Interleukin-8 as a Stratification Tool for Interventional Trials Involving Pediatric Septic Shock

Hector R. Wong¹, Natalie Cvijanovich², Derek S. Wheeler¹, Michael T. Bigham¹, Marie Monaco¹, Kelli Odoms¹, William L. Macias³, and Mark D. Williams³

¹Division of Critical Care Medicine, Cincinnati Children's Hospital Medical Center and Cincinnati Children's Research Foundation, and Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, Ohio; ²Division of Critical Care Medicine, Children's Hospital and Research Center Oakland, Oakland, California; and ³Eli Lilly Research Laboratories, Indianapolis, Indiana

Rationale: Interventional clinical trials involving children with septic shock would benefit from an efficient preenrollment stratification strategy.

Objectives: To test the predictive value of interleukin (IL)-8 for 28-day mortality in pediatric septic shock.

Methods: A training data set ($n = 40$) identified a serum IL-8 of greater than 220 pg/ml as having a 75% sensitivity and specificity for predicting 28-day mortality. This cutoff was then subjected to a series of validation steps.

Measurements and Main Results: Subjects were drawn from two large, independent pediatric septic shock databases. Prospective application of the IL-8 cutoff to validation data set 1 ($n = 139$) demonstrated 78% sensitivity and 64% specificity for 28-day mortality. A serum IL-8 level of 220 pg/ml or less, however, had a negative predictive value for 28-day mortality of 95% in validation data set 1, which was subsequently applied to an independently generated data set of children with septic shock (validation set 2, $n = 193$). A serum IL-8 level of 220 pg/ml or less had a negative predictive value for 28-day mortality of 94% when applied to validation set 2.

Conclusions: A serum IL-8 level of 220 pg/ml or less, obtained within 24 hours of admission, predicts a high likelihood of survival in children with septic shock. We propose that IL-8 can be used to exclude such patients from interventional clinical trials and ultimately derive a study population with a more favorable risk to benefit ratio when subjected to a study agent.

Keywords: pediatrics; septic shock; biomarkers; interleukin-8; stratification

Septic shock continues to be a major child health issue, leading to over 4,000 childhood deaths per year in the United States alone (1, 2). Apart from antibiotics, pediatric intensive care unit (PICU)-based supportive care, and vaccination strategies, there are no specific approved therapies for pediatric septic shock. Until the recent RESOLVE (Researching severe Sepsis and Organ dysfunction in children: a gLobal perspective) trial (3), progress in the field has been hampered by a lack of large randomized trials specifically dedicated to children with septic shock.

Conducting randomized trials involving children with septic shock is a difficult endeavor for two broad, interrelated reasons. First, septic shock is a syndrome rather than a distinct disease entity, and is thus characterized by broad heterogeneity with

AT A GLANCE COMMENTARY

Scientific Knowledge on the Subject

The development of an effective biomarker would be a useful stratification tool for clinical studies involving children with septic shock.

What This Study Adds to the Field

Serum interleukin-8 serum level, obtained within 24 hours of admission, has a 95% negative predictive value for mortality in children with septic shock.

regard to etiology, physiologic consequences, morbidity, and mortality (2, 4). Consequently, defining an optimal population for interventional clinical trials has been challenging, almost to the point of being not feasible. Second, compared with the adult septic shock population, the pediatric septic shock population is relatively smaller, as is the associated mortality rate (1, 5). Consequently, enrolling a sufficient number of patients, based on current stratification and eligibility criteria, can be prohibitive in terms of achieving sufficient power to demonstrate efficacy. One potential solution for achieving sufficient statistical power is to involve a large number of pediatric centers; the recent RESOLVE trial involved 104 pediatric centers from 18 countries (3). Although this strategy holds potential for a valid outcome having broad relevance, it also has the potential to introduce additional patient heterogeneity and high variability of "standard" therapy, both of which can be profound confounders for bringing successful trials to fruition (4).

Given these challenges, interventional clinical trials involving children with septic shock would significantly benefit from an effective, clinically feasible stratification strategy. The goal of this stratification strategy would be to identify patient populations having knowable risks for either poor or good outcome as a means for improving the risk to benefit ratio of a given intervention (4). If this goal is attainable, then it may become more feasible to conduct interventional clinical trials in children with septic shock using mortality as the primary outcome.

In an ongoing translational research program involving microarray-based, genome-wide expression profiling in children with septic shock, we recently reported that interleukin (IL)-8 mRNA levels, measured within 24 hours of presentation to the PICU, were increased in nonsurvivors of pediatric septic shock (6). These microarray data were corroborated by increased IL-8 protein in the parallel serum samples of nonsurvivors, thus raising the possibility that serum IL-8 can be used as a biomarker for stratifying mortality risk in children with septic shock.

The primary aim of the current study is to formally and prospectively test the hypothesis that a serum IL-8 level, measured within 24 hours of admission to the PICU, is a robust

(Received in original form January 21, 2008; accepted in final form May 19, 2008)

Supported by a grant from the National Institute of General Medical Sciences (RO1 GM064619). H.R.W. and the GPSS database are supported by the National Institutes of Health and the Cincinnati Children's Research Foundation. The RESOLVE database is funded by Eli Lilly.

Correspondence and requests for reprints should be addressed to Hector R. Wong, M.D., Division of Critical Care Medicine, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, OH 45229-3039. E-mail: hector.wong@cchmc.org

Am J Respir Crit Care Med Vol 178, pp 276–282, 2008

Originally Published in Press as DOI: 10.1164/rccm.200801-1310C on May 29, 2008

Internet address: www.atsjournals.org

biomarker for predicting the outcome of children with septic shock. The strategy for testing this hypothesis involves the prospective use of independent training and validation data sets.

METHODS

Genomics of Pediatric Septic Shock Database

The Genomics of Pediatric Septic Shock (GPSS) database was used to generate the training data set and validation data set 1. The GPSS database has been previously described in detail (6–8). Briefly, this noninterventonal database supports a translational research program focused on genome-level expression profiles (microarray-based) of children with septic shock. Eighteen PICUs in the United States have contributed samples to the GPSS (*see* ACKNOWLEDGMENT). The database contains extensive clinical data, RNA samples, and concomitant serum samples. RNA and serum samples are collected within 24 hours of admission to the PICU and 48 hours after the initial sample collection. In the current analyses, we used only the serum samples drawn within 24 hours of admission to the PICU. The subjects enrolled in the GPSS database include children (<10 yr of age) admitted to the PICUs of multiple institutions and having a diagnosis of septic shock. Pediatric-specific criteria are used to classify patients as having septic shock (9), and the only exclusion criteria are the inability to obtain informed consent or a white blood cell count of less than 1,000 cells/mm³. Standard clinical care is not under protocol for the enrolled subjects, and none of the subjects in the GPSS database are duplicated in the RESOLVE database (*see below*). Patients are monitored for 28 days to determine survival. Accordingly, all mortality-related data in this report refer to this 28-day period.

RESOLVE Database

The RESOLVE database consists of children (up to 17 yr of age) enrolled in a recent phase III, multicenter, interventional trial that evaluated the efficacy of drotrecogin alfa (activated) in children with septic shock (3). The RESOLVE trial is registered with clinicaltrials.gov (NCT00049764). The enrollment criteria for the RESOLVE trial were based on the same overall definition for septic shock as that of the GPSS database, except that enrollment in the RESOLVE trial required the presence of both cardiovascular and respiratory dysfunction 12 hours before entering the study. As would be expected for an interventional trial, the exclusion criteria for the RESOLVE trial were more extensive than that of the GPSS database (3). Four hundred seventy-four patients were analyzed in the RESOLVE trial and one of the secondary outcome variables was 28-day mortality. Of these 474 patients, 278 (59%) had serum IL-8 data available upon enrollment to the trial (i.e., within the first 24 h of PICU admission), and of these 278 patients, 193 (69%) were younger than 10 years. These 193 patients were used for validation data set 2 (*see RESULTS*).

The study protocols supporting the GPSS and RESOLVE databases were approved by the individual institutional review boards of each participating institution.

The patients in both databases are heterogeneous with regard to infecting organisms and comorbidities.

IL-8 Measurements

For the GPSS database, IL-8 measurements were conducted using an IL-8 ELISA kit as previously described (Biosource, Camarillo, CA) (6). For the RESOLVE database, IL-8 measurements were conducted using a validated multiplexed technique (FlowMetrix; Dynamic Systems Solutions, Herndon, VA) (10). Direct comparisons of ELISA and multiplex technology, when using the same antibodies and cytokine standards, demonstrate similar to better lower detection limits for the multiplex technology and correlation coefficients ranging from 0.857 to 1.0 (11–15).

Analytical Approach

Several aspects of the analytical approach will be further described in RESULTS because they will be important for understanding the data analyses. A receiver operating curve (ROC) was generated using SPSS base 16.0 for Windows (SPSS, Inc., Chicago, IL) and the previously reported cohort of patients in the GPSS database (6). Further information

on this patient cohort, including the microarray data, can be found at the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) and are accessible through GEO series accession number GSE4607 (<http://www.ncbi.nlm.nih.gov/geo/>). From this ROC we derived a serum IL-8 cutoff value that was applied prospectively, without *post hoc* calibrations, to the two validation data sets.

Specificities, sensitivities, positive predictive values, and negative predictive values were calculated by standard methods based on contingency tables. Confidence intervals for contingency table-derived data and likelihood ratios were calculated using the VassarStats website (<http://faculty.vassar.edu/lowry/clin1.html>). Continuous variables were compared by Student's *t* test. Categorical variables were compared by χ^2 or Fisher exact test when appropriate. When indicated, significance was set at 0.05.

RESULTS

Training Data Set

The first step in testing the ability of IL-8 to serve as biomarker for outcome in pediatric septic shock was generation of an ROC. The previously reported cohort of children with septic shock (32 survivors and 8 nonsurvivors) served as the training data set for ROC derivation (6). Serum IL-8 levels for this analysis, and all subsequent analyses, were obtained within 24 hours of admission to the PICU. Figure 1 demonstrates the ROC generated from this 40-patient cohort. From this ROC, we extrapolated that a serum IL-8 concentration greater than 220 pg/ml had 75% sensitivity (95% confidence interval [CI], 36–96%) and 75% specificity (95% CI, 56–88%) for predicting 28-day mortality in this training cohort. The positive and negative predictive values for mortality were of 43% (95% CI, 19–70%) and 92% (95% CI, 73–99%), respectively. This IL-8 concentration cutoff of 220 pg/ml was consistently used in all subsequent analyses, without *post hoc* calibrations.

Validation 1

We next sought to test the ability of a serum IL-8 concentration greater than 220 pg/ml to predict mortality for all patients in the GPSS database (validation data set 1). Validation data set 1 consisted of the 40 patients described above in the training data

Figure 1. Receiver operating curve (ROC) used for derivation of IL-8 cutoff greater than 220 pg/ml with a 75% specificity and a 75% sensitivity for predicting 28-day mortality in children with septic shock. The patient cohort used to generate the ROC consists of 40 patients with septic shock (8 nonsurvivors) as previously reported (6). Area under the curve = 0.857; 95% confidence interval = 0.676 to 0.999; standard error = 0.093; *P* value = 0.002.

set, plus an additional 99 consecutive patients with septic shock (10 nonsurvivors) enrolled in the GPSS database, but not previously reported (139 total patients). Table 1 demonstrates that the patients in the training data set and validation data set 1 were similar with respect to mortality rates, age, severity of illness (PRISM [Pediatric Risk of Mortality] score), sex distribution, and median IL-8 levels measured within 24 hours of admission to the PICU. Figure 2 represents the contingency table for patients in validation data set 1 having survivorship and serum IL-8 concentration as the contingency variables. A serum IL-8 concentration greater than 220 pg/ml had 78% sensitivity, 64% specificity, and 25% positive predictive value for mortality. In contrast, a serum IL-8 of 220 pg/ml or less had a negative predictive value for mortality of 95%. In other words, a serum IL-8 concentration of 220 pg/ml or less predicted survival with 95% accuracy in validation data set 1.

This analysis demonstrates that a serum IL-8 level greater than 220 pg/ml, measured within 24 hours of admission to the PICU, does not serve as a robust biomarker for predicting mortality in children with septic shock, at least in this cohort of patients. Importantly, however, when this same measurement generates a serum IL-8 concentration of 220 pg/ml or less, it may have the potential to predict survival with 95% certainty.

Validation 2

We next tested the negative predictive value for mortality of IL-8 by prospectively applying this cutoff (≤ 220 pg/ml) to an independent cohort of children with septic shock (validation data set 2, derived from patients in the RESOLVE database). One hundred and ninety-three patients in the RESOLVE database were younger than 10 years and had available serum IL-8 data upon study entry (i.e., within 24 h of admission to the PICU). Table 1 demonstrates that the patients in validation data set 1 and validation data set 2 were similar with respect to mortality rates, age, illness severity (PRISM score), and sex distribution. African Americans were underrepresented and Hispanics were overrepresented in validation data set 2, com-

pared with validation data set 1. The median IL-8 concentrations in validation data set 2 were higher compared with those of validation data set 1. This latter discrepancy is unlikely to reflect different levels of illness severity, as demonstrated by identical median PRISM scores when comparing validation data set 1 and validation data set 2. Rather, this discrepancy most likely reflects differences in the sensitivity of the methods used for measuring IL-8 (see METHODS).

Table 2 compares demographic data between patients in validation data set 2, the patients in the RESOLVE database 10 years or older with available IL-8 data, and the entire RESOLVE database. As would be predicted, patients in the RESOLVE database 10 years or older with available IL-8 data were older than the two other comparison groups. Otherwise, the three groups were similar with respect to mortality, illness severity, and sex distribution, thus indicating that the subgroup of patients drawn for validation data set 2 are relatively representative of the entire RESOLVE database.

Figure 3 represents the contingency table for validation data set 2 having survivorship and IL-8 concentration as the contingency variables. A serum IL-8 concentration of 220 pg/ml or less had a negative predictive value of 94% for mortality when prospectively applied to validation data set 2. This analysis demonstrates that a serum IL-8 concentration of 220 pg/ml or less, obtained within 24 hours of admission to the PICU, predicts survival with 94% certainty when prospectively applied to an independently generated cohort of children with septic shock.

Performance across All Patients

In this analysis, we combined all patients in validation data set 1 and validation data set 2 (332 total patients, 44 nonsurvivors) and generated a contingency table using the same conventions described above (Figure 4). A serum IL-8 concentration greater than 220 pg/ml had 82% sensitivity, 51% specificity, and 20% positive predictive value for mortality. In contrast, a serum IL-8 concentration of 220 pg/ml or less had a negative predictive value for mortality of 95%. Thus, when all patients in validation data sets 1 and 2 are examined as a single cohort, a serum IL-8 concentration of 220 pg/ml or less predicts survival with 95% accuracy and a narrow CI (90–98%).

To further illustrate the robustness of IL-8 for patient stratification, we again combined all of the patients in validation data sets 1 and 2 and placed them into one of two groups *post hoc*: those with a serum IL-8 concentration of 220 pg/ml or less ($n = 154$) and those with a serum IL-8 concentration greater than 220 pg/ml ($n = 178$). As shown in Figure 5, the group having a serum IL-8 concentration of 220 pg/ml or less had a significantly lower 28-day mortality rate and a significantly lower odds ratio for mortality than the group having a serum IL-8 concentration greater than 220 pg/ml. This *post hoc* analysis further demonstrates the ability of serum IL-8 to stratify patients into two patient populations having significantly different risks for 28-day mortality.

A final *post hoc* analysis examined adverse events in the RESOLVE trial based on IL-8 stratification. As shown in Table 3, the number of adverse events, among patients younger than 10 years and with available IL-8 data, were not significantly different between patients with an IL-8 serum level of 220 pg/ml or less and patients with an IL-8 serum level greater than 220 pg/ml.

DISCUSSION

IL-8 belongs to the CXC chemokine family, is the major neutrophil chemoattractant and activator in humans, and is well known to play a role in the host response to infection-related

TABLE 1. COMPARISON OF THE TRAINING DATA SET, VALIDATION DATA SET 1, AND VALIDATION DATA SET 2

	Training Data Set	Validation Data Set 1	Validation Data Set 2
No. of patients	40	139	193
No. of deaths (%; 95% CI)	8 (20, 10–36)	18 (13, 8–20)	26 (139–19)
Median age, yr (IQR)	2.8 (1.3–3.9)	2.2 (0.8–4.7)	3.2 (2.0–5.9)
Median PRISM score (IQR)	17 (11–22)	16 (9–22)	16 (12–21)
Sex, no. male/female	27/13	85/54	109/84
Ethnic origin: n (%)			
White	22 (55)	91 (65)	125 (64.8)
Hispanic	3 (7)	10 (7)	44 (22.8)*
African descent	11 (27)	28 (20)	8 (4.1)*
Asian	2 (5)	4 (3)	6 (3.1)
Other (mixed racial)	2 (5)	6 (4)	10 (4.7)
Median IL-8, pg/ml (IQR)	97 (0–800)	93 (0–501)	336 (123–1,502) [†]
No. of patients with IL-8 > 220 pg/ml (% mortality, 95% CI)	14 (43, 19–70) [‡]	57 (25, 14–38)	121 (18, 12–26)

Definition of abbreviations: CI = confidence interval; IQR = interquartile range; PRISM = Pediatric Risk of Mortality.

Differences are not significant unless otherwise specified.

* $P < 0.05$ versus training data set and validation data set 1 (χ^2).

[†] $P < 0.05$ versus training data set and validation data set 1 (analysis of variance).

[‡] $P = 0.20$ versus validation data set 1 (Fisher exact test), and $P = 0.07$ versus validation data set 2 (χ^2 test).

	Nonsurvivor	Survivor	
IL-8 > 220 pg/ml	14	43	→ PPV 25% (CI 14 to 38%) LR 2.2 (CI 1.6 to 3.1)
IL-8 ≤ 220 pg/ml	4	78	→ NPV 95% (CI 87 to 98%) LR 0.3 (CI 0.1 to 0.8)
	Sensitivity 78% CI 52 to 93%	Specificity 64% CI 55 to 73%	

Figure 2. Contingency table depicting sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), for 28-day mortality, of IL-8 cutoff (220 pg/ml) when applied to validation data set 1 (n = 139 patients). Also shown are the mortality likelihood ratios (LR) for a positive test (IL-8 > 220 pg/ml) and a negative test (IL-8 ≤ 220 pg/ml), respectively. Note that an IL-8 serum level ≤ 220 pg/ml has an NPV of 95% for 28-day mortality when applied to validation data set 1, with a confidence interval (CI) of 87 to 98%.

conditions such as septic shock and pneumonia (16). The ability to readily measure IL-8 in human samples, combined with its well-defined biological role, has provided the foundation for investigations centered on IL-8 as a biomarker in sepsis and septic shock. Multiple studies have demonstrated the potential utility of IL-8 in the early diagnosis of systemic bacterial infections, particularly in neonates (17–21). Of more direct relevance to the current study, previous studies have correlated increased serum IL-8 levels with increased illness severity and mortality in adults and children with septic shock (22–34). To our knowledge, however, the negative predictive value of IL-8 in pediatric septic shock has never been tested in the manner and with the magnitude of the current study.

On the basis of our previous report demonstrating increased IL-8 mRNA and protein in nonsurvivors of pediatric septic shock (6), we set out to test a singular hypothesis focused on the ability of serum IL-8 to serve as an outcome-specific biomarker. We used the initially reported cohort of patients as the training data set from which we derived a serum IL-8 cutoff value of 220 pg/ml. Importantly, we persisted with this particular value in all subsequent analyses, rather than deriving alternative *post hoc* values. When this serum IL-8 concentration was prospectively applied to a validation data set, generated by the same group of investigators (i.e., the GPSS database), the resulting sensitivities, specificities, and positive predictive values for mortality were not sufficiently robust for clinical utility. In contrast, the negative predictive value of this serum concentration was potentially very robust in that it could predict 28-day survival with 95% certainty.

To test the veracity of this observation (i.e., high negative predictive value), we prospectively applied the serum IL-8

cutoff to an independently generated cohort of children with septic shock (the RESOLVE database). The effectiveness and significance of this approach deserves emphasis. The RESOLVE database was generated by a completely different group of investigators (3). Therefore, the RESOLVE database likely has characteristics that are unique relative to those of the GPSS database, and vice versa. In addition, IL-8 measurements for the RESOLVE database were conducted using a different technique than that of the GPSS database. The use of two different assay techniques reflects the fact that the two study protocols were independently conducted by two distinct research groups. As discussed above, although the two techniques generally have good correlation, the multiplex technique appears to be more reliable at lower IL-8 concentrations (11–15). Despite these potentially profound confounders, we demonstrated that a serum IL-8 concentration of 220 pg/ml or less predicted survival with 94% accuracy in the RESOLVE database patients, thus considerably increasing the reliability and generalizability of our observations.

Because overall mortality for pediatric septic shock has been estimated to be approximately 10% (1, 5), and the overall patient cohort in this study had a mortality rate of 13%, one could argue that predicting survivorship with 95% certainty has limited utility. What then is the clinical utility of a serum IL-8 concentration of 220 pg/ml or less in children with septic shock? For the care of an individual child with septic shock, we do not envision any particular utility. Rather, we envision that this cutoff could be used as a stratification strategy for future interventional clinical trials involving children with septic shock. Although previous biomarker-based stratification strategies have focused on *a priori* selection, for inclusion of patients with the highest mortality risk (4, 35–37), the strategy that we are envisioning has a subtle, but important difference. The strategy we envision would exclude patients who would otherwise be eligible for enrollment, but who have a low mortality risk based on a study entry serum IL-8 concentration of 220 pg/ml or less. This type of strategy would serve to exclude patients who have a high probability (up to 95%) of surviving septic shock with standard care, thus ultimately generating a study population with a higher mortality rate (20% in the current cohort of patients). We are not aware of any other septic shock biomarker that can achieve a 95% negative predictive value for mortality. Other biomarkers (i.e., IL-6), or a combination of biomarkers, may have a similar (or better) negative predictive value, but this hypothesis has not been directly tested in the manner described in the current study.

By excluding patients with a low mortality risk, who would otherwise meet entry criteria, the risk to benefit ratio could be

TABLE 2. COMPARISON OF VALIDATION DATA SET 2, PATIENTS IN THE RESOLVE DATABASE ≥ 10 YEARS OF AGE WITH AVAILABLE IL-8 LEVELS, AND ALL PATIENTS IN THE RESOLVE DATABASE

	Validation Data Set 2	RESOLVE Database Patients ≥ 10 Years Old with IL-8 Levels	Entire RESOLVE Database
No. of patients	193	85	474
No. of deaths (%; 95% CI)	26 (13, 9–19)	19 (22, 14–33)	82 (17, 14–21)
Median age, yr (IQR)	3.2 (2.0–5.9)	14.3 (12.2–15.6)*	2.6 (0.7–7.8)
Median PRISM score (IQR)	16 (12–21)	19 (13–23)	16 (11–21)
Sex no. male/female	109/84	46/39	258/216

Differences are not significant unless otherwise specified.

* $P < 0.05$ versus validation data set 2 and the entire RESOLVE database (analysis of variance).

Figure 3. Contingency table depicting sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), for 28-day mortality, of IL-8 cutoff (220 pg/ml) when prospectively applied to validation data set 2 (n = 193). Also shown are the mortality likelihood ratios (LR) for a positive test (IL-8 > 220 pg/ml) and a negative test (IL-8 ≤ 220 pg/ml), respectively. Note that an IL-8 serum level ≤ 220 pg/ml has an NPV of 94% for 28-day mortality when prospectively applied to validation data set 2, with a confidence interval (CI) of 86 to 98%.

improved such that experimental interventions may have a higher likelihood of demonstrating efficacy. This assertion is predicated on an experimental intervention having a relatively constant risk profile across the range of illness severity. Conversely, the stratification strategy may not be as robust if the potential for harm of the experimental strategy significantly increases in proportion to illness severity. For example, of the 193 patients in the RESOLVE trial used in the current analysis, 72 (37%) would have been excluded under the current proposal based on study entry IL-8 levels. These 72 patients otherwise had a sufficient level of illness severity to warrant enrollment. These same 72 patients, however, had a mortality rate of 6% (compared with an overall mortality rate in the RESOLVE trial of 17%), whereas 5% of the overall study population in the RESOLVE trial had a study drug–related serious adverse event (3). In addition, our *post hoc* analysis of adverse events in the RESOLVE trial demonstrated similar adverse event rates in patients with serum IL-8 levels of 220 pg/ml or less and patients with serum IL-8 levels greater than 220 pg/ml. Thus, including these 72 patients could potentially expose them to a risk of a study drug–related serious adverse event that is virtually identical to that of the mortality risk.

Several potential limitations of the current study deserve discussion. First, the IL-8 measurement techniques varied between the two databases, as discussed above. Second, the GPSS database was developed with the goal of generating genome-level expression data in children (6–8). The decision to limit the GPSS database to children younger than 10 years was a programmatic decision seeking to distinguish our efforts

from that of other investigative groups studying adult septic shock. At the initiation of the GPSS database, we decided to only include patients younger than 10 years because this population would be exclusively prepubertal and would undeniably qualify as being representative of “children.” All children younger than 10 years and meeting criteria for septic shock are eligible for enrollment in the GPSS, with the only exclusion criteria being the inability to obtain informed consent or a white blood cell count less than 1,000 cells/mm³, which would preclude isolation of sufficient RNA for microarray analysis. Patients receiving chemotherapy, undergoing bone marrow transplantation, and/or receiving immunosuppressive medications, regardless of their expected outcomes, are included in the GPSS database. In contrast, the RESOLVE database was developed with the goal of testing a novel therapeutic strategy in patients between 38 weeks of corrected gestational age and 17 years of age (3). As a test agent efficacy trial, the patient population was more highly selected compared with that of the GPSS database. For example, patients with high risk of intracranial bleeding, expected to die before the 28-day study period from preexisting conditions, with end-stage renal or liver disease, with a platelet count less than 30,000/mm³, and/or having undergone bone marrow transplantation were excluded from the RESOLVE trial, whereas these types of patients were not excluded from the GPSS database. Patients undergoing chemotherapy and patients receiving immunosuppressive medications were included in the RESOLVE trial as long as they did not meet the exclusion criteria listed above. As stated above, despite these potentially profound confounders, we were

Figure 4. Contingency table depicting sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), for 28-day mortality, of IL-8 cutoff (220 pg/ml) when applied to the combined validation data sets 1 and 2 (n = 332). Also shown are the mortality likelihood ratios (LR) for a positive test (IL-8 > 220 pg/ml) and a negative test (IL-8 ≤ 220 pg/ml), respectively. Note that an IL-8 serum level ≤ 220 pg/ml has an NPV of 95% for 28-day mortality when applied to combined validation data sets 1 and 2, with a confidence interval (CI) of 90 to 98%.

Figure 5. An IL-8 cutoff value of 220 pg/ml segregates children with septic shock into two significantly different risk groups for 28-day mortality ($P < 0.001$, χ^2). Data are derived from 332 patients after combining validation data sets 1 and 2.

able to prospectively validate the negative predictive value of serum IL-8 using the RESOLVE database. The ability to overcome these potential limitations, while testing a single hypothesis, without *post hoc* modifications, strengthens the validity and generalizability of our data.

A third potential limitation is that all available patients in the GPSS database were used, whereas only a selected group of patients from the RESOLVE database were used for the current study (i.e., patients < 10 yr of age and having available IL-8 data). The rationale for selecting patients from the RESOLVE database was to derive a validation data set with a comparable age range as that of the training data set, but this approach excluded 60% of the patients in the RESOLVE database. Thus, we recognize that our data may be generalizable only to children with septic shock who are younger than 10 years. A *post hoc* analysis, however, involving all 278 patients in the RESOLVE database having available IL-8 data (58% of the total RESOLVE database and including children up to 17 yr of age) demonstrated that a serum IL-8 level of 220 pg/ml or less had a negative predictive value for 28-day mortality of 92% (95% CI, 85–96%).

In summary, we have demonstrated that a serum IL-8 concentration of 220 pg/ml or less, obtained within 24 hours of admission to the PICU, serves as a robust negative predictor for 28-day mortality in children younger than 10 years with septic shock. The assertion of robustness is supported by prospective validation using a patient cohort completely indepen-

dent of the training cohort. Furthermore, the ability to validate this cutoff using an independent cohort implies that the cutoff is valid across a broad spectrum of children with septic shock (generalizability). Generalizability is also supported by the fact that the data are derived from children admitted to multiple pediatric centers and from children with significant comorbidities. We propose that a serum IL-8 level of 220 pg/ml or less, obtained within 24 hours of admission, can be used as an effective strategy to exclude patients from interventional clinical trials involving children with septic shock. Such a strategy would serve to exclude patients with a high likelihood of surviving with standard care and ultimately derive a study population with a more favorable risk to benefit ratio when subjected to a study agent. Because reagents to measure IL-8 are widely available and serum is a readily obtainable biological sample, it is highly likely that a rapid, point-of-care test can be developed for measuring IL-8 at the bedside. Thus, the proposed approach is highly feasible.

Conflict of Interest Statement: H.R.W. submitted a provisional patent application in August 2007 for the use of IL-8 as a stratification tool in children with septic shock. N.C. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. D.S.W. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. M.T.B. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. M.M. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. K.O. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. W.L.M. is an employee and shareholder of Eli Lilly and Company. M.D.W. is an employee and shareholder of Eli Lilly and Company.

TABLE 3. ADVERSE EVENTS IN THE RESOLVE TRIAL BASED ON IL-8 STRATIFICATION

	IL-8 ≤ 220 pg/ml (n = 72)	IL-8 > 220 pg/ml (n = 121)	P Value
Serious adverse events, n (%)			
Day 0 to 6	5 (6.9)	11 (9.1)	0.6*
Day 0 to 28	12 (16.7)	21 (17.4)	0.9*
Serious bleeding events, n (%)			
Day 0 to 6	0 (0)	4 (3.3)	0.3†
Day 0 to 28	1 (1.4)	9 (7.4)	0.09†
Serious bleeding events, n (%)			
Day 0 to 6	0 (0)	0 (0)	—
Day 0 to 28	1 (1.4)	4 (3.3)	0.6†

* χ^2 .

† Fisher exact test.

Contributing investigators and centers for the GPSS database are as follows: Geoffrey L. Allen (Children's Mercy Hospital, Kansas City, MO); Nick Anas (Children's Hospital of Orange County, Orange, CA); Paul Checchia (St. Louis Children's Hospital, St. Louis, MO); Allan Doctor (St. Louis Children's Hospital, St. Louis, MO); Robert Fitzgerald (Devos Children's Hospital, Grand Rapids, MI); Robert J. Freishtat (Children's National Medical Center, Washington, DC); Jose Gutierrez (Pediatric Critical Care of Arizona, Phoenix, AZ); Meena Kalyanaraman (Newark Beth Israel Medical Center, Newark, NJ); Gary Kohn (Morristown Memorial Hospital, Morristown, NJ); Cheri Landers (Kentucky Children's Hospital, Lexington, KY); Richard Lin (The Children's Hospital of Philadelphia, Philadelphia, PA); Glenn McLaughlin (Jackson Memorial Hospital, Miami, FL); Keith Meyer (Miami Children's Hospital, Miami, FL); Scott Penfil (DuPont Hospital for Children, Wilmington, DE); Steve Shane (Washoe Medical Center, Reno, NV); Thomas P. Shanley (C.S. Mott Children's Hospital at the University of Michigan, Ann Arbor, MI); Neal J. Thomas (Penn State Children's Hospital, Hershey, PA); Nancy M. Tofil (The University of Alabama at Birmingham, Birmingham, AL); Margaret Winkler (The University of Alabama at Birmingham, Birmingham, AL); and Douglas Willson (University of Virginia, Charlottesville, VA).

References

- Watson RS, Carcillo JA, Linde-Zwirble WT, Clermont G, Lidicker J, Angus DC. The epidemiology of severe sepsis in children in the United States. *Am J Respir Crit Care Med* 2003;167:695–701.
- Shanley TP, Hallstrom C, Wong HR. Sepsis. In: Fuhrman BP, Zimmerman JJ, editors. *Pediatric critical care medicine*, 3rd ed. St. Louis: Mosby; 2006. pp. 1474–1493.
- 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;369:836–843.
- Marshall JC, Vincent JL, Guyatt G, Angus DC, Abraham E, Bernard G, Bombardier C, Calandra T, Jorgensen HS, Sylvester R, et al. Outcome measures for clinical research in sepsis: a report of the 2nd Cambridge Colloquium of the International Sepsis Forum. *Crit Care Med* 2005;33:1708–1716.
- Watson RS, Carcillo JA. Scope and epidemiology of pediatric sepsis. *Pediatr Crit Care Med* 2005;6(3, Suppl):S3–S5.
- Wong HR, Shanley TP, Sakthivel B, Cvijanovich N, Lin R, Allen GL, Thomas NJ, Doctor A, Kalyanaraman M, Tofil NM, Penfil S, Monaco M, et al. Genome-level expression profiles in pediatric septic shock indicate a role for altered zinc homeostasis in poor outcome. *Physiol Genomics* 2007;30:146–155.
- Shanley TP, Cvijanovich N, Lin R, Allen GL, Thomas NJ, Doctor A, Kalyanaraman M, Tofil NM, Penfil S, Monaco M, et al. Genome-level longitudinal expression of signaling pathways and gene networks in pediatric septic shock. *Mol Med* 2007;13:495–508.
- Wong HR. Pediatric septic shock treatment: new clues from genomic profiling. *Pharmacogenomics* 2007;8:1287–1290.
- Goldstein B, Giroir B, Randolph A. International pediatric sepsis consensus conference: definitions for sepsis and organ dysfunction in pediatrics. *Pediatr Crit Care Med* 2005;6:2–8.
- Fulton RJ, McDade RL, Smith PL, Kienker LJ, Kettman JR Jr. Advanced multiplexed analysis with the FlowMetrix system. *Clin Chem* 1997;43:1749–1756.
- Kellar KL, Kalwar RR, Dubois KA, Crouse D, Chafin WD, Kane BE. Multiplexed fluorescent bead-based immunoassays for quantitation of human cytokines in serum and culture supernatants. *Cytometry* 2001;45:27–36.
- Kellar KL, Iannone MA. Multiplexed microsphere-based flow cytometric assays. *Exp Hematol* 2002;30:1227–1237.
- Chen R, Lowe L, Wilson JD, Crowther E, Tzeggai K, Bishop JE, Varro R. Simultaneous quantification of six human cytokines in a single sample using microparticle-based flow cytometric technology. *Clin Chem* 1999;45:1693–1694.
- Cook EB, Stahl JL, Lowe L, Chen R, Morgan E, Wilson J, Varro R, Chan A, Graziano FM, Barney NP. Simultaneous measurement of six cytokines in a single sample of human tears using microparticle-based flow cytometry: allergies vs. non-allergies. *J Immunol Methods* 2001;254:109–118.
- Oliver KG, Kettman JR, Fulton RJ. Multiplexed analysis of human cytokines by use of the FlowMetrix system. *Clin Chem* 1998;44:2057–2060.
- Remick DG. Interleukin-8. *Crit Care Med* 2005;33(12, Suppl):S466–S467.
- Harris MC, D'Angio CT, Gallagher PR, Kaufman D, Evans J, Kilpatrick L. Cytokine elaboration in critically ill infants with bacterial sepsis, necrotizing enterocolitis, or sepsis syndrome: correlation with clinical parameters of inflammation and mortality. *J Pediatr* 2005;147:462–468.
- Verboon-Macielek MA, Thijsen SF, Hemels MA, Menses M, van Loon AM, Krediet TG, Gerards LJ, Fleer A, Voorbij HA, Rijkers GT. Inflammatory mediators for the diagnosis and treatment of sepsis in early infancy. *Pediatr Res* 2006;59:457–461.
- Ng PC, Lam HS. Diagnostic markers for neonatal sepsis. *Curr Opin Pediatr* 2006;18:125–131.
- Horisberger T, Harbarth S, Nadal D, Baenziger O, Fischer JE. G-CSF and IL-8 for early diagnosis of sepsis in neonates and critically ill children—safety and cost effectiveness of a new laboratory prediction model: study protocol of a randomized controlled trial. *Crit Care* 2004;8:R443–R450.
- Santana Reyes C, Garcia-Munoz F, Reyes D, Gonzalez G, Dominguez C, Domenech E. Role of cytokines (interleukin-1beta, 6, 8, tumour necrosis factor-alpha, and soluble receptor of interleukin-2) and C-reactive protein in the diagnosis of neonatal sepsis. *Acta Paediatr* 2003;92:221–227.
- Damas P, Canivet JL, de Groote D, Vrindts Y, Albert A, Franchimont P, Lamy M. Sepsis and serum cytokine concentrations. *Crit Care Med* 1997;25:405–412.
- Vermont CL, Hazelzet JA, de Kleijn ED, van den Dobbelsteen GP, de Groot R. CC and CXC chemokine levels in children with meningococcal sepsis accurately predict mortality and disease severity. *Crit Care* 2006;10:R33.
- Lin KJ, Lin J, Hanasawa K, Tani T, Kodama M. Interleukin-8 as a predictor of the severity of bacteremia and infectious disease. *Shock* 2000;14:95–100.
- Kern WV, Heiss M, Steinbach G. Prediction of gram-negative bacteremia in patients with cancer and febrile neutropenia by means of interleukin-8 levels in serum: targeting empirical monotherapy versus combination therapy. *Clin Infect Dis* 2001;32:832–835.
- Van der Kaay DC, De Kleijn ED, De Rijke YB, Hop WC, De Groot R, Hazelzet JA. Procalcitonin as a prognostic marker in meningococcal disease. *Intensive Care Med* 2002;28:1606–1612.
- Schots R, Kaufman L, Van Riet I, Ben Othman T, De Waele M, Van Camp B, Demanet C. Proinflammatory cytokines and their role in the development of major transplant-related complications in the early phase after allogeneic bone marrow transplantation. *Leukemia* 2003;17:1150–1156.
- Ozbalkan Z, Aslar AK, Yildiz Y, Aksaray S. Investigation of the course of proinflammatory and anti-inflammatory cytokines after burn sepsis. *Int J Clin Pract* 2004;58:125–129.
- Livaditi O, Kotanidou A, Psarra A, Dimopoulou I, Sotiropoulou C, Augustatou K, Papasteriades C, Armaganidis A, Roussos C, Orfanos SE, et al. Neutrophil CD64 expression and serum IL-8: sensitive early markers of severity and outcome in sepsis. *Cytokine* 2006;36:283–290.
- Bozza FA, Salluh JJ, Japiassu AM, Soares M, Assis EF, Gomes RN, Bozza MT, Castro-Faria-Neto HC, Bozza PT. Cytokine profiles as markers of disease severity in sepsis: a multiplex analysis. *Crit Care* 2007;11:R49.
- Rivers EP, Kruse JA, Jacobsen G, Shah K, Loomba M, Otero R, Childs EW. The influence of early hemodynamic optimization on biomarker patterns of severe sepsis and septic shock. *Crit Care Med* 2007;35:2016–2024.
- Hack CE, Hart M, van Schijndel RJ, Eerenberg AJ, Nuijens JH, Thijs LG, Aarden LA. Interleukin-8 in sepsis: relation to shock and inflammatory mediators. *Infect Immun* 1992;60:2835–2842.
- Fujishima S, Sasaki J, Shinozawa Y, Takuma K, Kimura H, Suzuki M, Kanazawa M, Hori S, Aikawa N. Serum MIP-1 alpha and IL-8 in septic patients. *Intensive Care Med* 1996;22:1169–1175.
- Hein OV, Misterek K, Tessmann JP, van Dossow V, Krimphove M, Spies C. Time course of endothelial damage in septic shock: prediction of outcome. *Crit Care* 2005;9:R323–R330.
- Marshall JC. Biomarkers of sepsis. *Curr Infect Dis Rep* 2006;8:351–357.
- Marshall JC. The staging of sepsis: understanding heterogeneity in treatment efficacy. *Crit Care* 2005;9:626–628.
- Dellinger RP, Abraham E, Bernard G, Marshall JC, Vincent JL. Controversies in sepsis clinical trials: proceedings of a meeting of the International Sepsis Forum, Lausanne, Switzerland, September 29, 2001. *J Crit Care* 2006;21:38–47.