	Level of severity of the disease and causal pathogen affects the accuracy of the 2-biomarker test. The accuracy of this test improves significantly when excluding mild cases (sensitivity: 79% specificity: 78%)
27552617 J. A. Herberg (2016)[10]	240 children (median age 19 months) Validation group: 130 children	Identify a blood RNA expression signature that distinguishes bacterial from viral infection in febrile children	Bacterial infection were distinguished from those with viral infection using a 2-transcript signature (FAM89A and IFI44L). Expression of IFI44L was increased in patients with viral infection and FAM89A was increased in patients with bacterial infection, relative to healthy children. When this 2-transcript signature was implemented as a disease risk score in the validation group, all patients with microbiologically confirmed definite bacterial infection were classified as bacterial (sensitivity, 100%) and 27 of 28 patients with definite viral infection were classified as viral (specificity, 96.4%)	Lack of a reference standard. Validation of the signatures was undertaken in groups that included a high proportion of patients requiring intensive care
29377961 Hannah A. DeBerg (2018)[11]	164 children (under 10 years) Validation group: 29 children	Prove that pathogens causing diarrheal diseases would elicit distinct patterns of gene expression in host peripheral blood that could be used to characterize immune responses associated with each pathogen and with disease severity	The largest numbers of differentially expressed genes were observed in Shigella and rotavirus infections. Differential expression of genes related to chemokine receptors or inflammasome signaling in Shigella cases, such as CCR3, CXCR8, and NLRC4, and interferon response genes, such as IFI44 and OASL, in rotavirus cases. Rotavirus, but not Shigella or health controls (HC) groups, showed an increase in interferon response gene expression accompanied by an increase in the proportion of neutrophils ($r = 0.20$, $p = 0.24$ for Shigella; $r = 0.80$, $p = 8.4e{-}13$ for rotavirus; $r = 0.01$, $p = 0.95$ for HC). Complement expression was positively correlated with proportion of neutrophils in Shigella and rotavirus groups, but not in HC ($r = 0.40$, $p = 0.016$ for Shigella; $r = 0.62$, $p = 9.1e{-}7$ for rotavirus; $r = 0.05$, $p = 0.77$ for controls)	Other pathogens (like parasites) can contribute to childhood diarrheal disease. The transcriptional pathways are shared with these or other species that were outside of the scope of the study
32722616 Martina Wahlund (2020)[12]	45 children with 63 febrile neutropenia episodes; the RNA samples collected in 43 cases (0–18 years)	Examine if gene-expression profiling was feasible in children with febrile neutropenia who were undergoing cancer treatment	Gene-expression profiling is not suitable for determining the etiology of febrile neutropenia in immunosuppressed children during cancer treatment, because children with low WBCs or ANCs and, hence, an elevated risk of infection, have too few immune cells in their blood for reliable gene expression analysis	The cohort was heterogeneous. In one-third of the episodes had RNA concentrations that were too low for RNA sequencing

Table 1. continued

Reference First author Date	Study population	Aim (s)	Main result (s)	Study limitations
27837008 Jacob Silterra Jan (2017)[13]	68 children Validation group: 37 children (<10 years)	Perform RNA sequencing and analyze the transcriptomes of pediatric patients with well-characterized clinical phenotype to identify transcriptional features associated with each disease class	Bacterial markers are enriched in response/defense response to bacteria as well as in response to wounding and in inflammatory response, while viral markers are enriched in response/defense response to viruses and in interferon signaling.Transcriptional signatures of host response can be used to classify the underlying etiology for pediatric patients in tropical settings presenting with pneumonia. Myo10 (MyosinX) increased in bacterial infections. OLM4 in Gram positive. The major viral responses were type I interferon and cytokine signaling. Malarial responses, less well differentiated, included leukocyte activation and generic immune responses	Chest radiograph was used as gold standard for bacterial infections
28588308 D. L. Sampson (2017)[14]	1337 cases and 1106 controls (GEO dataset) Various Population	Discovered and validated a four-gene expression signature in whole blood, indicative of a general host systemic response to many types of viral infection. The signature's genes are: Interferon Stimulated Gene 15 (ISG15), interleukin 16 (IL16), 2′,5′-oligoadenylate synthetase like (OASL), and adhesion G protein coupled receptor E5 (ADGRE5)	The signature provides statistically significant ($p < 0.05$) discrimination between viral and non-viral conditions. The signature may have clinical utility for differentiating host systemic inflammation (SI) due to viral versus bacterial or non-infectious causes	The validation datasets were generated from multiple sample types (blood, liver biopsy, cultured hepatocytes) using multiple experimental methods (microarrays, RNA-seq). A group was under-represented in the validation data. The dataset analyzed did not include OASL
31409879 Alberto Gómez-Carballa (2019)[15]	25 children Validation group: 10 children Age 1.4–8.7	Evaluate a RT-qPCR assay for a 2-transcript host expression signature (FAM89A and IFI44L genes) inferred from microarray data that allow to differentiate between viral and bacterial infection in febrile children	FAM89A gene is highly expressed with respect to IFI44L gene in bacterial samples. This assay was able to discriminate viral from bacterial infections (sensitivity = 90.9%; specifcity = 85.7%). Unexpectedly, the monogenic IFI44L expression signature yielded better results than those obtained from the 2-transcript test (sensitivity = 90.9%; specifcity=92.8%). This gene alone differentiates between both groups, thus saving time, effort, and costs	Te DRS 2-transcript test was not able to differentiate between control (uninfected) and bacterial infected patients but it clearly distinguished the viral group from bacterial group and viral from control group (but, in the clinical practice, the utility of this test is the ability to efficiently separate viral from bacterial children with fever, without a control group)
25526183 X. D. Chen (2014)[16]	230 children (from 13 months to 14 years old) Validation group: 250 children	The changes in cytokine profiles from before and after infection in children with hemopathy in the bone marrow inhibition phase were analyzed to explore the utility of using this method to detect characteristics of infection and provide a basis for the rational use of antibiotics based on clinical CBA profiling	The effective rate (86.05%, $N = 43$) of infectious cases predicted was significantly higher than that obtained using traditional methods for selecting antibiotics based on clinical indications. Cytokine prophiles indicated the status of the corresponding bacterial infection. The sensitivity of cytokine detection was greater than that of CRP levels or blood culture, and the cytokine profiles during infection gave a better indication of the infection control state than could CRP levels. Gram-negative bacterium-infected children showed increased IL-6 and IL-10 levels in most cases; Gram-positive bacteria-infected children showed only a mild increase of IL-6 levels, and fungus-infected children showed only a mild increase of IFN-γ levels but not IL-6 or IL-10	This method could not predict infections with 100% accuracy. Therefore, other potential concurrent pathogenic factors or polyinfection might have been occurring in some cases
32127629 Ci-Xiu Li (2020)[17]	48 children 0–10 years Validation group: 7 children	Using total RNA sequencing to reveal the full spectrum of microbes associated with pediatric acute respiratory illness: determine the presence and abundance of viral, bacterial, and eukaryotic pathogens, and to reveal mixed infections, pathogen genotypes/subtypes, evolutionary origins, epidemiological history, and antimicrobial resistance	Total RNA sequencing is an efficient, accurate, and powerful means to characterize the infectome of a target tissue in clinical cases, especially in viruses A large proportion of the viruses, especially RNA viruses, were not identified in the classic diagnostic laboratory tests, and sometimes the tests were unavailable or not performed	Small number of children in the validation group. In the case of bacteria and fungi infections the meta-transcriptomics and standard diagnostic tests returned both consistent and inconsistent results, although transcriptomics had a lower bacterial detection rate

Table 1. continued

Reference First author Date	Study population	Aim (s)	Main result (s)	Study limitations
30473680 Kiyoshi F. Fukutanl (2018)[18]	58 children 6–23 months Validation group: 7 adults	Determine microbial (viral and bacterial) loads and perform immune transcriptional profiling in nasopharyngeal aspirates of children with ARI. Systems immunology approaches identified modular communities of immune response genes that are correlated with microbial load, thus distinguishing microbe-related immune signatures	Confirm a strong innate immune response in nasopharingeal swab, due to the presence of evolutionarily conserved type I Interferon (IFN)-stimulated genes (ISG), which was correlated with total bacterial and/or viral load. In comparison with indeterminate nasopharingeal swab, adaptive immunity transcripts discriminated among viral, bacterial, and codetected microbial profiles. In viral NPAs, B cell transcripts were significantly enriched among differentially expressed immune transcripts, while only type III IFN was correlated with viral load. In bacterial NPAs, myeloid cells and co-inhibitory transcripts were enriched and significantly correlated with bacterial load	Due to the cross-sectional design of the study and ethical limitations, the authors were unable to sample healthy age-matched infants
30425971 Rebecca G. Wallihan (2018)[19]	152 children Age 2.1–8.0 Validation group: 39 children	Evaluate the value of MDTH to assess disease severity in children hospitalized with CAP. The diagnosis of specific pathogens or discrimination between bacterial and viral etiologies was not the main aim.	Statistical group comparisons identified 5,675 differentially expressed transcripts between children with pyogenic bacteria CAP and healthy controls, 1,456 transcripts between those with detection of *M. pneumoniae* and healthy controls, and 4,104 transcripts between children with detection of only respiratory viruses and healthy controls. This gene list of 6,726 genes was then used to calculate the MDTH scores. MDTH scores were significantly higher in all pneumonia groups when compared with healthy controls. Children with detection of pyogenic bacteria had higher MDTH, CRP, and PCT values than those with detection of respiratory viruses. In children hospitalized with CAP, MDTH score measured within 24 h of admission was independently associated with longer duration of hospitalization, regardless of the pathogen detected.	Complete microbiologic data are lacking in some patients. The main aim was not the diagnosis of specific pathogens or discrimination between bacterial and viral etiologies but rather the identification of a biomarker to objectively assess disease severity. Healthy controls were not comprehensively tested for respiratory pathogens, analyses were not corrected for duration of illness prior to hospitalization
30962590 Matthew C. Altman (2019)[20]	106 children with 154 cold events 0–17 years	The main aim was to identify patterns of gene expression induced during events that progressed to asthma exacerbations versus those events that resolved without treatment with systemic corticosteroids. Virus PCR and partial sequencing were used to identify events associated with a respiratory virus or virus-negative events	High type 2 inflammation and low type I IFN response gene expression in nasal samples at baseline predicts exacerbation risk. Epithelial-associated SMAD3 signaling is upregulated and lymphocyte response pathways are downregulated early in exacerbation, followed by later upregulation of effector pathways including epidermal growth factor receptor signaling. There was a set of multiple inflammatory cell pathways involved in virus-associated exacerbations, in contrast to squamous cell pathways associated with non-viral exacerbations	The authors use upper-airway samples as a proxy for lower-airway events. The distinction between viral and non-viral exacerbation was a secondary endpoint
27552618 Prashant Mahajan (2016)[21]	279 randomly selected infants from 1883 children (median age 37 days) 19 afebrile healthy infants	Assess whether RNA biosignatures can distinguish febrile infants aged 60 days or younger with and without serious bacterial infections.	66 classifier genes distinguished infants with and without bacterial infections with 87% sensitivity and 89% specificity. Ten classifier genes distinguished infants with bacteremia from those without bacterial infections with 94% sensitivity and 95% specificity	RNA biosignatures were analyzed on frozen samples. Bacterial cultures were considered the reference standard for purposes of analysis, despite a known substantial rate of false-positive and false-negative results. Viral testing was not consistently performed for all study participants. The ability to interpret the biosignatures of infants with potential bacterial-viral coinfections was limited.

Table 1. continued

Reference First author Date	Study population	Aim (s)	Main result (s)	Study limitations
31983492 Fran Balamuth (2020)[22]	66 children with confirmed bacterial or viral infections Age between 57 days and 18 years 12 healthy controls	Primary outcome was source pathogen, defined as confirmed bacterial source from sterile body fluid or confirmed viral source. Secondary outcome was establish sepsis severity, defined as maximum therapy required for shock reversal in the first 3 hospital days	558 differentially expressed genes, many related to interferon signaling or viral immunity. Authors did not find statistically significant gene expression differences in patients according to sepsis severity	Single center study. Convenience sample of subjects. The expression signatures observed could represent differences in cell types, rather than gene expression (but authors performed an adjustment for white blood cells count). RNA expression patterns could change over time
23858444 Xinran Hu (2013)[23]	30 febrile children Age between 2 and 36 months 22 afebrile controls	Analyze blood leukocyte transcriptional profiles associated with specific viral infection and distinguish children with viral infection from children with bacterial infection. Distinguish if symptomatic and asymptomatic children infected with the same virus have distinctively different leukocyte transcriptional profiles	Blood leukocyte transcriptional profiles clearly distinguished virus-positive febrile children from both virus-negative afebrile controls and afebrile children with the same viruses present in the febrile children. Virus-specific gene expression profiles could be defined. The IFN signaling pathway was uniquely activated in febrile children with viral infection, whereas the integrin signaling pathway was uniquely activated in children with bacterial infection. Transcriptional profiles classified febrile children with viral or bacterial infection with better accuracy than white blood cell count in the blood	The number of children positive for each virus under study was limited, samples were obtained at different times and were not always from the same time period with respect to the onset of the illness. Cases and controls were not matched according to demographic characteristics, especially race

bacterial versus non-bacterial infections but offered better discrimination in viral versus non-infectious illness or bacterial/viral coinfection. This study suggests that transcriptomics can be used not only as an isolated diagnostic tool, but also with other standard tests, increasing sensitivity and specificity of the diagnosis.

In the only study where no benefit was found, Wahlund et al.[12] analyzed 63 episodes of febrile neutropenia in 45 children suffering from different oncological pathologies. The authors summarize that immunosuppressed children do not have enough immune cells in the blood for a reliable analysis of gene expression, suggesting that transcriptomics may not be used to accurately determine the etiology of infections in febrile neutropenia. However, single-cell RNA expression technology could be used to overcome the concern with cell number/relative proportions.[28] The role of transcriptomic in immunosuppressed patients should therefore be carefully evaluated in future studies also using new single-cell RNA technologies, even if in the other study conducted on oncohematological patients[16] the authors showed increased specific cytokines levels within 24 h after infection that correlate with the corresponding bacterial infection. Sensitivity of cytokine detection soon after infection was greater than that of CRP or blood culture levels, and cytokine profiles during infection provided a better indication of infection control status than CRP levels. The use of transcriptomics in frail patients (such as oncological children or other immunosuppressed patients) would be essential to guide antibiotic therapies during infectious processes, which are much more frequent than in the general population. Immune deficiencies make the use of these methods more difficult; therefore, future studies should investigate this relationship to evaluate their real effectiveness.

A potential important application of transcriptomics in infectious diseases is the possibility of establishing targeted and personalized therapies instead of general empirical treatments. However, there is no single laboratory test that can accurately identify sepsis during the early stages of disease; among the studies we have analyzed, one of the most frequently reported limitations is the lack of a gold standard for diagnosing infection. For example, Mahajian et al.[21] consider bacterial etiology as the bacterial detection in culture, despite a known substantial rate of false-positive and false-negative results. Negative cultures may reflect antibiotic abuse, low numbers of bacteria in the culture or inaccessible sites of infection. Difficulty in differentiating bacterial, viral, and non-infectious etiologies of children diseases contributes to antibiotic overuse. It is known that the inappropriate use of antibiotic therapy (both in terms of duration and spectrum of action of the drug) can lead to the development of bacterial resistance through genetic and epigenetic modifications[29,30] favoring the onset of subsequent infections that are difficult to treat.[31] This is especially true for pediatric patients, who often develop fever due to viruses and are still treated with antibiotic therapies to cover up any bacterial infections.[32] A more accurate diagnosis would reduce the improper use of antibiotics and would allow the use of targeted drugs, with less prescription of empirical therapies. Sweeney et al. used their previously published 11-gene Sepsis MetaScore together with the new bacterial/viral classifier to build an integrated antibiotics decision model. In a pooled analysis of 1057 samples from 20 cohorts, the integrated antibiotics decision model had a sensitivity and specificity for bacterial infections of 94.0% and 59.8%, respectively (negative likelihood ratio, 0.10).[33] However, they excluded infants, which represent one of the most challenging populations in pediatric practice, where clinical signs and symptoms are often subtle and difficult to understand, and discriminating viral from bacterial infections is particularly important but, at the same time, especially difficult. Similarly, the definition of the spectrum of conditions behind a febrile child can help advancing clinical practice and pathophysiologic understanding of underlying mechanisms.

The mentioned difficulties in defining and, therefore, recognizing bacterial and viral infections highlight a relevant point to be taken into account in the assessment of new biomarkers: the clinical definition of viral and bacterial infections, which ultimately affect all clinical decisions and also the proper assessment of every study. There is increasing understanding that the historic dichotomous distinction of febrile children in the cohort of bacterial and viral infections is not only difficult to achieve (due to sensitivity problems of microbiological gold standards), but it is probably not appropriate from a clinical point of view.[34] There is, in fact, emerging evidence for a complex relationship between bacterial and viral pathogens and their interplay in the etiology, clinical presentation, and determination of disease severity.[35] Moreover, biomarkers can have significant overlaps that cannot simply be explained by a dichotomous classification, as found by a recent and innovative approach proposed by Nijman et al.[34] The PERFORM consortium (Personalized Risk assessment in febrile children to optimize Real-life Management across the European Union) recently proposed an algorithm that showed better discrimination between patients with bacterial and viral infection, aligned better with host-response biomarker concentrations and clinically relevant, patients-centered outcomes, providing a framework for phenotyping children with infections based on the trends in the different biomarkers in relation to the certainty of the diagnosis of either bacterial or viral categories. In this new algorithm, a more granular classification that also adds the "probable bacterial" or "probable viral" infections groups gives a wider range of clinical phenotypes that better reflects clinical practice. We strongly support this view, and encourage future studies in the field of biomarkers to prospectively consider this perspective and include these categories to have a proper interpretation of biomarkers according to these outcomes. The last decade has seen a similar change in the field of tuberculosis, where the classic dichotomous classification of active tuberculosis and latent infection has now been substituted by the so-called TB-spectrum, which better reflects both clinical observations and biomolecular phases of the complex interplay between the host and *M. tuberculosis*.[36,37] Similarly, the definition of the spectrum of conditions behind a febrile child can help advancing clinical practice and pathophysiologic understanding of underlying mechanisms.

The transcriptomic profile of infections is not only influenced by the underlying pathogen but also by the immune system of the involved person; the expression signatures observed could represent differences in cell types and levels, rather than gene expression. It is therefore possible that higher cellular concentrations can provide higher amounts of RNA, while low levels are associated with an often unsatisfactory transcriptomic profile. In many of the other studies analyzed there was no normalization of the values for the number of white blood cells, and this could be an important confounding factor. Balamuth et al.[22] excluded from the cohort of his study children with absolute neutrophil counts $<1000 \times 10^9$/L because they were not likely associated with sufficient quantities of RNA for transcriptomic analysis. Therefore, this aspects is particularly relevant when the transcriptomics techniques are implemented in a center, or in the evaluation of results from different studies. Similarly, this problem should always be taken into account when specific cohort of patients with impaired white blood cell count for underlying conditions are included in the analyses, since false-negative results may have serious implications in the most fragile patients, such as the oncoheamatologic ones. However, as previously mentioned, new single-cell technologies can potentially bypass these problems.[28]

The importance of the contribution of white blood cell counts to the interpretation of RNA-transcriptomic has been addressed by Monaco et al., which highlighted how studying o peripheral blood mononuclear cells (PBMCs) in their entirety often contributes to inconclusive or difficult to interpret results, as it not always possible to accurately ascertain which specific immune cell types are responsible for any given transcriptomic signal of interest. Moreover, the proportion of immune cell subsets in the blood can vary during disease, age, or clinical interventions.[38] To address these issues, they identified sets of genes that are specific, are co-expressed, across different cell types, provided mRNA abundance revealing cell type specificity and developed transcriptomics signatures normalized by mRNA abundance, eventually providing transcriptomic data specific for different immune subsets. However, studies including patients of different ethnic backgrounds and age groups will be still necessary in order to allow a better characterization of specific cohort of patients and a more standardized clinical application.

Importantly, the method of collecting, transporting, and storing samples for transcriptomic analysis can also influence the results of gene expression profiles[39] causing ex vivo alterations in mRNA transcripts. High-quality sample preparation and immediate RNA isolation are prerequisites for the best transcriptional analysis[40] although many times the samples are stored and processed later. For example, Prashant et al.[21] used frozen samples for RNA analyses and specifies it in the limitations of their work.

One of the problems that hinder the application of transcriptomics in clinical practice is the high cost and the need for dedicated skills for the interpretation of results. The bed side use of this technology is the key and fundamental point for the clinical application, even if to date the transcripts to be analyzed for a correct differential diagnosis are several and, in the context of the creation of a device for the detection of signatures, evidence is available only for some transcripts. However, incredible improvements have been achieved in the last decade in the implementation of easier platforms that are now able to allow transcriptomic analysis at a single-cell resolution, simultaneous analysis of host and pathogen transcriptomes (dual RNA-seq), and sequencing of full-length transcripts to allow detailed analysis of transcript isoforms and direct analysis of RNA.[41] In the future, techniques such as laser capture microdissection may be coupled with RNA-seq to allow host cells and their interacting pathogens to be isolated and studied as defined cell groups or dual single-cell analyses. Technical advances are also behind the corner. Measurement of the release of hydrogen ions during the nucleotide incorporation process with a pH semiconductor system[42] could facilitate in the future the creation of devices for the analysis of nucleic acids in real time using portable and low-cost devices. Moreover, studies are underway that attempt to simplify the use of these methods, such as the one conducted by Pennisi et al.;[43] they proved that EMR1-ADGRE (replacing FAM89A) and IFI144L transcript signatures can be detected using an inexpensive semiconductor-based sensing platform. Using this device, host RNA signatures can be detected rapidly and cost-effectively, at a lower cost than traditional diagnostic methods, in a format suitable for development as a point-of-care diagnostic test that might be applied to a range of different clinical settings. Last, new and less invasive routes for host RNA signatures have been tested, including the nasal environment. Jinsheng et al.[44] found that gene expression signatures were as good or better for discriminating between respiratory syncytial virus (RSV) infection, symptomatic non-RSV respiratory virus infection, asymptomatic rhinovirus infection, and virus-negative asymptomatic controls. Although the authors did not include a control group of bacterial or probable-bacterial infections, this is a promising area of investigation.

Limitations

Our scoping review has some limitations. First of all, a formal assessment of publication bias was not performed. We know also that transcriptomics is a very recent topic, which is being studied not only for infections but also for other pathologies such as oncological ones; our systematic review includes all relevant

publications up to February 2021 but it is likely that further works about transcriptomic signatures and pediatric infection will be published during these months. Additionally, it is always possible in a systematic review that relevant publications were not identified in the search; however, the search term was extensively calibrated by authors. In addition, the available studies used different gold standards to define sepsis and SBIs, and the lack of a clinically relevant gold standard (children can have a SBI or sepsis with negative blood cultures) further limits the generalization of results of this review. Last, differences in the size of the datasets, heterogeneity of included patients, and wide variation of biosignatures analyzed between articles further limit the generalizability of our studies, and should be taken into account for future researches.

Future studies should target their biomarker research for a new diagnostic test intended for use in children and should be performed in diverse geographical places, involving different populations, because the diagnostic test can be applied in different settings.

CONCLUSIONS

Transcript host RNA signatures may have the potential role to support clinicians in differentiating bacterial from viral infections in febrile children and can be used with traditional diagnostic methods to determine if and when to start antibiotic therapy. A more accurate diagnosis of infectious diseases would improve diagnosis of SBIs in febrile children and optimize the use of drugs, reducing adverse effects such antibiotic resistance or dysregulation of pediatric microbiota. While the summarized results are promising, our review identified a wide variability of target populations and gold standards used to define sepsis and SBIs, limiting the generalization of our findings. Importantly, a new classification of the spectrum of viral and bacterial infections in children, which better reflects routine practice, seems necessary to have more standardized and comparable outcome measures, and should be considered in future studies that should address the role of transcript host RNA signatures in defining this spectrum.

Given the importance of the topic, we hope that this systematic review will contribute to provide targeted guidance for further scientific explorations toward the eventual development of a new diagnostic tool to differentiate bacterial and viral infections in children.

DATA AVAILABILITY

The main results of studies included in the review are available in the main text and Supplementary material.

REFERENCES

1. Goldstein, B., Giroir, B. & Randolph, A. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr. Crit. Care Med.* **6**, 2–8 (2005).
2. Liu, L. et al. Global, regional, and national causes of child mortality in 2000–13, with projections to inform post-2015 priorities: an updated systematic analysis. *Lancet* **385**, 430–440 (2015).
3. Memar, M. Y., Varshochi, M., Shokouhi, B., Asgharzadeh, M. & Kafil, H. S. Procalcitonin: the marker of pediatric bacterial infection. *Biomed. Pharmacother.* **96**, 936–943 (2017).
4. Oeser, C. et al. PCR for the detection of pathogens in neonatal early onset sepsis. *PLoS ONE* **15**, e0226817 (2020).
5. Patel, K. & McElvania, E. Diagnostic challenges and laboratory considerations for pediatric sepsis. *J. Appl. Lab. Med.* **3**, 587–600 (2019).
6. Woll, C., Neuman, M. I. & Aronson, P. L. Management of the febrile young infant: update for the 21st century. *Pediatr. Emerg. Care* **33**, 748–753 (2017).
7. Suarez, N. M. et al. Superiority of transcriptional profiling over procalcitonin for distinguishing bacterial from viral lower respiratory tract infections in hospitalized adults. *J. Infect. Dis.* **212**, 213–222 (2015).
8. Lydon, E. C. et al. Validation of a host response test to distinguish bacterial and viral respiratory infection. *EBioMedicine* **48**, 453–461 (2019).
9. Barral-Arca, R. et al. A 2-transcript host cell signature distinguishes viral from bacterial diarrhea and it is influenced by the severity of symptoms. *Sci. Rep.* **8**, 8043 (2018).
10. Herberg, J. A. et al. Diagnostic test accuracy of a 2-transcript host RNA signature for discriminating bacterial vs viral infection in febrile children. *JAMA* **316**, 835–845 (2016).
11. DeBerg, H. A. et al. Shared and organism-specific host responses to childhood diarrheal diseases revealed by whole blood transcript profiling. *PLoS ONE* **13**, e0192082 (2018).
12. Wahlund, M. et al. The feasibility of host transcriptome profiling as a diagnostic tool for microbial etiology in childhood cancer patients with febrile neutropenia. *Int. J. Mol. Sci.* **21**. https://doi.org/10.3390/ijms21155305 (2020).
13. Silterra, J. et al. Transcriptional categorization of the etiology of pneumonia syndrome in pediatric patients in malaria-endemic areas. *J. Infect. Dis.* **215**, 312–320 (2017).
14. Sampson, D. L. et al. A four-biomarker blood signature discriminates systemic inflammation due to viral infection versus other etiologies. *Sci. Rep.* **7**, 2914 (2017).
15. Gómez-Carballa, A. et al. A qPCR expression assay of IFI44L gene differentiates viral from bacterial infections in febrile children. *Sci. Rep.* **9**, 11780 (2019).
16. Chen, X. D. et al. Effectiveness of bacterial infection-related cytokine profile (BIRCP) determination for monitoring pathogen infections in children with hemopathy in the bone marrow inhibition phase. *Genet. Mol. Res.* **13**, 10622–10631 (2014).
17. Li, C.-X. et al. High resolution metagenomic characterization of complex infectomes in paediatric acute respiratory infection. *Sci. Rep.* **10**, 3963 (2020).
18. Fukutani, K. F. et al. In situ immune signatures and microbial load at the nasopharyngeal interface in children with acute respiratory infection. *Front. Microbiol.* **9**, 2475 (2018).
19. Wallihan, R. G. et al. Molecular distance to health transcriptional score and disease severity in children hospitalized with community-acquired pneumonia. *Front. Cell. Infect. Microbiol.* **8**, 382 (2018).
20. Altman, M. C. et al. Transcriptome networks identify mechanisms of viral and nonviral asthma exacerbations in children. *Nat. Immunol.* **20**, 637–651 (2019).
21. Mahajan, P. et al. Association of RNA biosignatures with bacterial infections in febrile infants aged 60 days or younger. *JAMA* **316**, 846–857 (2016).
22. Balamuth, F. et al. Gene expression profiles in children with suspected sepsis. *Ann. Emerg. Med.* **75**, 744–754 (2020).
23. Hu, X., Yu, J., Crosby, S. D. & Storch, G. A. Gene expression profiles in febrile children with defined viral and bacterial infection. *Proc. Natl Acad. Sci. USA* **110**, 12792–12797 (2013).
24. Pontrelli, G. et al. Accuracy of serum procalcitonin for the diagnosis of sepsis in neonates and children with systemic inflammatory syndrome: a meta-analysis. *BMC Infect. Dis.* **17**, 302 (2017).
25. Gliddon, H. D., Herberg, J. A., Levin, M. & Kaforou, M. Genome-wide host RNA signatures of infectious diseases: Discovery and clinical translation. *Immunology* **153**, 171–178 (2018).
26. Kaforou, M., Wright, V. J. & Levin, M. Host RNA signatures for diagnostics: an example from paediatric tuberculosis in Africa. *J. Infect.* **69**, S28–S31 (2014).
27. Ng, S. et al. Whole blood transcriptional responses of very preterm infants during late-onset sepsis. *PLoS ONE* **15**, e0233841 (2020).
28. Potter, S. S. Single-cell RNA sequencing for the study of development, physiology and disease. *Nat. Rev. Nephrol.* **14**, 479–492 (2018).
29. Ghosh, D. et al. Antibiotic resistance and epigenetics: more to it than meets the eye. *Antimicrob. Agents Chemother.* **64**, e02225-19 (2019).
30. Bakhit, M. et al. Resistance decay in individuals after antibiotic exposure in primary care: a systematic review and meta-analysis. *BMC Med.* **16**, 126 (2018).
31. Nicolini, G., Sperotto, F. & Esposito, S. Combating the rise of antibiotic resistance in children. *Minerva Pediatr.* **66**, 31–39 (2014).
32. Esposito, S. et al. Approach to neonates and young infants with fever without a source who are at risk for severe bacterial infection. *Mediators Inflamm.* **2018**, 1–11 (2018).
33. Sweeney, T. E., Wong, H. R. & Khatri, P. Robust classification of bacterial and viral infections via integrated host gene expression diagnostics. *Sci. Transl. Med.* **8**, 346ra91 (2016).
34. Nijman, R. G. et al. A novel framework for phenotyping children with suspected or confirmed infection for future biomarker studies. *Front. Pediatr.***9**, 688272 (2021).
35. de Steenhuijsen Piters, W. A. A. et al. Nasopharyngeal microbiota, host transcriptome, and disease severity in children with respiratory syncytial virus infection. *Am. J. Respir. Crit. Care Med.* **194**, 1104–1115 (2016).
36. Buonsenso, D. et al. The tuberculosis spectrum: translating basic research into pediatric clinical practice. *Med. Hypotheses* **141**, 108091 (2020).
37. Pai, M. et al. Tuberculosis. *Nat. Rev. Dis. Prim.* **2**, 16076 (2016).

38. Monaco, G. et al. Rna-seq signatures normalized by mrna abundance allow absolute deconvolution of human immune cell types. *Cell Rep.* **26**, 1627–1640.e7 (2019).
39. Donohue, D. E. et al. Gene expression profiling of whole blood: a comparative assessment of RNA-stabilizing collection methods. *PLoS ONE* **14**, e0223065 (2019).
40. Debey, S. et al. Comparison of different isolation techniques prior gene expression profiling of blood derived cells: impact on physiological responses, on overall expression and the role of different cell types. *Pharmacogenomics J.* **4**, 193–207 (2004).
41. Lee, H. J. et al. (2018). Transcriptomic studies of malaria: a paradigm for investigation of systemic host-pathogen interactions. *Microbiol. Mol. Biol. Rev.* **82**, https://doi.org/10.1128/MMBR.00071-17.
42. Toumazou, C. et al. Simultaneous DNA amplification and detection using a pH-sensing semiconductor system. *Nat. Methods* **10**, 641–646 (2013).
43. Pennisi, I. et al. Translation of a host blood RNA signature distinguishing bacterial from viral infection into a platform suitable for development as a point-of-care test. *JAMA Pediatr.* **175**, 417 (2021).
44. Yu, J. et al. Host gene expression in nose and blood for the diagnosis of viral respiratory infection. *J. Infect. Dis.* **219**, 1151–1161 (2019).

ACKNOWLEDGEMENTS

All of the authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work. This research received no external funding.

AUTHOR CONTRIBUTIONS

Conceptualization: D.B. and P.V.; methodology, resources, and writing—original draft preparation: G.S. and D.B.; writing—review and editing: D.B. and G.S.; supervision: P.V. All authors have read and agreed to the published version of the manuscript.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41390-021-01890-z.

Correspondence and requests for materials should be addressed to Danilo Buonsenso.