Published in final edited form as:

Pediatr Crit Care Med. 2011 March ; 12(2): 165–173. doi:10.1097/PCC.0b013e3181e28876.

Biomarker discovery and development in pediatric critical care medicine

Jennifer M. Kaplan, M.D., M.S. and Hector R. Wong, M.D.

Division of Critical Care Medicine, Cincinnati Children's Hospital Medical Center and Cincinnati Children's Research Foundation. Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH

Abstract

Objective—To frame the general process of biomarker discovery and development, and to describe a proposal for the development of a multi-biomarker based risk model for pediatric septic shock.

Data Source—Narrative literature review and author generated data.

Main Results—Biomarkers can be grouped into four broad classes, based on the intended function: diagnostic, monitoring, surrogate, and stratification. Biomarker discovery and development requires a rigorous process, which is frequently not well followed in the critical care medicine literature. Very few biomarkers have successfully transitioned from the candidate stage to the true biomarker stage. There is great interest in developing diagnostic and stratification biomarkers for sepsis. Procalcitonin is currently the most promising diagnostic biomarker for sepsis. Recent evidence suggests that interleukin-8 can be used to stratify children with septic shock having a high likelihood of survival with standard care. Currently, there is a multi-institutional effort to develop a multi-biomarker based sepsis risk model intended to predict outcome and illness severity for individual children with septic shock.

Conclusions—Biomarker discovery and development is an important portion of the pediatric critical care medicine translational research agenda. This effort will require collaboration across multiple institutions and investigators. Rigorous conduct of biomarker-focused research holds the promise of transforming our ability to care for individual patients and our ability to conduct clinical trials in a more effective manner.

INTRODUCTION

In 2001 the National Institutes of Health convened a panel to develop standard definitions and a conceptual framework for biomarker discovery and development [1]. Based on this consensus conference a biomarker is broadly defined as a characteristic that can be objectively measured and evaluated as an indicator of normal biological processes, pathological processes, or pharmacological responses to a therapeutic intervention. Further, biomarkers can be subdivided into “Type 0” and “Type 1” biomarkers. A Type 0 biomarker provides a signal of the natural history of a disease that correlates longitudinally with known clinical indices and outcomes. A Type 1 biomarker captures the effects of a therapeutic intervention in accordance with its mechanism of action.

Based on this broad definition it is evident that we use biomarkers daily in the pediatric intensive care unit (PICU) without acknowledging them as such. For example, we routinely

measure serum creatinine as an indicator of kidney function or evolving kidney injury (a Type 0 biomarker). We also routinely measure serum glucose levels to monitor the biological effects of insulin therapy (a Type 1 biomarker). Other examples of everyday biomarkers used in the PICU include white blood cell counts in cerebral spinal fluid, urinary protein, and intravascular pressure measurements. However, the typical intensivist thinks of a “biomarker” as a novel “blood test” that provides us with patient information that is not readily obtainable using our current diagnostic and monitoring tools. This relatively narrow, but nonetheless highly relevant concept of biomarkers will provide the overall framework for this review. We will first provide a framework for the general discovery and development of biomarkers. Subsequently, we will illustrate several of the key points relating to biomarker discovery and development by describing a recent proposal for the derivation a multi-biomarker based risk model for pediatric septic shock.

BIOMARKER APPLICATION

Table 1 lists four broad functional classes of biomarkers as proposed in recent literature [2,3]. Diagnostic biomarkers serve to establish the presence or absence of a disease or other clinical conditions. The majority of diagnostic biomarkers are of most value within a specific clinical context and with the addition of adjunctive diagnostic tests. For example, myocardial enzymes (i.e. troponin, creatine phosphokinase) are most useful for accurately diagnosing myocardial infarction when there is well-founded clinical suspicion of myocardial infarction (context) and with the addition of other tests (e.g. electrocardiogram). Some diagnostic biomarkers are of most value for screening purposes. A screening biomarker, which is sub-type of a diagnostic biomarker, is intended to provide a simple, low cost, and low risk approach to check for the possibility of a disease in a population (e.g. heel sticks for metabolic disease screening in newborns, or stool guaiac testing for colon cancer screening in adults), rather than to firmly establish a given diagnosis. Ideal screening biomarkers have high sensitivity, without the need for high specificity. Therefore, a positive result derived from a screening biomarker typically requires follow-up with more definitive (i.e. specific) and often more complex diagnostic approaches. Conversely, ideal diagnostic biomarkers should have a high specificity, while the sensitivity becomes relatively less important. Sensitivity, specificity, and related concepts will be discussed further in a subsequent section.

Another important concept regarding diagnostic biomarkers, particularly in the field of critical care medicine, involves disease severity and the therapeutic options available for a given disease. The clinical significance of a diagnostic biomarker is directly proportional to disease severity and the therapeutic options for that disease. Thus, the most clinically important biomarkers in critical care medicine, and consequently what should be the focus of our translational research agenda, are those that diagnose or indicate risk for a severe illness or outcome for which there is an effective therapy.

Monitoring biomarkers serve to provide a “readout” that informs the clinician of the effectiveness of a given therapy for the purpose of titration. Robust monitoring biomarkers must therefore be functionally or biologically linked to the mechanism of action of a given therapeutic intervention. As previously mentioned, glucose monitoring in the context of insulin therapy is a relevant example of a monitoring biomarker routinely used in the PICU. Unfortunately, most of the monitoring biomarkers that we use in the PICU only provide indirect readouts of therapeutic efficacy, and consequently our titration capabilities are limited for many important therapies. For example, blood tacrolimus (FK-506) levels are routinely followed in solid organ transplant patients admitted to the PICU. The purpose of tacrolimus monitoring is two-fold: to minimize toxicity and to assess the adequacy of immune suppression. The latter purpose, which is singularly important in the management

of these patients, is only achieved indirectly. We infer “adequacy of immune suppression” based on a range of blood tacrolimus levels, without actually knowing the true state of immune activity/suppression. Thus, the field of pediatric critical care medicine needs further development of monitoring biomarkers that efficiently and feasibly measure bioactivity directly linked to therapeutic interventions.

Surrogate biomarkers serve a similar general function to that of monitoring biomarkers (i.e. they provide a readout of efficacy), but are intended to predict clinical outcome, rather than for titration of therapy [2,3]. As the name implies, surrogate biomarkers serve as proxy or substitute endpoints for patient-centered outcomes such as mortality or serious morbidities. The use of surrogate biomarkers has the theoretical advantage of lowering the number of subjects required to conduct a trial having sufficient power to detect efficacy. Robust surrogate biomarkers need to be directly affected by a given therapy, reliably predict the outcome of interest, and be mechanistically linked to the biological processes that affect the outcome of interest. The cardiology field has effectively used surrogate biomarkers (e.g. lipid profiles, c-reactive protein) in the conduct of clinical trials [4–6], but the development of surrogate biomarkers in the field of critical care medicine is a daunting challenge, which may not be completely feasible. Nonetheless, the development of effective surrogate biomarkers is required in the field of critical care medicine in order to better conduct Phase 1 and 2 clinical trials.

Stratification biomarkers (a.k.a. staging biomarkers) serve to sub-classify or stage diseases based on severity and/or outcome. Biomarker-based stratification can be applied to individual patients or in the context of a clinical trial. In the case of an individual patient, an effective stratification biomarker can identify higher risk patients that would be eligible for more toxic, more aggressive, or higher risk therapies (e.g. ECMO). Alternatively, an effective stratification biomarker can identify patients that are likely to have good outcomes with standard therapies, and thereby can objectively exclude patients from being exposed to unnecessary therapies. In either case, the general goal is to optimize the risk to benefit ratio of a given therapy. The oncology field is replete with excellent examples of effective therapies being predicated on data garnered from robust stratification biomarkers [6–9].

The field of critical care medicine would benefit tremendously from the development of effective biomarker-based stratification strategies. As intensivists we generally deal with syndromes rather than true “diseases” (e.g. septic shock and acute respiratory distress syndrome). Accordingly, the “diseases” that we encounter daily in the PICU are highly heterogeneous from both a clinical and biological standpoint. Relative to this high level of heterogeneity, however, our therapeutic approaches are highly homogenous. We generally apply the same therapeutic approaches to heterogeneous syndromes such as septic shock. We also conduct interventional clinical trials in the PICU without directly addressing the heterogeneous nature of the study cohort. Not surprisingly, the critical care medicine literature is replete with interventional clinical trials that have failed to demonstrate efficacy despite being based on sound preclinical data [10]. A major reason for these multiple failures is the inability to effectively manage and account for the profound heterogeneity of our patient populations, rather than fundamentally flawed biological principles. Thus, the development of robust stratification strategies is another important component of the critical care medicine research agenda, and biomarker-based strategies have great potential to directly address this agenda.

CANDIDATE BIOMARKER DISCOVERY

The biomarker field has become a highly competitive endeavor in the biomedical and pharmaceutical industries. The first step in biomarker development is discovery of “candidate” biomarkers [11]. As the name implies, a candidate biomarker has the potential to serve one of the aforementioned functions in the context of a given disease/syndrome, but remains to be formally tested and validated before being legitimately classified as a biomarker. The path from candidate biomarker status to actual biomarker status is very challenging and very few candidate biomarkers achieve actual biomarker status.

There are two broad approaches to candidate biomarker discovery: knowledge-based approaches and unbiased approaches. The knowledge-based approach to biomarker discovery embodies the traditional scientific method (i.e. hypothesis-driven) and has a long-standing history in all medical fields, including critical care medicine. In this approach, the investigator has a strong working knowledge of the pathophysiology of a given disease and related biological data. Based on this knowledge the investigator derives a hypothesis linking the candidate biomarker and the disease of interest, and subsequently designs a study to test the hypothesis. The critical care medicine literature is replete with studies in which “protein/molecule/metabolite X” (e.g. a cytokine) is demonstrated to be increased (or decreased) in the blood (or other biological fluid) of patients with “disease Y” (e.g. sepsis) [12–17]. While these studies are of value for initial identification of candidate biomarkers, the majority fall short of establishing a more definitive, clinically useful link between the candidate biomarker and the disease of interest. Simply reporting a correlation between a candidate biomarker level and a disease or disease outcome, is merely the first step in biomarker discovery.

The advantages of the knowledge-based approach are that it is grounded in the traditional scientific method and tends to be focused. This type of focused, hypothesis-driven approach is generally favorably regarded by review panels of funding agencies. The main disadvantage of the approach is that it is potentially biased, and is limited by the current state of knowledge in the field.

Unbiased approaches take a large-scale approach to candidate biomarker discovery without making few, if any, *a priori* decisions or assumptions regarding the potential candidates. This type of approach leverages the high throughput technologies related to the fields of proteomics, transcriptomics, and metabolomics, which allow for the simultaneous, semi-quantitative measurements of thousands of proteins, gene transcripts, or metabolites, respectively, in biological samples. Accordingly, these platforms provide unprecedented opportunities to compare differential expression of proteins, gene transcripts, or metabolites between biological samples representing disease states, gradations of disease states, and normal states. Human blood, a primary target for biomarker discovery and development, has been described as a highly comprehensive and readily accessible proteome potentially providing a representation of all body tissues during health and disease [18].

As the name implies, the primary advantage of these large-scale approaches is that they are relatively unbiased in that no particular protein, transcript, or metabolite becomes the focus of investigation. In other words, all proteins, genes, or metabolites “have a chance” to be discovered as candidate biomarkers through the development of comprehensive biomolecular signatures of disease states. This type of approach is particularly well-suited for the discovery phase of biomarker development, rather than the latter stages of derivation and validation. The main disadvantages of these approaches are that they are costly, can have limited detection capabilities, and that the large data sets present unprecedented analytical challenges requiring complex bioinformatics prone to generation of false positives and false

negatives. In addition, some funding agency review panels view these approaches as non-hypothesis driven “fishing expeditions.”

Another important component of biomarker discovery is the biological material one chooses to use for analysis. Since the biomarker field is, in large part, focused on the development of “blood tests,” blood is a logical source for the discovery phase of candidate biomarkers. However, other biological materials (“proximal fluids”) may be more appropriate in the discovery phase for several reasons including, proximity to the diseased tissue of interest, enrichment for disease-specific candidate biomarkers, and less complex backgrounds [11]. Examples of human proximal fluids include urine, cerebral spinal fluid, ascites, and bronchoalveolar lavage fluid. Proximal fluids can also be derived from appropriate animal models and cell culture models during the initial discovery phase.

Neutrophil gelatinase-associated lipocalin (NGAL) is now recognized as a robust biomarker for acute kidney injury (AKI) in critically ill patients, including children [19–21]. The discovery of NGAL as a candidate biomarker for AKI well illustrates the above concepts of analyzing a proximal fluid (urine) derived from an animal model, and using an unbiased approach (transcriptomics and proteomics). NGAL was initially identified as candidate AKI biomarker in rodent models of kidney ischemia [22,23]. Analyses of the kidney parenchymal transcriptome and the urine proteome demonstrated that NGAL was one of the most abundant genes expressed in rodents subjected to experimental renal ischemia. The use of kidney tissue and urine as the biological materials was a key component of the discovery phase in that they directly represent or are in close proximity to the tissue of interest (i.e. the kidney), are therefore likely to be enriched for kidney-specific candidate biomarkers, and the urine proteome is several orders of magnitude less complex than the blood proteome. Thus, an unbiased approach based on biological samples from an experimental animal model allowed the discovery of a candidate biomarker (i.e. NGAL) that may have not been readily evident using a knowledge-based approach.

BIOMARKER DERIVATION AND VALIDATION

After the candidate biomarker discovery phase, the next phase of biomarker development involves derivation and validation. The derivation phase involves the establishment of “cut-off” values or ranges at which the biomarker will provide clinically useful information. For example, in the case of a diagnostic biomarker a cut-off value allows one to determine whether or not the disease of interest is present (i.e. sepsis). In the case of a stratification biomarker, a cut-off value allows one to predict a particular outcome (i.e. mortality).

The derivation phase should typically involve a relatively large cohort of study subjects with and without the disease or outcome of interest. Other commonly used terms for derivation cohorts are “training” cohorts and “learning” cohorts. After biomarker cut-off values/ranges are established in the derivation cohort, the performance of these cut-off values/ranges needs to be formally tested in a validation cohort of patients. The ideal validation cohort is also large and completely independent of the original derivation cohort. Unfortunately, a substantial proportion of the critical care medicine literature involving candidate biomarkers does not effectively transition from initial derivation to subsequent formal validation.

The derivation and validation phases of biomarker development serve to determine the validity of the biomarker in providing the correct answer of interest. To determine validity there needs to be an established method or readout to indicate the correct answer (i.e. a gold standard). Some diseases, but not all, have a diagnostic gold standard. For example, isolation of a known pathogen in cerebral spinal fluid is the diagnostic gold standard for meningitis. Many studies in sepsis use a positive blood culture as the gold standard, however false negative culture results may occur from sampling error and false positive results may occur

from contamination. For a particular disease the gold standard may change over time based on development of newer diagnostic methods. For example, to diagnose a pulmonary embolism the gold standard test historically was pulmonary angiography. This test has now been replaced by helical computed tomography pulmonary angiography. Therefore tests of validity, including sensitivity and specificity, must be interpreted in the context of the clinical situation and the available gold standards.

In the following paragraphs we will outline common approaches and concepts for determining biomarker validity, also known as test performance characteristics. The approaches described primarily apply to single biomarker strategies. More complex approaches, such as principle component analysis and multivariable logistic regression, are required for multi-biomarker-based approaches.

Sensitivity and Specificity

Sensitivity and specificity calculations provide information regarding intrinsic test validity and are not dependent on the prevalence of a disease or condition. As such, specificity and sensitivity are said to be “fixed” for a given test. In order to calculate sensitivity and specificity data must be presented as dichotomous outcomes. For analysis of survival data this is straightforward (dead vs. not dead). For diagnostic testing, however, a cutoff value is necessary to distinguish the two groups (e.g. glucose levels ≥ 200 mg/dl vs. glucose levels < 200 mg/dl), and decisions involving cut-off values directly affect sensitivity and specificity calculations. Sensitivity and specificity testing involve administering the test to two groups: patients with the disease and another group without the disease. Once the testing is performed the results can be placed in a 2×2 contingency table and measures of test accuracy can be performed (Figure 1). Sensitivity is defined as the proportion of subjects with the disease in whom the test gives a positive result (e.g., patients with bacteremia who have a positive blood culture), whereas specificity is the proportion of subjects without the disease in whom the test gives a negative result (e.g., patients without bacteremia who have a negative blood culture). A highly sensitive test has a low false-negative rate and a highly specific test has a low false-positive rate. Therefore, tests that have a high sensitivity are often used as screening tests because they are helpful for ruling out a diagnosis in patients when the test is negative. Tests with high specificity are used to rule in a disease where a positive test indicates a high probability that the patient has the disease.

Positive and negative predictive values

Sensitivity and specificity define the operating characteristics of a given test. Most of the time, however, a clinician’s interest lies in knowing what to do with a test result. Herein lays the utility of positive predictive values (PPV) and negative predictive values (NPV). Predictive values require an index of suspicion of the disease and are affected by the prevalence of the disease in the population. A PPV assesses the validity of a positive test, and puts test specificity in the context of disease prevalence. A high disease prevalence usually results in high PPV, whereas low disease prevalence usually results in low PPV. NPV assesses the validity of a negative test, and puts test sensitivity in the context of disease prevalence. Low disease prevalence usually results in high NPV.

PPV and NPV can also be calculated using a 2×2 table (Figure 1). PPV reflects the proportion of time that a patient with a positive diagnostic test result has the disease (e.g. the probability of bacteremia, given a positive blood culture). NPV reflects the proportion of time that a patient with a negative diagnostic test result does not have the disease (e.g. the probability of not having bacteremia, given a negative blood culture).

Likelihood Ratios

The likelihood ratio (LR) is based on odds rather than probabilities, simultaneously takes sensitivity and specificity into account, and is not dependent on prevalence. A positive LR (+LR) expresses the odds that a positive test result will be observed in patients with the disease versus the odds that the same result will be observed in patients without the disease: sensitivity/(1-specificity). A negative LR (−LR) expresses the odds that a negative test result will be observed in patients with the disease versus the odds that the same result will be observed in patients without the disease: (1-sensitivity)/specificity. An LR of 1 indicates that the test result is just as likely to occur in patients with the disease as in patients without the disease. Clinically useful +LRs are typically >10, whereas clinically useful −LRs are typically <0.1.

Receiver Operating Characteristic Curves

When tests are measured on a continuum, sensitivity and specificity levels depend on where the cutoff is set between positive and negative results. The relationship between sensitivity and specificity can be graphically displayed using a receiver operating characteristic (ROC) curve (Figure 2). Cut-off points are selected and the sensitivity and specificity are determined at each point. The resulting graph depicts the Y-axis as the true-positive rate (sensitivity) as a function of the false-positive rate (1-specificity) on the x-axis. The proximity of the curve to the upper left-hand corner of the graph corresponds to increasing accuracy (i.e. the true positive rate is closer to 1 and the false-positive rate is closer to 0). The overall accuracy of a test can be demonstrated as the area under the ROC curve (AUC). An AUC of 0.5 indicates that the test is positive or negative by chance alone, whereas the “perfect” test would yield an AUC of 1.

Confidence Intervals

Confidence intervals refer to a range of values about the sample mean or proportion. Confidence intervals reflect the range of values in which the means of different samples drawn from a given population are likely to fall. It is a measure of precision; therefore a narrow confidence interval is more precise than a larger confidence interval. Commonly reported confidence intervals are 95% and 99%. A 95% confidence interval can be interpreted to indicate that 95% of the sample means drawn from a population will fall within the reported interval. Therefore, narrow confidence intervals provide data with a higher level of validity. Large patient cohorts are more likely to yield narrow confidence intervals, whereas small patient cohorts are more likely to yield wide confidence intervals. Confidence intervals should ideally be reported for all of the indices discussed above, including LR. For ROC curves, interpretation of an AUC with a confidence interval including 0.5 would suggest that the test could be positive or negative by random chance.

SEPSIS-RELATED BIOMARKERS

Since sepsis and septic shock are such important conditions to the daily clinical practice of critical care medicine, there has been a great deal of interest in developing sepsis-related biomarkers. All four classes of biomarkers (diagnostic, monitoring, surrogate, and stratification) have potential applications in sepsis. For example, a monitoring biomarker (e.g. lactate) could assist in titrating cardiovascular support. Alternatively, the development of surrogate biomarkers for sepsis could potentially be a major advancement in the conduct of clinical trials, which currently require very large sample sizes to detect a survival benefit. The most well developed areas thus far, however, relate to diagnostic and stratification biomarkers for sepsis. Accordingly, in the following section we will discuss diagnostic and stratification biomarkers for sepsis in greater detail. The discussion will also help to highlight many of the key concepts described in previous sections.

DIAGNOSTIC BIOMARKERS FOR SEPSIS

Procalcitonin

A daily conundrum in the ICU is reliably distinguishing which patients meeting criteria for systemic inflammatory response syndrome (SIRS) are infected and which are not infected. Accordingly, there have been extensive efforts to develop diagnostic biomarkers for sepsis/infection in critically ill patients. Procalcitonin (PCT) has emerged as the most promising of these diagnostic biomarkers.

PCT is a 116-amino-acid peptide and a precursor of calcitonin [24,25]. Low levels of PCT circulate in the serum, but are substantially increased during systemic infections [26]. PCT is normally secreted from the thyroid parafollicular or clear (C) cells, however in sepsis PCT can originate from extra-thyroidal sites including liver, spleen, and adipose tissue [27,28]. PCT can be detected 2–3 hours after an injection of endotoxin in normal human volunteers [29].

In critically ill patients PCT has consistently demonstrated greater accuracy than C-reactive protein (CRP) in determining sepsis, but demonstrates a wide range of sensitivity and specificity across individual studies [30–34]. Trauma patients who developed sepsis had higher PCT levels on admission compared with trauma patients who did not develop sepsis [35]. Using a cutoff value of 1.09 ng/ml and positive blood cultures as the gold standard, PCT had a high sensitivity (100%) and negative predictive value (100%) for diagnosing sepsis, but low specificity and positive predictive value, and no mortality was associated with “missed” infections. In a meta-analysis focused on adults admitted to intensive care units after surgery or trauma, PCT levels in infected patients were about 16 times higher than in patients without infection [31]. However, a separate meta-analysis including a more generalized population of patients in intensive care units, emergency departments (ED), or hospital wards concluded that PCT cannot accurately distinguish sepsis from SIRS [36].

PCT may have value in diagnosing and staging sepsis in the ED. In ED patients, fever, chills, and PCT levels were the strongest predictors for the diagnosis of bacterial infection [37]. Using a rapid semi-quantitative method to determine PCT levels Oh et al. demonstrated that in patients with suspected infection, who presented to the ED and met SIRS criteria, a cutoff value of >2 ng/ml was useful for distinguishing between SIRS/sepsis and severe sepsis/septic shock [37].

In the pediatric population PCT also proves to be a better diagnostic marker for sepsis than CRP [33,38,39]. In ROC curve analysis, PCT was better than CRP or the immature to total neutrophil ratio in diagnosing sepsis [32]. In infants < 3 months of age with fever of unknown origin, PCT has a higher diagnostic value than CRP or leukocyte count in predicting serious bacterial infections [40]. Additionally, Han et al. demonstrated that PCT levels were significantly increased in patients with bacterial sepsis compared to controls, but PCT levels were not significantly increased in patients with viral, fungal, or culture negative sepsis compared with controls [41].

The most important clinical applicability of PCT as a biomarker may be in determining the initiation (diagnostic biomarker) and duration (monitoring biomarker) of antibiotic use. In a randomized control trial patients were divided into a PCT-guided treatment plan or a standard antibiotic treatment plan [42]. In the PCT group, antibiotics were discontinued if clinical signs and symptoms of the infection had improved and if the PCT level decreased. The duration of antibiotic treatment in the PCT-guided group was 5.9 ± 1.7 days shorter than in the control group ($p < 0.001$). Further support of the use of PCT in guiding antibiotic therapy was demonstrated in a large, multicenter, randomized trial in patients with lower

respiratory tract infections [43]. The use of a PCT-guided treatment algorithm resulted in lower rates of antibiotic exposure and antibiotic-associated adverse effects compared with standard guidelines. Similar findings were recently published in another randomized, PCT-guided algorithm trial [44].

Interleukin-18 (IL-18)

IL-18 is also gaining attention as a potential diagnostic biomarker for sepsis, but the overall data are currently not as robust as that for PCT. IL-18 is a pro-inflammatory cytokine and a member of the IL-1 super family that activates NF- κ B and MAP kinases [45]. IL-18 is produced from a precursor protein and is cleaved to a mature form by caspase-1. Increased IL-18 levels are found in many human inflammatory conditions including rheumatoid arthritis, neonatal infections, and sepsis [46–48].

Plasma IL-18 levels were significantly increased in patients with sepsis (588 ± 83 pg/ml) compared to trauma patients (182 ± 13 pg/ml) on day of admission or at time of sepsis diagnosis [49]. In this same cohort, patients with septic shock had higher IL-18 plasma levels compared to patients without shock. IL-18 plasma levels were also elevated in patients who developed sepsis following major visceral surgery compared to patients who did not develop sepsis [50]. Furthermore, non-survivors with sepsis had significantly higher IL-18 levels compared to survivors from sepsis and IL-18 levels continuously increased during sepsis in non-survivors. Other data suggest that IL-18 may allow for distinction between gram positive and gram negative bacterial infections in critically ill patients [45,51].

In summary, the role of IL-18 as a diagnostic biomarker for sepsis remains to be further developed, and limited data exist in the pediatric population (thus providing a potential area for translational research). In contrast, there is considerably better developed data indicating that IL-18 may be a robust diagnostic biomarker for AKI [52–54].

STRATIFICATION BIOMARKERS FOR SEPTIC SHOCK

Rationale

Septic shock in humans, as opposed to animal models, is a heterogeneous syndrome having highly variable physiological and biological manifestations in a given patient population [10]. This heterogeneity continues to pose a major challenge to the rational conduct of clinical trials focused on septic shock. Indeed, the vast majority of clinical trials targeted at improving the outcome of septic shock have failed despite being founded on quality pre-clinical data. One major reason for these multiple failures appears to be the inability to manage the inherent heterogeneity of clinical septic shock in the context of a clinical trial. Accordingly, it has been proposed by Marshall that a key challenge in the field is to reduce and manage this heterogeneity by more effectively stratifying patients for the purpose of more effective clinical research [10].

Interleukin-8 (IL-8) as a stratification biomarker in pediatric septic shock

We recently reported that IL-8 can reliably predict survival in critically ill children with septic shock receiving standard care [55]. The identification, derivation, and validation of IL-8 as a stratification biomarker illustrates many of the concepts described in the previous sections.

The initial discovery of IL-8 as a candidate stratification biomarker in pediatric septic shock involved a hybrid approach, which was partially unbiased and partially knowledge-based. The unbiased approach consisted of genome-wide analysis of gene expression

(transcriptomics) using whole blood-derived RNA obtained within 24 hours of admission to the PICU. These microarray-based studies interrogated over 30,000 genes simultaneously, and revealed that IL-8 was one of the most highly expressed genes in non-survivors of pediatric septic shock compared to survivors [56]. Because IL-8 is a well known chemo-attractant and activator of neutrophils, we hypothesized that serum IL-8 protein levels could be used as a stratification biomarker in pediatric septic shock (knowledge-based approach).

We then measured serum IL-8 protein levels in the same cohort of patients used for the microarray studies ($n = 42$, training data set) to construct an ROC curve for 28-day mortality, which yielded an AUC of 0.857. From these procedures we derived an IL-8 cut-off level of >220 pg/ml as having a 75% sensitivity and specificity for predicting 28-day mortality in the training data set. We next applied this cut-off value to a validation data set of 139 children with septic shock and constructed a contingency table to calculate the sensitivity, specificity, positive predictive value, and negative predictive value of this IL-8 cut-off for predicting 28-day mortality. The calculated sensitivity, specificity, and positive predictive value were not sufficiently robust for clinical application. However, the negative predictive value for mortality was 95%, with narrow 95% confidence intervals (87 to 98%), and a negative likelihood ratio for mortality of 0.3. In other words, these data suggested that a serum IL-8 level ≤ 220 pg/ml, measured within 24 hours of admission to the PICU, may have the ability to predict survival in children with septic shock with 95% probability. Subsequently, we tested the veracity of this negative predictive value in a second validation data set, completely independent of our own database: children enrolled in the Phase 3 trial of activated protein C (RESOLVE Trial) [57]. Application of the IL-8 cutoff to the RESOLVE database demonstrated an equally robust negative predictive value and negative likelihood ratio for 28-day mortality.

Based on these data, we have proposed that serum IL-8 measurements, conducted within the first 24 hours of admission to the PICU, can be used to identify children with a high probability of survival with standard care (i.e. stratification) [55]. These low risk patients could then be excluded from interventional clinical trials. By excluding patients in this manner, one could improve the risk to benefit ratio of an experimental intervention by generating a study cohort that would be more likely to realize a benefit from the experimental intervention, particularly if the intervention carries more than minimal risk for serious adverse events. Using a similar approach as that described above, we also recently demonstrated that chemokine (C-C motif) ligand 4 (CCL4, a.k.a. MIP-1 β) has a 98% negative predictive value for mortality in children with septic shock [58].

A multi-biomarker-based risk model for pediatric septic shock stratification

The strengths of the individual IL-8- and CCL4-based stratification strategies described above are simplicity and very high negative predictive values for mortality. However, both stratification biomarkers have positive predictive values, sensitivities, and specificities that are too low to develop a comprehensive pediatric septic shock stratification tool meeting a wide variety of clinical and research needs [55,58]. Given the complexity of septic shock, it is possible that a multi-biomarker based stratification system could fulfill all of these needs. Herein we describe our plan to derive the pediatric sepsis biomarker risk model (heretofore called PERSEVERE: PEdiatRic SEpsis biomarkEr Risk modEl) using 15 candidate stratification biomarkers.

The candidate biomarkers for derivation of PERSEVERE were first identified using an unbiased approach that leveraged the discovery potential of a well-annotated microarray data set (transcriptomics) consisting of 98 children with septic shock and representing the first 24 hours of admission to the PICU [59]. We employed two complementary but distinct unbiased approaches to the discovery of candidate biomarker genes. In the first approach we

used normal controls, septic shock survivors, and septic shock non-survivors as the comparison groups, and performed a 3 group ANOVA using all gene probes on the array (87,933 gene probes). The ANOVA consisted of a Welch test with correction for multiple comparisons via a Benjamini-Hochberg False Discovery Rate of 5%. This test yielded > 20,000 differentially regulated gene probes between the 3 comparison groups. We then performed a *post-hoc* Student-Newman-Keuls test to determine the specific inter-group differences in gene expression. This *post-hoc* test yielded 137 gene probes differentially regulated between septic shock survivors and non-survivors. This 137 candidate biomarker gene list will heretofore be called “gene list A” for ease of reading.

In the second unbiased approach to candidate biomarker discovery, we applied Support Vector Machines-based dichotomous class prediction modeling to identify candidate biomarker genes [60]. Starting with all genes on the array (87,933) we attempted to predict “survivor” and “non-survivor” classes via leave-one-out cross validation procedures. This class prediction modeling approach was able to correctly predict survival or non-survival in 84 of the 98 patients (86% correct class prediction). More specifically, the model correctly predicted 15 of the 17 non-survivors (88%) and 69 of the 81 survivors (85%). Further, we used the Fisher test for gene selection, and extracted the top 5% class predictor gene probes out of the starting 87,933 gene probes (4,397 gene probes). This 4,397 candidate biomarker gene list will heretofore be called “gene list B” for ease of reading.

Having derived candidate biomarker gene lists A and B by two distinct methods, we next conducted Venn analysis to determine which genes are common to the two gene lists. As shown in Figure 3, we found 117 gene probes common to gene lists A and B. To derive a final list of candidate biomarker genes we examined the above 117 gene list for genes meeting 2 simultaneous, *a priori* criteria (i.e. a knowledge-based approach): 1) the gene has a reasonable level of biological plausibility with regard to the pathobiology of pediatric septic shock, the host response to infection, and/or the host inflammatory response; and 2) the gene product (protein) can be feasibly measured in the blood compartment. Based on these 2 criteria we derived a final working list of 15 candidate biomarker genes as shown in Table 2.

We propose that the candidate biomarker list shown in Table 2 provides a robust foundation for derivation of PERSEVERE. Several features of the biomarker list support the assertion of robustness. First, the patient cohort from which the genes were derived represents 11 major pediatric centers in the United States (see acknowledgment section), thus taking into account any potential variability in standard care and thereby speaking to the potential generalizability of the candidate biomarkers. Second, the initial working gene lists were generated in a systematic, objective, and rigorous manner, based on a combined statistical approach and a class prediction approach. Third, the selection process has limited, if any, bias. The selection process began with the entire probe set on the array. Thus, all genes “had a chance” to be selected. The final step in selecting the candidate biomarkers (i.e. a knowledge-based approach) is the only potential source of bias. Finally, the microarray data from which the genes were selected represent the first 24 hours of presentation to the PICU. This early time point is ideal and clinically relevant for the goals of PERSEVERE.

Robust illness severity scores are currently available for critically ill pediatric populations, but it is clear that these scores are not appropriate tools to stratify individual patients for the purposes of clinical trials, nor for making individual patient decisions [61]. PERSEVERE is intended to predict outcome and illness severity for individual children with septic shock. The 15 candidate biomarkers will be quantified in serum samples using a multi-plex platform. PERSEVERE will be derived in a formal derivation cohort of patients using statistical modeling approaches, which integrate the candidate biomarker measurements with

principal component analysis and multivariable logistic regression. The final output will be an equation into which an individual patient's biomarker measurements can be input to derive an individual risk assessment. Subsequently, PERSEVERE will be validated prospectively in a separate validation cohort. If PERSEVERE comes to fruition, we expect that it will provide an unprecedented decision and stratification tool for the care of individual children with septic shock and for the conduct of interventional clinical trials.

CONCLUDING PERSPECTIVES

Biomarkers for diagnosing, monitoring, and stratifying various forms of critical illness hold the promise of dramatically changing how we practice critical care medicine on a daily basis, and how we conduct clinical trials in the future [2]. Realizing this promise, however, requires rigorous methodologies, validation procedures, and thoughtful application of high throughput technologies. In addition, realizing this promise will require collaboration across many centers and investigators. In our opinion, it is unlikely that biomarker discovery and development specific to the field of pediatric critical care medicine will be a priority for the bio-technology and pharmaceutical industries. Accordingly, clinician-scientists based in pediatric critical care medicine will need to lead biomarker discovery and development specific to the field.

Acknowledgments

Supported by the National Institutes of Health: R01GM064619, RC1HL100474, and K12HD028827

Contributing investigators and centers for genomic database that generated the candidate biomarkers described for PERSEVERE: Thomas P. Shanley (C.S. Mott Children's Hospital at the University of Michigan, Ann Arbor, Michigan); Natalie Cvijanovich (Children's Hospital and Research Center Oakland, Oakland, CA); Richard Lin (The Children's Hospital of Philadelphia, Philadelphia, PA); Geoffrey L. Allen (Children's Mercy Hospital, Kansas City, MO); Neal J. Thomas (Penn State Children's Hospital, Hershey, PA); Douglas F. Willson (University of Virginia, Charlottesville, VA); Robert J. Freishtat (Children's National Medical Center, Washington, D.C.); Nick Anas (Children's Hospital of Orange County, Orange, CA); Keith Meyer (Miami Children's Hospital, Miami, FL); and Paul Checchia (St. Louis Children's Hospital, St. Louis, MO)

References

1. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001; 69(3):89–95. [PubMed: 11240971]
2. Marshall JC, Reinhart K. Biomarkers of sepsis. Crit Care Med. 2009; 37(7):2290–2298. [PubMed: 19487943]
3. Puntmann VO. How-to guide on biomarkers: biomarker definitions, validation and applications with examples from cardiovascular disease. Postgrad Med J. 2009; 85(1008):538–545. [PubMed: 19789193]
4. Ridker PM, Rifai N, Rose L, Buring JE, Cook NR. Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. N Engl J Med. 2002; 347(20):1557–1565. [PubMed: 12432042]
5. Ridker PM, Cannon CP, Morrow D, Rifai N, Rose LM, McCabe CH, Pfeffer MA, Braunwald E. C-reactive protein levels and outcomes after statin therapy. N Engl J Med. 2005; 352(1):20–28. [PubMed: 15635109]
6. Ridker PM, Morrow DA, Rose LM, Rifai N, Cannon CP, Braunwald E. Relative efficacy of atorvastatin 80 mg and pravastatin 40 mg in achieving the dual goals of low-density lipoprotein cholesterol <70 mg/dl and C-reactive protein <2 mg/l: an analysis of the PROVE-IT TIMI-22 trial. J Am Coll Cardiol. 2005; 45(10):1644–1648. [PubMed: 15893181]
7. Watanabe T, Wu TT, Catalano PJ, Ueki T, Satriano R, Haller DG, Benson AB 3rd, Hamilton SR. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. N Engl J Med. 2001; 344(16):1196–1206. [PubMed: 11309634]

8. Sturgeon CM, Duffy MJ, Stenman UH, Lilja H, Brunner N, Chan DW, Babaian R, Bast RC Jr, Dowell B, Esteva FJ, et al. National Academy of Clinical Biochemistry laboratory medicine practice guidelines for use of tumor markers in testicular, prostate, colorectal, breast, and ovarian cancers. *Clin Chem*. 2008; 54(12):e11–79. [PubMed: 19042984]
9. Duffy MJ, Crown J. A personalized approach to cancer treatment: how biomarkers can help. *Clin Chem*. 2008; 54(11):1770–1779. [PubMed: 18801934]
10. Marshall JC. Sepsis: rethinking the approach to clinical research. *J Leukoc Boil*. 2008; 83(3):471–482.
11. Rifai N, Gillette MA, Carr SA. Protein biomarker discovery and validation: the long and uncertain path to clinical utility. *Nat Biotechnol*. 2006; 24(8):971–983. [PubMed: 16900146]
12. Wheeler DS, Devarajan P, Ma Q, Harmon K, Monaco M, Cvijanovich N, Wong HR. Serum neutrophil gelatinase-associated lipocalin (NGAL) as a marker of acute kidney injury in critically ill children with septic shock. *Crit Care Med*. 2008; 36(4):1297–1303. [PubMed: 18379258]
13. Giuliano JS Jr, Lahni PM, Harmon K, Wong HR, Doughty LA, Carcillo JA, Zingarelli B, Sukhatme VP, Parikh SM, Wheeler DS. Admission angiopoietin levels in children with septic shock. *Shock*. 2007; 28(6):650–654. [PubMed: 18092380]
14. Wheeler DS, Lahni P, Odoms K, Jacobs BR, Carcillo JA, Doughty LA, Wong HR. Extracellular heat shock protein 60 (Hsp60) levels in children with septic shock. *Inflamm Res*. 2007; 56(5):216–219. [PubMed: 17588138]
15. Wheeler DS, Fisher LE Jr, Catravas JD, Jacobs BR, Carcillo JA, Wong HR. Extracellular hsp70 levels in children with septic shock. *Pediatr Crit Care Med*. 2005; 6(3):308–311. [PubMed: 15857530]
16. Wong HR, Doughty LA, Wedel N, White M, Nelson BJ, Havrilla N, Carcillo JA. Plasma bactericidal/permeability-increasing protein concentrations in critically ill children with the sepsis syndrome. *Pediatr Infect Dis J*. 1995; 14(12):1087–1091. [PubMed: 8745023]
17. Wong HR, Carcillo JA, Burckart G, Shah N, Janosky JE. Increased serum nitrite and nitrate concentrations in children with the sepsis syndrome. *Crit Care Med*. 1995; 23(5):835–842. [PubMed: 7736740]
18. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics*. 2002; 1(11):845–867. [PubMed: 12488461]
19. Haase M, Bellomo R, Devarajan P, Schlattmann P, Haase-Fielitz A. Accuracy of neutrophil gelatinase-associated lipocalin (NGAL) in diagnosis and prognosis in acute kidney injury: a systematic review and meta-analysis. *Am J Kidney Dis*. 2009; 54(6):1012–1024. [PubMed: 19850388]
20. Devarajan P. Neutrophil gelatinase-associated lipocalin (NGAL): a new marker of kidney disease. *Scand J Clin Lab Invest Suppl*. 2008; 241:89–94. [PubMed: 18569973]
21. Mishra J, Dent C, Tarabishi R, Mitsnefes MM, Ma Q, Kelly C, Ruff SM, Zahedi K, Shao M, Bean J, et al. Neutrophil gelatinase-associated lipocalin (NGAL) as a biomarker for acute renal injury after cardiac surgery. *Lancet*. 2005; 365(9466):1231–1238. [PubMed: 15811456]
22. Mishra J, Ma Q, Prada A, Mitsnefes M, Zahedi K, Yang J, Barasch J, Devarajan P. Identification of neutrophil gelatinase-associated lipocalin as a novel early urinary biomarker for ischemic renal injury. *J Am Soc Nephrol*. 2003; 14(10):2534–2543. [PubMed: 14514731]
23. Devarajan P. Proteomics for biomarker discovery in acute kidney injury. *Semin Nephrol*. 2007; 27(6):637–651. [PubMed: 18061846]
24. Snider RH Jr, Nylen ES, Becker KL. Procalcitonin and its component peptides in systemic inflammation: immunochemical characterization. *J Investig Med*. 1997; 45(9):552–560.
25. Becker KL, Snider R, Nylen ES. Procalcitonin in sepsis and systemic inflammation: a harmful biomarker and a therapeutic target. *Br J Pharmacol*. 2009
26. Assicot M, Gendrel D, Carsin H, Raymond J, Guilbaud J, Bohuon C. High serum procalcitonin concentrations in patients with sepsis and infection. *Lancet*. 1993; 341(8844):515–518. [PubMed: 8094770]
27. Linscheid P, Seboek D, Nylen ES, Langer I, Schlatter M, Becker KL, Keller U, Muller B. In vitro and in vivo calcitonin I gene expression in parenchymal cells: a novel product of human adipose tissue. *Endocrinology*. 2003; 144(12):5578–5584. [PubMed: 12960010]

28. Muller B, White JC, Nylen ES, Snider RH, Becker KL, Habener JF. Ubiquitous expression of the calcitonin-i gene in multiple tissues in response to sepsis. *J Clin Endocrinol Metab.* 2001; 86(1): 396–404. [PubMed: 11232031]
29. Dandona P, Nix D, Wilson MF, Aljada A, Love J, Assicot M, Bohuon C. Procalcitonin increase after endotoxin injection in normal subjects. *J Clin Endocrinol Metab.* 1994; 79(6):1605–1608. [PubMed: 7989463]
30. Brunkhorst FM, Eberhard OK, Brunkhorst R. Discrimination of infectious and noninfectious causes of early acute respiratory distress syndrome by procalcitonin. *Crit Care Med.* 1999; 27(10): 2172–2176. [PubMed: 10548201]
31. Uzzan B, Cohen R, Nicolas P, Cucherat M, Perret GY. Procalcitonin as a diagnostic test for sepsis in critically ill adults and after surgery or trauma: a systematic review and meta-analysis. *Crit Care Med.* 2006; 34(7):1996–2003. [PubMed: 16715031]
32. McMaster P, Park DY, Shann F, Cochrane A, Morris K, Gray J, Cottrell S, Belcher J. Procalcitonin versus C-reactive protein and immature-to-total neutrophil ratio as markers of infection after cardiopulmonary bypass in children. *Pediatr Crit Care Med.* 2009; 10(2):217–221. [PubMed: 19057447]
33. Simon L, Saint-Louis P, Amre DK, Lacroix J, Gauvin F. Procalcitonin and C-reactive protein as markers of bacterial infection in critically ill children at onset of systemic inflammatory response syndrome. *Pediatr Crit Care Med.* 2008; 9(4):407–413. [PubMed: 18496408]
34. Simon L, Gauvin F, Amre DK, Saint-Louis P, Lacroix J. Serum procalcitonin and C-reactive protein levels as markers of bacterial infection: a systematic review and meta-analysis. *Clin Infect Dis.* 2004; 39(2):206–217. [PubMed: 15307030]
35. Castelli GP, Pognani C, Cita M, Paladini R. Procalcitonin as a prognostic and diagnostic tool for septic complications after major trauma. *Crit Care Med.* 2009; 37(6):1845–1849. [PubMed: 19384224]
36. Tang BM, Eslick GD, Craig JC, McLean AS. Accuracy of procalcitonin for sepsis diagnosis in critically ill patients: systematic review and meta-analysis. *Lancet Infect Dis.* 2007; 7(3):210–217. [PubMed: 17317602]
37. de Kruif MD, Limper M, Gerritsen H, Spek CA, Brandjes DP, ten Cate H, Bossuyt PM, Reitsma PH, van Gorp EC. Additional value of procalcitonin for diagnosis of infection in patients with fever at the emergency department. *Crit Care Med.* 38(2):457–463. [PubMed: 20083920]
38. Rey C, Los Arcos M, Concha A, Medina A, Prieto S, Martinez P, Prieto B. Procalcitonin and C-reactive protein as markers of systemic inflammatory response syndrome severity in critically ill children. *Intensive Care Med.* 2007; 33(3):477–484. [PubMed: 17260130]
39. Enguix A, Rey C, Concha A, Medina A, Coto D, Dieguez MA. Comparison of procalcitonin with C-reactive protein and serum amyloid for the early diagnosis of bacterial sepsis in critically ill neonates and children. *Intensive Care Med.* 2001; 27(1):211–215. [PubMed: 11280637]
40. Olaciregui I, Hernandez U, Munoz JA, Emparanza JI, Landa JJ. Markers that predict serious bacterial infection in infants under 3 months of age presenting with fever of unknown origin. *Arch Dis Child.* 2009; 94(7):501–505. [PubMed: 19158133]
41. Han YY, Doughty LA, Kofos D, Sasser H, Carcillo JA. Procalcitonin is persistently increased among children with poor outcome from bacterial sepsis. *Pediatr Crit Care Med.* 2003; 4(1):21–25. [PubMed: 12656537]
42. Hochreiter M, Kohler T, Schweiger AM, Keck FS, Bein B, von Spiegel T, Schroeder S. Procalcitonin to guide duration of antibiotic therapy in intensive care patients: a randomized prospective controlled trial. *Crit Care.* 2009; 13(3):R83. [PubMed: 19493352]
43. Schuetz P, Christ-Crain M, Thomann R, Falconnier C, Wolbers M, Widmer I, Neidert S, Fricker T, Blum C, Schild U, et al. Effect of procalcitonin-based guidelines vs standard guidelines on antibiotic use in lower respiratory tract infections: the ProHOSP randomized controlled trial. *Jama.* 2009; 302(10):1059–1066. [PubMed: 19738090]
44. Bouadma L, Luyt CE, Tubach F, Cracco C, Alvarez A, Schwebel C, Schortgen F, Lasocki S, Veber B, Dehoux M, et al. Use of procalcitonin to reduce patients' exposure to antibiotics in intensive care units (PRORATA trial): a multicentre randomised controlled trial. *Lancet.*

45. Tschoeke SK, Oberholzer A, Moldawer LL. Interleukin-18: a novel prognostic cytokine in bacteria-induced sepsis. *Crit Care Med*. 2006; 34(4):1225–1233. [PubMed: 16540967]
46. Rooney T, Roux-Lombard P, Veale DJ, Fitzgerald O, Dayer JM, Bresnihan B. Synovial tissue and serum biomarkers of disease activity, therapeutic response and radiographic progression. analysis of a proof-of-concept randomized clinical trial of cytokine blockade. *Ann Rheum Dis*. 2009
47. Cusumano V, Midiri A, Cusumano VV, Bellantoni A, De Sossi G, Teti G, Beninati C, Mancuso G. Interleukin-18 is an essential element in host resistance to experimental group B streptococcal disease in neonates. *Infect Immun*. 2004; 72(1):295–300. [PubMed: 14688108]
48. Grobmyer SR, Lin E, Lowry SF, Rivadeneira DE, Potter S, Barie PS, Nathan CF. Elevation of IL-18 in human sepsis. *J Clin Immunol*. 2000; 20(3):212–215. [PubMed: 10941829]
49. Oberholzer A, Steckholzer U, Kurimoto M, Trentz O, Ertel W. Interleukin-18 plasma levels are increased in patients with sepsis compared to severely injured patients. *Shock*. 2001; 16(6):411–414. [PubMed: 11770036]
50. Emmanuilidis K, Weighardt H, Matevossian E, Heidecke CD, Ulm K, Bartels H, Siewert JR, Holzmann B. Differential regulation of systemic IL-18 and IL-12 release during postoperative sepsis: high serum IL-18 as an early predictive indicator of lethal outcome. *Shock*. 2002; 18(4):301–305. [PubMed: 12392271]
51. Feezor RJ, Oberholzer C, Baker HV, Novick D, Rubinstein M, Moldawer LL, Pribble J, Souza S, Dinarello CA, Ertel W, et al. Molecular characterization of the acute inflammatory response to infections with gram-negative versus gram-positive bacteria. *Infect Immun*. 2003; 71(10):5803–5813. [PubMed: 14500502]
52. Parikh CR, Devarajan P. New biomarkers of acute kidney injury. *Crit Care Med*. 2008; 36(4 Suppl):S159–165. [PubMed: 18382188]
53. Parikh CR, Mishra J, Thiessen-Philbrook H, Dursun B, Ma Q, Kelly C, Dent C, Devarajan P, Edelstein CL. Urinary IL-18 is an early predictive biomarker of acute kidney injury after cardiac surgery. *Kidney Int*. 2006; 70(1):199–203. [PubMed: 16710348]
54. Washburn KK, Zappitelli M, Arikian AA, Loftis L, Yalavarthy R, Parikh CR, Edelstein CL, Goldstein SL. Urinary interleukin-18 is an acute kidney injury biomarker in critically ill children. *Nephrol Dial Transplant*. 2008; 23(2):566–572. [PubMed: 17911094]
55. Wong HR, Cvijanovich N, Wheeler DS, Bigham MT, Monaco M, Odoms K, Macias WL, Williams MD. Interleukin-8 as a stratification tool for interventional trials involving pediatric septic shock. *Am J Respir Crit Care Med*. 2008; 178(3):276–282. [PubMed: 18511707]
56. Wong HR, Shanley TP, Sakthivel B, Cvijanovich N, Lin R, Allen GL, Thomas NJ, Doctor A, Kalyanaraman M, Tofil NM, et al. Genome-level expression profiles in pediatric septic shock indicate a role for altered zinc homeostasis in poor outcome. *Physiol Genomics*. 2007; 30(2):146–155. [PubMed: 17374846]
57. Nadel S, Goldstein B, Williams MD, Dalton H, Peters M, Macias WL, Abd-Allah SA, Levy H, Angle R, Wang D, et al. Drotrecogin alfa (activated) in children with severe sepsis: a multicentre phase III randomised controlled trial. *Lancet*. 2007; 369(9564):836–843. [PubMed: 17350452]
58. Nowak JE, Wheeler DS, Harmon KK, Wong HR. Admission chemokine (C-C motif) ligand 4 levels predict survival in pediatric septic shock*. *Pediatr Crit Care Med*. 2009
59. Wong HR, Cvijanovich N, Lin R, Allen GL, Thomas NJ, Willson DF, Freishtat RJ, Anas N, Meyer K, Checchia PA, et al. Identification of pediatric septic shock subclasses based on genome-wide expression profiling. *BMC Med*. 2009; 7:34. [PubMed: 19624809]
60. Byvatov E, Schneider G. Support vector machine applications in bioinformatics. *Appl Bioinformatics*. 2003; 2(2):67–77. [PubMed: 15130823]
61. Vincent JL, Opal SM, Marshall JC. Ten reasons why we should NOT use severity scores as entry criteria for clinical trials or in our treatment decisions. *Crit Care Med*. 38(1):283–287. [PubMed: 19730252]

	Actual	
	Predicted Positive	Predicted Negative
Positive	100	10
Negative	5	95

Sensitivity = $\frac{100}{100 + 10} = 0.91$
Specificity = $\frac{95}{5 + 95} = 0.95$
PPV = $\frac{100}{100 + 5} = 0.95$
NPV = $\frac{95}{10 + 95} = 0.90$

Figure 1.
Two-by-two contingency table illustrating calculations for sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV).

Figure 2.

Hypothetical receiver operating characteristic (ROC) curve. The straight diagonal line represents an ROC curve having an area under the curve (AUC) of 0.5, in which case the test is positive or negative by chance alone. In contrast, the hypothetical ROC curve has an AUC of 0.857.

Figure 3.
Venn analysis comparing gene lists A and B. See text for gene list derivations.

Table 1

Classes of biomarkers.

Biomarker Class	Potential Clinical Use
Diagnostic	To establish the presence or absence of a disease or clinical condition. Screening biomarkers can be considered as a diagnostic biomarker subtype.
Monitoring	To provide a readout that informs the clinician of the effectiveness of a given therapy for the purpose of titration.
Surrogate	To provide a readout that informs the clinician of the effectiveness of a given therapy for the purpose predicting clinical outcome.
Stratification	To sub-classify or stage diseases in terms of severity and/or outcome.

Table 2

Candidate biomarker gene list for derivation of PERSEVERE.

Gene Symbol	Description
CCL3	C-C chemokine ligand 3; a.k.a. MIP-1 α
LCN2	Lipocalin 2; a.k.a. NGAL
MMP8	Matrix metalloproteinase 8
RETN	Resistin
THBS	Thrombospondin 1
GZMB	Granzyme B
HSPA1B	Heat shock protein 70kDa 1B
ORM1	Orosomucoid 1, acute phase protein with unknown function
LTF	Lactotransferrin
ELA2	Neutrophil elastase 1
IL1A	Interleukin 1 α
SULF2	Sulfatase 2; extracellular modulator of heparan sulfate proteoglycans
FGL2	Fibrinogen-like 2; acute phase protein similar to fibrinogen