REVIEW ARTICLE

Pediatric sepsis biomarkers for prognostic and predictive enrichment

Hector R. Wong

© The Author(s), under exclusive licence to the International Pediatric Research Foundation, Inc 2021

Sepsis is a major public health problem in children throughout the world. Given that the treatment guidelines emphasize early recognition, there is interest in developing biomarkers of sepsis, and most attention is focused on diagnostic biomarkers. While there is a need for ongoing discovery and development of diagnostic biomarkers for sepsis, this review will focus on less well-known applications of sepsis biomarkers. Among patients with sepsis, the biomarkers can give information regarding the risk of poor outcome from sepsis, risk of sepsis-related organ dysfunction, and subgroups of patients with sepsis who share underlying biological features potentially amenable to targeted therapeutics. These types of biomarkers, beyond the traditional concept of diagnosis, address the important concepts of prognostic and predictive enrichment, which are key components of bringing the promise of precision medicine to the bedside of children with sepsis.

Pediatric Research (2022) 91:283–288; https://doi.org/10.1038/s41390-021-01620-5

INTRODUCTION

Current treatment guidelines for sepsis emphasize a high index of suspicion for early detection of sepsis and timely administration of antibiotics as key principles for improving outcomes. Consequently, there is considerable interest in sepsis biomarkers, which provide information regarding the treatment guideline application. When considering “sepsis biomarkers” it is important to understand the clinical context and what information one seeks to derive from the biomarker data.

Evaluating febrile children, who are otherwise clinically undifferentiated, is a common challenge for pediatricians. In this context, the biomarker data can be useful in differentiating those with infection from those with a sterile source of systemic inflammation as the underlying cause of fever. Further, the biomarkers can also be useful in differentiating those with bacterial infection from those with a viral infection, as this would help in deciding whether or not to prescribe antibiotics. Procalcitonin, although imperfect, has emerged as a widely used biomarker to address these challenges in both the ambulatory and inpatient settings. There are also emerging gene expression panels that show substantial promise in addressing these clinical challenges. These panels have the potential to capture the complexity of the human response to infection more accurately than single biomarker strategies, and therefore can have added value when compared with single biomarker strategies.

While diagnostic biomarkers for sepsis are of great clinical importance, this important topic will not be the focus of this review, as it is well covered in the recent review literature. Rather, this review will address other important clinical questions, less well covered in the existing literature, that could be addressable by the discovery and development of reliable biomarkers. The clinical questions addressed in this review refer to patients diagnosed with sepsis based on the current criteria and include:

Among patients with sepsis, what is the risk for a poor outcome? Among patients with sepsis, what is the risk for clinically important organ failure? Among patients with sepsis, do there exist identifiable subgroups that share common underlying biological features?

As will be discussed in subsequent sections, effectively addressing these questions, via the discovery and development of biomarkers, is fundamental to bringing the promise of precision medicine to the bedside of children with sepsis.

BIOMARKERS TO ESTIMATE THE RISK OF POOR OUTCOME FROM SEPSIS

Mortality secondary to sepsis remains a major public health problem for children in the United States and around the world. Reliably estimating baseline mortality risk from pediatric sepsis is a form of prognostic enrichment, which refers to the selection of a patient cohort more likely to have an outcome of interest when compared with an unselected cohort. Prognostic enrichment is fundamental to clinical decision making, resource allocation, quality improvement efforts, and research. While many biomarkers have been evaluated to estimate the risk for a poor outcome from sepsis, few have been validated in a sufficiently rigorous manner to warrant clinical adoption as a prognostic enrichment strategy.

The Pediatric Sepsis Biomarker Risk Model (PERSEVERE) incorporates a panel of serum protein biomarkers, measured within 24 h of a sepsis diagnosis, to estimate baseline risk of mortality among critically ill children with septic shock. Figure 1 depicts the general approach for identifying the PERSEVERE biomarkers by leveraging the discovery potential of transcriptomics based on whole-blood-derived RNA. In a cohort of...
Generically ill children with septic shock, complementary bioinformatics and machine learning approaches were employed to derive a list of candidate predictor genes for 28-day mortality\cite{20,24}. From over 100 candidate genes, the list was further defined using two a priori criteria: (1) biological plausibility linking the candidate predictor gene and sepsis-related pathobiology and (2) the ability to measure the gene product (i.e., protein) in the blood compartment. Based on these criteria, 12 candidate biomarkers were selected for further evaluation (Table 1).

The serum protein concentrations of these 12 biomarkers were then measured in a cohort of children with septic shock to model the risk of 28-day mortality using Classification and Regression Tree methodology\cite{28}. The resulting model, PERSEVERE, included five of the original 12 candidate biomarkers (CCL3, IL8, HSPA1B, GZMB, and MMP8) and age as predictor variables. Subsequently, PERSEVERE was recalibrated to yield PERSEVERE II, which includes admission platelet count as an additional predictor variable\cite{29}. The PERSEVERE II decision rules were locked and prospectively applied to a validation cohort of 69% (95% CI 64–73) for distinguishing non-survivors of the original 12 candidate biomarkers\cite{29}. Among these, 11 overlapped with the genes identified in the initial discovery phase of the PERSEVERE biomarkers (Table 2), thus adding a measure of internal validity among the various models. A recent approach employing artificial intelligence and transcriptomic data further confirmed the reliability of these models with respect to the identified classifier genes\cite{29}.

### Biomarkers to estimate the risk of sepsis-associated organ dysfunction

While all organs are at risk of dysfunction during sepsis, acute kidney injury (AKI) is a particularly common complication of sepsis, and independently contributes to increased sepsis-related morbidity and mortality\cite{30,31}. The underlying pathology of sepsis-associated AKI involves a complex and incompletely understood interplay of deranged perfusion, oxidant stress, and inflammation. Consequently, there is interest in developing strategies to identify children at greater risk of developing sepsis-associated AKI\cite{31}. This form of prognostic enrichment focused on the development of sepsis-associated AKI rather than mortality, could enable the targeted application of renal protective strategies and renal replacement therapy, and in the future could serve to inform enrollment for interventional clinical trials to mitigate sepsis-associated AKI.

Recently, the PERSEVERE biomarkers were used to develop a model to estimate the risk of sepsis-associated AKI in a cohort of critically ill children with septic shock\cite{26}. The primary outcome variable for this study was the development of severe AKI on day 3 of septic shock. Severe sepsis-associated AKI was defined as Kidney Disease Improving Global Outcomes stage 2 AKI or...
greater, which is at least a two-fold increase in serum creatinine concentration compared to baseline.

The derived model estimated the risk of severe sepsis-associated AKI with an AUROC of 0.95 (95% CI 0.92–0.98), a sensitivity of 92% (95% CI 82–97), and a specificity of 89% (95% CI 85–92). A ten-fold cross-validation procedure yielded a summary AUROC of 0.88, and when tested using a historical cohort, the model had an AUROC of 0.83 (95% CI 0.79–0.88). In addition, the model was able to identify patients with evidence of AKI at the time of admission to the pediatric intensive care unit, but who would go on to recover renal function by day 3 of septic shock. This model requires prospective validation, but its performance characteristics warrant further exploration as a potential prognostic enrichment strategy for sepsis-associated AKI.

In a conceptually related study, the PERSEVERE biomarkers were evaluated for their ability to identify the risk of clinical deterioration among immunocompromised hospitalized patients being evaluated for possible infection. The enrollment criteria captured patients in whom a blood culture was obtained as part of the workup for suspected infection. The PERSEVERE biomarkers were measured at the same time as the blood culture sample. Clinical deterioration was defined by a set of criteria to codify evolving organ failure requiring intensive care and was tracked up to 72 h from enrollment. This study showed that the PERSEVERE biomarkers were able to identify patients more likely to experience clinical deterioration (i.e., evolving organ dysfunction) with a sensitivity of 86% and a negative predictive value of 97%. This was a single-center study that requires validation, as well as expansion to patients being evaluated for infection who are not immunocompromised. For example, it could be productive to evaluate the PERSEVERE biomarkers among immune-competent children who are being evaluated in the ambulatory setting for possible infection. In this context, the PERSEVERE biomarkers could be evaluated for their ability to identify patients with a greater likelihood of clinical deterioration, therefore informing treatment and disposition decisions.

**Gene expression-based endotypes of septic shock**

Endotypes represent subgroups of heterogeneous clinical syndromes as defined by commonalities of underlying biology. Recently, there has been considerable interest in identifying endotypes of septic shock using genome-wide expression profiling. The underlying premise of these efforts is that unbiased transcriptomic studies can reveal endotypes of sepsis based on coordinated gene expression patterns (i.e., biomarkers) shared among endotypes. The general approach involves unsupervised clustering algorithms based on commonalities of gene expression, agnostic to patient characteristics, clinical trajectory, and outcome. Ultimately, the goal is to leverage gene expression-based endotypes for personalized treatment and disposition decisions.

**Table 2.** Genes common to a community approach for identification of prognostic enrichment genes and the original working gene list for the PERSEVERE biomarkers.

| Gene symbol | Description |
|-------------|-------------|
| CD24        | CD24 molecule |
| CEACAM8     | Carcinoembryonic antigen-related cell adhesion molecule 8 |
| CXS2        | CDC28 protein kinase regulatory subunit 2 |
| CX3CR1      | C-X3-C motif chemokine receptor 1 |
| DDX4        | DNA damage-inducible transcript 4 |
| EMR3        | Adhesion G-protein-coupled receptor E3 |
| G0S2        | G0/G1 switch 2 |
| IL8         | Interleukin-8, C-X-C motif chemokine ligand 8 |
| MAFF        | MAF bZIP transcription factor F |
| RGS1        | Regulator of G-protein signaling 1 |
| TGFBI       | Transforming growth factor beta induced |

**Fig. 2 The PERSEVERE II decision tree.** All patients begin at the root node and are subsequently partitioned to daughter nodes using biomarker-based criteria as indicated at the top of each node. All biomarker data are shown as pg/ml, and platelet data are shown as the number of platelets (K/μl). Daughter nodes that can no longer be partitioned are called terminal nodes (TNs) and shown in red outline. The terminal nodes are used to assign a baseline mortality probability to a patient classified to a given terminal node. The baseline mortality probability (MP) corresponding to each terminal node is indicated in red font within each terminal node and is derived from the published PERSEVERE II model. These baseline mortality risks are used for the construction of the AUROC. For calculation of the diagnostic test characteristics, the mortality probability is dichotomized into those who are predicted to survive and those who are predicted to not survive by 28 days. Patients allocated to TN1, TN2, TN5, and TN8 (mortality risk 0.000–0.019) are classified as predicted survivors. Patients allocated to TN3, TN4, TN6, TN7, and TN9 are classified as predicted non-survivors (mortality risk 0.167–0.444).
expression-based endotyping as a predictive enrichment tool. Distinct from prognostic enrichment, predictive enrichment refers to the selection of a patient cohort in whom an intervention is more likely to be effective based on underlying biology when compared with an unselected cohort. Like prognostic enrichment, predictive enrichment is thought to be a key enabler of precision medicine.

In 2016, Davenport and colleagues reported two gene expression-based endotypes of sepsis among adults, termed sepsis response signature 1 and 2 (SRS1 and SRS2). The gene expression profile defining SRS1 corresponded to endotoxin tolerance, repression of human leukocyte antigen class II, and T cell exhaustion. Patients assigned to SRS1 had a higher mortality rate compared with patients assigned to SRS2. The existence of SRS1 and SRS2 was validated in a subsequent study of adults with sepsis and appears to be applicable among patients who have either a pulmonary or intra-abdominal source of sepsis.

In 2017, Scicluna et al. identified four gene expression-based endotypes among adults with sepsis, termed Mars1 to 4. Patients assigned to the Mars1 endotype had the highest risk of mortality, and like the SRS1 endotype, the Mars1 endotype was characterized by repression of gene programs corresponding to innate and adaptive immunity.

In 2018, Sweeney et al. leveraged virtually all publicly available transcriptomic data involving adults, children, and neonates with sepsis to identify three endotypes of sepsis. These were termed inflammatory, adaptive, and coagulopathic, reflecting their respective endotype-defining gene expression signatures. The coagulopathic endotype was associated with the highest mortality rate among the three endotypes. Recently, the existence of these three endotypes and their association with mortality was reported among critically ill patients with coronavirus disease 2019.

Analogous gene expression endotypes of septic shock were reported among critically ill children in 2009, and subsequently validated. Pediatric septic shock endotypes A and B are defined by 100 genes corresponding to the adaptive immune system and glucocorticoid receptor signaling. Endotype A assignment is independently associated with higher mortality and organ failure burden when compared to endotype B assignment. Approximately one-third of children with septic shock change endotype assignment during the initial 3 days of septic shock. Children who persist as endotype A during the entire 3-day acute period of septic shock are at the highest risk for poor outcome. Finally, an interaction between younger age, assignment to endotype A, and increased risk of mortality was reported among adults with septic shock.

The aforementioned studies share an approach to endotype identification based on transcriptomic data generated using whole-blood-derived RNA and clustering algorithms to classify patients with sepsis into subgroups that share common patterns of gene expression. All the studies report associations between endotype assignment and mortality, but this is not the primary intent of identifying endotypes. Rather, the primary intent of endotype identification among patients with sepsis is to identify subgroups with common underlying biological features that could potentially be amendable to targeted therapies. This embodies the concepts of predictive enrichment and precision medicine. For example, all the endotyping strategies ostensibly identify a subgroup in whom dysregulation of the adaptive immune system is inferred based on the endotype-defining gene expression profile. If these endotyping strategies are truly identifying subgroups of patients with alterations of the adaptive immune system, it could have major implications for designing trials targeting modulation of the adaptive immune system among patients with sepsis. On a broader level, these endotyping strategies have the potential to manage the challenge of heterogeneity of treatment effect typically observed in interventional trials involving patients with septic shock.

Heterogeneity of treatment effect among patients with sepsis is perhaps best exemplified when considering the role of adjunctive corticosteroids for patients with septic shock, which has been intensely debated since the 1960s. Adjunctive corticosteroids are thought to have a potential therapeutic role in septic shock given their pluripotent effects on the cardiovascular system, inflammation, and immunity. Across multiple, large, randomized trials, adjunctive corticosteroids consistently reduce vasopressor requirements, but there has not been a consistently positive effect on patient survival. Recently, it has been suggested that gene expression-based endotyping could provide a means to better understand and manage the heterogeneity of treatment effect for adjunctive corticosteroids among patients with septic shock.

The Vasopressin vs. Norepinephrine as Initial Therapy in Septic Shock (VANISH) trial randomized adults with septic shock to hydrocortisone or placebo upon reaching a predefined vasopressor dose. A subset of the VANISH cohort also had corresponding transcriptomic data, and Antcliff et al. used these data to assign the cohort to the aforementioned SRS1 or SRS2 endotypes. They found that hydrocortisone use was independently associated with increased odds of mortality among patients assigned to the SRS2 endotype.

Among children with septic shock, observational data have consistently demonstrated that corticosteroid prescription is associated with increased odds of mortality among patients assigned to pediatric septic shock endotype A. This observation remains after adjusting for baseline illness severity, age, and co-morbidity burden. Recently, the publicly available transcriptomic data from the VANISH trial was used to cross-assign adults with septic shock to pediatric septic shock endotypes A or B. This study corroborated the association between corticosteroids and increased mortality risk among patients assigned to endotype A.

Another line of investigation used observational data to determine if there is a subgroup of children with septic shock that is more likely to benefit from adjunctive corticosteroids. This study combined PERSEVERE-based mortality risk stratification with endotyping and reported that among endotype B patients with an intermediate to the high baseline risk of mortality, adjunctive corticosteroids were associated with a more than ten-fold reduction of risk for poor outcome. This study, albeit based on observational data, illustrates the concept of how prognostic enrichment (PERSEVERE) and predictive enrichment (endotyping) can potentially be combined to identify a subset of patients more likely to have a beneficial response to adjunctive corticosteroids. The concept is currently being tested prospectively among adults (NCT04280497) and children (NCT03401398) with septic shock. A major next step for implementing these strategies at the bedside is the development of rapid assay platforms that can generate clinically actionable information in a clinically relevant timeframe.

CONCLUSION

Leveraging biomarkers to identify which patients have sepsis continues to be an important and well-recognized area of study. Equally important, but less well recognized, is the discovery and development of biomarkers for prognostic and predictive enrichment among patients with sepsis. This area has progressed considerably over the last few years in the pediatric domain. While a great deal of validation and development work remains, biomarkers for prognostic and predictive enrichment hold the promise of bringing precision medicine to the bedside of children with sepsis. This would include the enhancement of clinical trial design, the development of more targeted therapies, and bringing biological precision to quality improvement efforts.
ACKNOWLEDGEMENTS
This work was funded by the National Institute of General Medical Sciences, R35GM126943 (to H.R.W.).

COMPETING INTERESTS
H.R.W. and the Cincinnati Children’s Hospital Research Foundation hold US patents for the PERSEVERE biomarkers and the pediatric septic shock endotyping strategy described in the manuscript.

ADDITIONAL INFORMATION
Correspondence and requests for materials should be addressed to H.R.W.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.