EBioMedicine 2 (2015) 2087–2093

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Article

Prospective Testing and Redesign of a Temporal Biomarker Based Risk Model for Patients With Septic Shock: Implications for Septic Shock Biology

Hector R. Wong [a,b,*], Natalie Z. Cvijanovich [c], Nick Anas [d], Geoffrey L. Allen [e], Neal J. Thomas [f], Michael T. Bigham [g], Scott L. Weiss [h], Julie Fitzgerald [h], Paul A. Checchia [i], Keith Meyer [j], Michael Quasney [k], Mark Hall [l], Rainer Gedeit [m], Robert J. Freishtat [n], Jeffrey Nowak [o], Shekhar S. Raj [p], Shira Gertz [q], Kelli Howard [a], Kelli Harmon [a], Patrick Lahni [a], Erin Frank [a], Kimberly W. Hart [r], Christopher J. Lindsell [r]

[a] Division of Critical Care Medicine, Cincinnati Children's Hospital Medical Center and Cincinnati Children's Research Foundation, Cincinnati, OH, United States
[b] Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH, United States
[c] UCSF Benioff Children's Hospital Oakland, Oakland, CA, United States
[d] Children's Hospital of Orange County, Orange, CA, United States
[e] Children's Mercy Hospital, Kansas City, MO, United States
[f] Penn State Hershey Children's Hospital, Hershey, PA, United States
[g] Akron Children's Hospital, Akron, OH, United States
[h] The Children's Hospital of Philadelphia, Philadelphia, PA, United States
[i] Texas Children's Hospital and Baylor College of Medicine, Houston, TX, United States
[j] Miami Children's Hospital, Miami, FL, United States
[k] CS Mott Children's Hospital at the University of Michigan, Ann Arbor, MI, United States
[l] Nationwide Children's Hospital, Columbus, OH, United States
[m] Children's Hospital of Wisconsin, Milwaukee, WI, United States
[n] Children's National Medical Center, Washington, DC, United States
[o] Children's Hospital and Clinics of Minnesota, Minneapolis, MN, United States
[p] Riley Hospital for Children, Indianapolis, IN, United States
[q] Hackensack University Medical Center, Joseph M. Sanzari Children's Hospital, Hackensack, NJ, United States
[r] Department of Emergency Medicine, University of Cincinnati College of Medicine, Cincinnati, OH, United States

ARTICLE INFO

Article history:
Received 5 October 2015
Received in revised form 7 November 2015
Accepted 19 November 2015
Available online 22 November 2015

Keywords:
Sepsis
Biomarkers
Outcome
Prediction
Modeling
Inflammation
Immune Suppression
Endotype

ABSTRACT

The temporal version of the pediatric sepsis biomarker risk model (tPERSEVERE) estimates the risk of a complicated course in children with septic shock based on biomarker changes from days 1 to 3 of septic shock. We validated tPERSEVERE performance in a prospective cohort, with an a priori plan to redesign tPERSEVERE if it did not perform well. Biomarkers were measured in the validation cohort (n = 168) and study subjects were classified according to tPERSEVERE. To redesign tPERSEVERE, the validation cohort and the original derivation cohort (n = 299) were combined and randomly allocated to training (n = 374) and test (n = 93) sets. tPERSEVERE was redesigned using the training set and CART methodology. tPERSEVERE performed poorly in the validation cohort, with an area under the curve (AUC) of 0.67 (95% CI: 0.58–0.75). Failure analysis revealed potential confounders related to clinical characteristics. The redesigned tPERSEVERE model had an AUC of 0.83 (0.79–0.87) and a sensitivity of 93% (68–97) for estimating the risk of a complicated course. Similar performance was seen in the test set. The classification tree segregated patients into two broad endotypes of septic shock characterized by either excessive inflammation or immune suppression.

1. Introduction

Septic shock is a dynamic clinical and biological syndrome (Cohen et al., 2015). Patient outcomes are highly variable, reflecting a complex, time-dependent interplay between inflammation, immunity, pathogen-related factors, patient heterogeneity, and therapeutic interventions. We have attempted to navigate this complexity at the individual patient

* Corresponding author at: Division of Critical Care Medicine, MLC 2005, Cincinnati Children's Hospital Medical Center, 3333 Burnet Avenue, Cincinnati, OH 45229, United States.
E-mail address: hector.wong@cchmc.org (H.R. Wong).

http://dx.doi.org/10.1016/j.ebiom.2015.11.035

level by developing biomarker-based models to estimate the probability of poor outcomes in patients with septic shock (Alder et al., 2014; Wong et al., 2012, 2014a,b; Kaplan and Wong, 2011).

Analogous to longstanding concepts in the oncology field, we contend that understanding baseline probability of poor outcome is fundamental to clinical practice and research in the field of septic shock. Prognostic models that reliably estimate the risk of poor outcome have the potential to serve as tools for enrichment of clinical trials, inform individual patient decision-making, to serve as a benchmark for quality improvement efforts, and to facilitate risk-stratified analyses of clinical data. Further, such models have the potential to provide insight regarding the pathogenesis of septic shock and how it varies among different patients.

The Pediatric Sepsis Biomarker Risk Model (PERSEVERE) incorporates a panel of biomarkers and age into a decision tree estimating the baseline risk of mortality in children with septic shock (Wong et al., 2012, 2014b). The PERSEVERE biomarkers are proteins measured in the blood compartment on day 1 of presentation to the intensive care unit with septic shock. To reflect change in risk over time, we developed a temporal version of the model (tPERSEVERE) (Wong et al., 2014c). tPERSEVERE considers how the PERSEVERE biomarker concentrations change from day 1 to day 3 of septic shock, and how these changes are associated with poor outcome.

We envisioned that a reliable temporal model could potentially serve as a monitor for therapeutic efficacy, in combination with traditional clinical parameters. For example, if the model signals a decreasing risk for poor outcome over time, this could be indicative of therapeutic efficacy. Alternatively, if the model signals increasing risk or unchanged risk from a high baseline risk, this could indicate lack of efficacy and could potentially trigger a reassessment of the therapeutic regimen.

Models such as tPERSEVERE require prospective testing in order to assess validity and generalizability. In the current study, we prospectively tested the performance of tPERSEVERE in an independent validation cohort, and use the results to explore how biological variation may be associated with the pathogenesis of poor outcomes in septic shock.

2. Methods

2.1. Study Subjects and Data Collection

The validation cohort consisted of 168 subjects prospectively enrolled since the initial derivation of tPERSEVERE. The protocol for collection and use of biological specimens and clinical data was approved by the Institutional Review Boards of each of 18 participating institutions. Children ≤18 years of age admitted to the pediatric intensive care unit (PICU) and meeting pediatric-specific consensus criteria for septic shock were eligible for enrollment (Goldstein et al., 2005; Wong et al., 2007). The only exclusion criterion was the inability to obtain informed consent, which was obtained from parents or legal guardians prior to any data or sample collection.

Serum samples were obtained within 24 h of first meeting the criteria for septic shock in the PICU, which was typically at presentation. These are referred to as “day 1” samples. “Day 3” samples were collected 48 h after the day 1 samples. Clinical and laboratory data were collected daily while in the PICU. Organ failure data were tracked up to day seven of septic shock using previously published criteria (Goldstein et al., 2005). Mortality was tracked for 28 days after enrollment. Complicated course was defined as the persistence of two or more organ failures at day seven of septic shock or 28-day mortality (Wong et al., 2015). Illness severity was estimated using PRISM scores (Pollack et al., 1997). Baseline mortality probability was estimated using PERSEVERE (Wong et al., 2012, 2014b).

2.2. PERSEVERE Biomarkers

PERSEVERE includes C–C chemokine ligand 3 (CCL3), interleukin 8 (IL8), heat shock protein 70 kDa 1B (HSPA1B), granzyme B (GZMB), and matrix metallopeptidase 8 (MMP8) (Wong et al., 2012). Serum concentrations of these biomarkers were measured using a multi-plex magnetic bead platform (MILLIPLEX™ MAP, EMD Millipore Corporation, Billerica, MA). Biomarker concentrations were measured in a Luminex® 100/200 System (Luminex Corporation, Austin, TX), according the manufacturers' specifications. Assay performance data were previously published (Wong et al., 2012).

2.3. Statistical Analysis and Validation of tPERSEVERE

Initially, data are described using medians, interquartile ranges, frequencies, and percentages. Comparisons between groups used the Mann–Whitney U-test, Chi-square, or Fisher's exact tests as appropriate. Descriptive statistics and comparisons used SigmaStat Software (Systat Software, Inc., San Jose, CA).

Each study subject was assigned a probability of a complicated course using the previously published tPERSEVERE model (Wong et al., 2014c). tPERSEVERE performance is reported using diagnostic test statistics with 95% confidence intervals computed using SPSS 23.0 (IBM Corporation, Armonk, NY), R (base version 3.1.1) and package epiR (version 0.9–62) (R Core Team, 2014; Stevenson et al., 2015).

2.4. Redesigning tPERSEVERE

A priori, we determined that if the area under the receiver operating curve was less than 0.7 in the validation cohort, we would redesign tPERSEVERE. Initially, we explored reasons for failure by comparing the validation cohort to the original cohort, and by characterizing false negatives and comparing them to true positives. Then, to redesign tPERSEVERE we combined the 168 prospectively enrolled subjects, and the 299 previously reported subjects. From this pooled cohort ($n = 467$) we randomly selected 80% of the subjects for a training cohort ($n = 374$) and the remaining 20% were retained as a test cohort ($n = 93$).

The modeling procedures for redesigning tPERSEVERE used CART methodology (Salford Predictive Modeler v6.6, Salford Systems, San Diego, CA) (Che et al., 2011; Muller and Mockel, 2008). The primary outcome variable for the modeling was a complicated course, as defined above (Wong et al., 2015). Using complicated course as the primary outcome variable allows for the exploration of association between temporal biomarker changes and nuances of sepsis severity beyond the dichotomy of “alive” vs. “dead”. Continuous, predictor variables for the modeling procedure included the PERSEVERE-based mortality probability, day 1 and day 3 PERSEVERE biomarker values, age, and a derived variable termed “delta”, which subtracted the day 1 value for a given biomarker from the respective day 3 value. Dichotomous predictor variables included gender, and the presence of any co-morbidity, malignancy, immune suppression, or previous bone marrow transplantation. Weighting of cases and the addition of cost for misclassification were not used in the modeling procedures. The code and data used to generate the model is available from the authors.

2.5. Funding

The study was funded by the National Institutes of Health, National Institute of General Medical Sciences. The funder played no other role in the study, in the writing of the manuscript, or in the decision to submit the manuscript.

3. Results

3.1. Prospective Validation of tPERSEVERE

Table 1 shows the demographic and clinical characteristics of the validation cohort ($n = 168$). Sixty-three subjects (38%) had a complicated course. Among these, 25 (40%) died by study day 28. Compared