Published in final edited form as:

Ann Surg. 2009 October ; 250(4): 531–539. doi:10.1097/SLA.0b013e3181b8fbd5.

Validation of the Riboleukogram to Detect Ventilator-Associated Pneumonia After Severe Injury

J. Perren Cobb, MD, Ernest E. Moore, MD, Doug L. Hayden, MA, Joseph P. Minei, MD, Joseph Cuschieri, MD, Jingyun Yang, PhD, Qing Li, BS, Nan Lin, PhD, Bernard H. Brownstein, PhD, Laura Hennessy, RN, Philip H. Mason, PhD, William S. Schierding, MS, David J. Dixon, PhD, Ronald G. Tompkins, MD, ScD, H. Shaw Warren, MD, David A. Schoenfeld, PhD, and Ronald V. Maier, MD

* Center for Critical Illness and Health Engineering; Washington University in St. Louis, St. Louis, MO

† The NIGMS Inflammation and Host Response to Injury Program; Massachusetts General Hospital, Boston, MA

Abstract

Objective—We hypothesized that circulating leukocyte RNA profiles or “riboleukograms” detect ventilator-associated pneumonia after blunt trauma.

Summary Background Data—A pilot microarray study of 11 ventilator-associated pneumonia (VAP) patients suggested that 85 leukocyte genes can be used to diagnose VAP. Validation of this gene set to detect VAP was tested using data from an independent patient cohort.

Methods—A total of 158 intubated blunt trauma patients were enrolled at 5 centers, where 57 (36%) developed VAP. Patient age was 34.2 ± 11.1 years; 65% were male. Circulating leukocyte GeneChip U133 2.0 expression values were measured at time 0.5, 1, 4, 7, 14, 21, and 28 days after injury. DChip normalized leukocyte transcriptional profiles were analyzed using repeated measures logistic regression. A compound covariate model based on leukocyte gene transcriptional profiles in a training subset of patients was tested to determine predictive accuracy for VAP 4 days prior to clinical diagnosis in the test subset.

Results—Using gene expression values measured on each study day at an FDR <0.05, 27 (32%) of the 85 genes were associated with the diagnosis of VAP 1 to 4 days before diagnosis. However, the compound covariate model based on these 85-genes did not predict VAP in the test cohort better than chance ($P = 0.27$). In contrast, a compound covariate model based upon de novo transcriptional analysis of the 158 patients predicted VAP better than chance 4 days before diagnosis with a sensitivity of 57% and a specificity of 69%.

Conclusion—Our results validate those described in a pilot study, confirming that riboleukograms are associated with the development of VAP days prior to clinical diagnosis. Similarly, a riboleukogram predictive model tested on a larger cohort of 158 patients was better than chance at predicting VAP days prior to clinical diagnosis.

Keywords

trauma; sepsis; micro array; genechip; diagnosis; RNA; glue grant

Reprints: J. Perren Cobb, MD, FACS, Campus Box 8109, 660 S. Euclid Ave, St. Louis, MO 63110. cobb@wustl.edu.

None of the authors has a financial interest in the work presented in this manuscript.

There is no commercial sponsor for the work described in this manuscript. Reprints will not be available from the author.

Ventilator-associated pneumonia (VAP) is the most common nosocomial infection in the critically ill, occurring in 10% to 20% of those patients receiving mechanical ventilation.¹ The rate of VAP is highest in trauma patients and is independently associated with death in those that are less severely injured.^{2,3} Because VAP also increases days on the ventilator, the resulting increase in trauma-associated costs is substantial, as high as \$57,000 per episode.⁴ These significant increases in morbidity, mortality, and cost have made lowering VAP rates an obvious target for quality care. Recent reports indicate that preventative measures are effective in significantly decreasing the rates of VAP after injury, sparking interest in VAP as a metric for performance improvement.^{5–7}

Improvements in the treatment of VAP, however, are confounded by the absence of accurate and timely diagnostics. Without a “gold standard,” the diagnostic criteria for VAP remain controversial, as it is difficult to distinguish VAP from tracheal colonization in patients with other forms of lung injury.^{6,8,9} Thus, “the establishment of an appropriate diagnosis of VAP is one of the most crucial and difficult issues in the care of critically ill patients.”¹⁰ There is a tendency for clinicians to over diagnose VAP, because of the dire consequences of untreated or inappropriately treated pneumonia.^{6,9,11,12} This practice not only risks the harm of unnecessary antibiotics, but also increases the likelihood of infection with resistant organisms.⁹ Advances in molecular diagnostics thus hold great promise for VAP, especially if these tests could determine earlier the offending organism(s).

Recent reports suggested the clinical utility of riboleukograms, that is, circulating leukocyte transcriptional profiles over time, used to track the host response to and recovery from critical illness complicated by pneumonia.^{13,14} On the basis of the report of an 85-gene riboleukogram,¹³ coupled with our own preclinical data,^{15–17} we hypothesized that gene expression analysis could be used to classify host responses to severe injury complicated by VAP. We tested this hypothesis using data from an existing cohort of 158 intubated trauma patients to validate the ability of these 85 genes to detect VAP and to predict its occurrence days previous to the clinical diagnosis. The patient data are housed in a Trauma-related Database (TRDB)¹⁸ created by the Inflammation and Host Response to Injury Program (Trauma Glue Grant),¹⁹ a 10-year, large-scale collaborative research program funded by the National Institute of General Medical Sciences. Sample collection and processing protocols had been developed in a series of experiments that sought to validate the informational value of changes in the relative abundance of mRNA in circulating leukocytes.^{20,21}

MATERIALS AND METHODS

Patients

An existing TRDB data set derived from 166 patients with severe blunt trauma was used to test the hypothesis that riboleukograms can detect VAP and predict its occurrence days prior to clinical diagnosis.²² Access to TRDB was provided in compliance with an IRB protocol approved by Washington University in St. Louis, in accordance with University and federal requirements for access to protected patient information.

Of the 166 patients in TRDB, 158 patients were at risk for VAP, that is, they were intubated and mechanically ventilated at some point after injury. Each of these patients was classified into “VAP” and “non-VAP” categories. The diagnosis of VAP was made using criteria consistent with the consensus criteria published by the American Thoracic Society,^{9,23} with the exception that a diagnosis of VAP could be made by the attending physician with patients on the ventilator <48 hours. Specifically, a classification of “VAP” was made if the patient was on a ventilator and culture results were positive from secretions obtained by invasive diagnostics, either bronchoalveolar lavage (BAL), protected specimen brushing, or

aspiration of tracheal secretions. The remaining subjects in the 158 patient cohort were classified as “non-VAP.” Clinical parameters were compared using the statistical software package SAS (SAS Institute, Inc, Cary, NC). Clinical Pulmonary Infection Score (CPIS) variables were analyzed using random effects analysis of variance to test each variable for an aggregate time effect, a group effect, and a group by time interaction.

Gene Expression Values

In addition to clinical information, TRDB also contains circulating leukocyte gene expression data collected from patients at frequent points after injury: 0.5, 1, 4, 7, 14, 21, and 28 days.²² The methods used to obtain these cells (buffy coat) from blood and to generate transcriptional profiles have been reported previously.^{20,21} The Affymetrix U133 2.0+ GeneChip system was used to determine hybridization intensity and relative changes in RNA abundance (gene expression). dChip-normalized expression values²⁴ for 54,613 probe sets were available in TRDB to test our hypotheses.

Validation of the 85-Genes List to Detect VAP

In an earlier report, a list of 85 genes from buffy coat was used to create riboleukograms that mapped the clinical trajectories of patients who developed VAP as a complication of critical illness or injury.¹³ In the current study, gene expression data for these 85 genes was used to determine the association of gene expression with initial onset of VAP, employing repeated measures logistic regression using generalized estimating equations (GEE, PROC GENMOD, SAS Version 9.1). It was considered clinically relevant to assess the odds of initial VAP onset within 3 or 4 days of each sampling time point, balancing the risks of delayed antibiotic treatment in those with VAP (detecting signal too late) versus (unnecessary) antibiotic treatment in those without VAP (trying to detect signal and therefore treat too early). Patients were not considered to be at risk for VAP after day of initial VAP onset or on the day of death. The GEE model included a term for gene expression and sample day, the latter to account for any general temporal trend due simply to recovery from severe injury. To account for multiple observations per patient, “patient” was treated as a cluster effect with an exchangeable correlation structure. Each of the 85 genes was analyzed separately. A principal component plot of the 85-gene riboleukogram was performed using methods reported previously.¹³ A permutation test was used to test the statistical significance of the observed spread between the VAP and non-VAP trajectories in this plot. To carry out the test, each gene was first standardized to zero mean and unit variance to remove any difference in scale across the 85 genes. Patients were randomly permuted between the VAP and non-VAP groups and the Euclidean distance between the 2 trajectories was calculated at each of the 7 time points as a measure of spread between the 2 trajectories. The *P* values for each time point are reported both unadjusted and Bonferroni-adjusted to account for the 7 tests. In addition, the maximum of the 7 Euclidean distances was computed as a global test of any spread between the 2 trajectories. The *P* values are based on 2000 random permutations.

Testing of 85 Genes to Predict VAP

The 158 patient cohort was randomly divided into a training set (2/3, 105 patients) and test set (1/3, 53 patients), equally balanced between VAP and non-VAP patients. The training and test set approach were used to explore a variety of predictive models while avoiding the danger of over-fitting. The test set of 53 patients was reserved to estimate the prediction accuracy of the single model found to have the best cross-validation error rate within the training set.

A variety of models were considered to predict the onset of VAP within 3 or 4 days of each time point; the optimal “prediction window” is unknown (Table 1). As a general approach,

the expression value, together with a vector of features specific to the sample day (rather than the patient) was treated as the unit of analysis. Ignoring within patient correlation between samples should tend to inflate cross-validation prediction accuracy. Thus, a model with poor cross-validation prediction accuracy was (safely) rejected. The longitudinal gene expression data available provided both a conceptual and practical challenge to predictive modeling. We considered 2 methods of summarizing the longitudinal gene expression at each sample day to use as an additional potential predictive feature. The first method was to simply calculate, for each gene, the gene expression slope from baseline to the current sample day for each patient and each sample using simple linear regression (referred to as the “slope”). This slope estimates the average change in gene expression per unit time since time 0. Thus, we have 170 features on each sample day to use in the prediction model; 85 genes each having a slope and the current gene expression value (we will call this “expression”). The second method to summarize longitudinal gene expression is more complex and required artificially aligning the sample data to fall exactly on study days 0, 1, 4, and 7. Next, for each gene and each of study days 0, 1, 4, and 7 a logistic regression model was fit to predict VAP onset following that day. This model included a separate coefficient for each study day’s gene expression.

From each of these models we compute the estimated probability of VAP onset for each gene, for each patient, on each study day (the “score”). The 85 scores for each sample are then used as the feature vector in the machine learning algorithms to predict VAP onset. Note that for each gene, the score is a complete summary of the (linear) relationship between prior gene expression history and VAP onset at each study day. These scores should be recalculated within the retained folds during *k*-fold cross validation; however, we ignored this step as it increases cross-validation accuracy, allowing us to safely reject models with poor cross-validation accuracy. The single model that was found to have the best cross-validation error rate within the training set thereafter was applied to the test set to estimate the prediction error rate.

De Novo Leukocyte Gene Expression Analysis and Predictive Testing in 158 Patients

The analysis described above was on the basis of the use of the previously reported list of 85 genes used to generate riboleukograms for a small, heterogeneous cohort of critically ill or injured patients.¹³ To leverage the statistical power of the much larger cohort of 158 patients, we also performed de novo analysis on the existing leukocyte gene expression data in TRDB using the same model tested above on 85 gene transcripts. Specifically, the Compound Covariate Predictor model²⁵ was applied to the list of 3821 probe sets that had a coefficient of variation of 0.5 or greater in the training data set, using both gene expression and slope as parameters to predict VAP within a 4-day window in the test data set. A second principle component analysis plot was made based upon the genes found to be significantly associated with VAP onset in this de novo analysis. Since these genes were chosen based on their observed significance in this cohort, further significance (permutation) testing of the observed spread between the VAP and non-VAP trajectories was not performed.

RESULTS

Patient Characteristics

Demographic information for the 158 intubated subjects in the TRDB is included in Table 2. Fifty 7 (36.1%) of 158 patients developed VAP, diagnosed by BAL in 41 (72%) and protected specimen brushings in 5 (9%) of the cases. The remaining 19% of VAP cases were diagnosed by tracheal aspirates. Death after injury (4% overall) did not differ between VAP and non-VAP groups. The average day of VAP diagnosis was 6.9 days (with a standard deviation of 4.5 days); the range was 1 to 18 days after injury (Fig. 1). Note that of those in

the VAP cohort, the number of patients decreases over time while the proportion of those who have been diagnosed with VAP increases (Table 3). Four patients were taken off the ventilator for 3 days or less with bronchial or tracheal secretions that were culture positive and obtained before extubation; these patients were classified as VAP. One patient on the ventilator had secretions culture positive for *Candida* species; that patient was classified as non-VAP. Age and geographic ancestry did not differ between the VAP and non-VAP groups. Similarly, there was no difference in the frequency of VAP across the 7 clinical sites. However, patients who developed VAP tended to be men and were more severely injured. Gram-negative bacteria were the most common organism class cultured from the patients who developed VAP (58% of culture-positive isolates, including *Enterobacter*, *Acinetobacter*, and *Hemophilus*), but *Staphylococcus aureus* was the most common species isolated (26%). Antibiotic data are not available from TRDB. Plots of the CPIS variables used commonly to diagnose VAP (temperature, WBC, and P/F ratios)²⁶ are presented in Figure 2. All 3 variables show a highly significant aggregate change over time ($P < 0.0001$), while temperature showed a significant group-by-time interaction ($P = 0.0378$) and P/F ratio showed a significant group effect ($P = 0.0006$).

Validation of 85 Genes and Prediction of VAP

For the purpose of qualitative comparisons, riboleukogram principal component plots were made using the 85 gene transcripts associated previously with the onset of VAP¹³ (Fig. 3A). The VAP and non-VAP trajectories appeared to bifurcate at some point after day 1, separating maximally between days 4 and 7, coincident with the average day of VAP diagnosis in this patient cohort. Coincident with initiation of antibiotic therapy, the curves converged during the remainder of the study, out to 28 days after injury. The single time point for the normal volunteers appears to be in line with the healing trajectory of both groups of patients. The global permutation test for these trajectories showed a weak trend towards significance ($P = 0.11$). However, the difference between the trajectories at day 4 was highly significant ($P < 0.0001$), both adjusted and unadjusted. The difference at day 7 showed a trend towards significance adjusted ($P = 0.0805$) but was significant unadjusted ($P = 0.0115$). Moreover, 27 (32%) of the 85 genes were significantly associated with VAP in the current study at an FDR of 0.05 (Table 4).

As 36% of the subjects in this 158 patient cohort developed VAP, the default prediction of “non-VAP” would be correct in the majority of cases (85% of the samples 3–4 days before diagnosis were “non-VAP”). Thus, in the absence of other accuracy measures, total accuracy must exceed this 85% rate for the model to be considered a good predictor. Using the previously reported list of 85 genes,¹³ 3- and 4-day prediction windows before VAP appeared to be equally limited; similarly, adding a time-element (slope) did not appear to help. Of the several models tested (Table 1), the single best model appeared to be the Compound Covariate Predictor using expression and slope as features and the 4-day prediction window. Using the training patient cohort, the model’s accuracy was estimated at 68% (32/47) sensitivity, 74% (194/263) specificity, 32% (32/101) positive-predictive value, and 93% (194/209) negative predictive value. The accuracy of this model applied to the test cohort data was 48% (11/23) sensitivity, 65% (84/129) specificity, 20% (11/56) positive predictive value, and 88% (84/96) negative predictive value. These results are not significantly better than chance (flipping a weighted coin).

De novo Leukocyte Gene Expression Analysis and Prediction of VAP

In light of the small sample size (11 patients) used to generate the list of 85 gene transcripts previously associated with VAP,¹³ we tested the potential of this larger cohort of 158 patients to generate a more robust predictor using the same Compound Covariate Predictor model (expression, slope, and the 4-day prediction window). A simple coefficient-of-

variation filter of greater than or equal to 0.5 was used to reduce the number from 54,613 probe sets to 3821 gene transcripts; of these, the model chose 1837 to predict the onset of VAP. The accuracy of this model applied to the test cohort data was 57% (13/23) sensitivity, 69% (89/129) specificity, 25% (13/53) positive predictive value, and 90% (89/99) negative predictive value. These results are significantly better than chance ($P < 0.05$), in contrast to the results obtained above using the 85 gene list derived from the pilot 11-patient cohort.

Gene Biology

Of the 27 gene transcripts validated in this independent cohort of 158 patients, the majority of the gene products are associated with early neutrophil activation in response to (and for the subsequent clearance of) bacterial infections. Five of these 27 genes were also in the list of 1837 genes that were included in the Compound Covariate Predictor model: neurogranin, NLR family apoptosis protein (NAIP), alpha-1/alpha-3 defensins (DEFA1/DEFA3), carcinoembryonic antigen-related cell adhesion molecule 8 (CEACAM8), and ATPase, H⁺ transporting, lysosomal 42kDa, V1 subunit C1 (ATP6V1C1).

DISCUSSION

We hypothesized that circulating leukocyte gene expression analysis could be used to detect VAP as a complication in severely injured patients days prior to the clinical diagnosis. Using methods reported previously to explore qualitative differences in expression,¹³ 85-gene riboleukograms were generated on an independent cohort of 158 patients, showing that, on average, the immunoinflammatory trajectories of patients who developed VAP as a complication bifurcated shortly after injury from those who did not develop VAP. Twenty-seven (32%) of the 85 genes tested were validated to detect VAP, but these genes were no better than chance in predicting VAP 3 to 4 days before the clinical diagnosis in a test cohort of 1/3 of the 158 patients. Likewise, clinical (CPIS) variables commonly used to diagnosis VAP were altered significantly in these patients but have been reported to have similarly unreliable predictive accuracy.⁹ In contrast, VAP prediction using the same model based upon de novo analysis of the gene expression data from the 158 patients was significantly better than chance, although this level of diagnostic accuracy is not clinically useful. Specifically, a compound covariate predictor model using 1837 genes produced 57% sensitivity and 69% specificity in the test cohort, providing a 25% positive predictive value and 90% negative predictive value 4 days prior to clinical diagnosis.

These gene expression findings, derived from the largest cohort of critically ill patients yet reported, are consistent with published reports on the use of transcriptional profiling to improve diagnostics. For example, riboleukograms (that is, relative changes in circulating leukocyte transcriptional profiles) can be used to track the human systemic response over time to an inflammatory or infectious challenge.^{13–15} Attempts followed using gene expression profiles to provide molecular diagnostics proof-of-feasibility for acute infections in both children and adults, discriminating between infectious and noninfectious states with accuracies ranging from 80% to 95% in validation cohorts.^{27–33} We also reported that transcriptional profiles of circulating leukocytes collected shortly after severe injury were significantly associated with adverse outcomes that developed later during hospitalization, including multiple organ failure, duration of ventilation, length of hospital stay, and infection rate.²² The current report tests the predictive ability of a list of genes to diagnose acute infection using a rigorous experimental design: microarray data derived from an independent patient cohort, collected across multiple centers, by investigators who employed a different microarray platform. The smaller list of 85 genes derived from a small, heterogeneous group of patients¹³ failed to accurately predict VAP 3 to 4 days before infection, perhaps not surprisingly. In contrast, the list of 1837 genes derived using a larger, more homogenous training cohort was better than chance in diagnosing VAP applying the

same predictive model to the same test cohort of patients. The importance of larger patient cohorts to improving patient classification using gene expression profiles is well described.^{34,35}

Of interest, both the 85 and 1837 gene list predictive models were more successful in ruling out, as compared with ruling in, VAP in the test cohort. Likewise, gene expression profiles of peripheral blood mononuclear cells were first used to rule out (not rule in) moderate to severe graft rejection in patients after cardiac transplantation.³⁶ This may stem in part from the fact that for the diagnosis of complications, always predicting “not present” will be correct in the majority of cases. Nevertheless, with additional experience, circulating leukocyte gene expression profiling has proven to be accurate in devising low-, intermediate- and high-risk categories for prediction of future cardiac transplant rejection, permitting development of discrete surveillance strategies, and an food and drug administration-approved device.^{37,38} The field of cancer experienced a similar evolution with microarray-based diagnostics, starting with the observation that the published lists of genes touted as predictive of cancer recurrence were highly unstable and that the gene expression signatures depended strongly on patient selection in the training cohorts (at the time, 5 of 7 published gene lists performed no better than flipping a coin).³⁴ Yet today, as a result of more rigorous experimental designs, new computational approaches, and prospective validation, panels of 21³⁹ or 70⁴⁰ genes can be used to assess the likelihood of breast cancer recurrence. These examples, drawn from the heart transplant and cancer literature provide a roadmap for demonstrating the clinical efficacy of leukocyte gene expression diagnostics for sepsis. Solutions are needed for the ongoing challenges (roadblocks) of cost, variance in blood processing protocols, technology standardization, underpowered single-center studies, and independent prospective validation.¹⁴

The principal component plot of the riboleukogram (Fig. 3), used as qualitative tool to explore global effects, also provides novel biologic insight into the human systemic inflammatory response, both at the organism and molecular levels. For instance, both the VAP and non-VAP trajectories started similarly early after injury, despite significant differences at baseline (time = 0.5 days) in gender, APACHE II score, and the degree of lung injury (VAP patients were more frequently men and had higher injury scores). The VAP and non-VAP trajectories diverged significantly thereafter, several days prior to the average day of clinical diagnosis of VAP during a well-documented period of immunosuppression,^{41–44} suggesting the possibility for earlier VAP diagnosis and treatment with antibiotics. Coincident with the initiation of antibiotic therapy, the trajectories converged and appeared to be headed towards a point of immune recovery, as depicted by the single point representing the 28 volunteers. In addition, as compared with the non-VAP trajectory, points on the VAP trajectory on average are at a greater distance from those of the normal volunteers, consistent with the previously reported state of immune health, that is, an “immune attractor” for the critically ill and injured.¹³ These observations lead us to speculate that riboleukograms may one day be used at the bedside not only as a quantitative measure of how far a given patient is from immune recovery but also to optimally guide initiation or discontinuation of antimicrobial therapy.

At the molecular level, gene expression analysis has been applied successfully to study generic and pathogen-specific human leukocyte responses *in vitro*.⁴⁵ In addition, network analysis has helped focus studies on the role of individual genes on outcome in mouse models of sepsis (for example, *bcl-2* and *bim*).^{46,47} However, genotypic and phenotypic variance in humans, the diagnostic challenge of sepsis, and differences in experimental design have made the application of these approaches more challenging for patient studies.^{13,48,49} For instance, gene expression analysis of circulating (mixed) leukocyte populations found little to no overlap of gene lists from septic adult patients after injury (54 genes),²⁸

pediatric patients with acute infections (61 genes),²⁷ and septic critically ill adults (85 genes¹³ and 138 genes³¹). Some investigators have reported the ability to distinguish between infecting pathogens in patients (eg, viral versus bacterial infections)²⁷; other have not.^{13,30,31} In the current study, we used a rigorous, conservative approach to validate 27 genes previously associated with the human response to VAP as a complication of critical illness and injury¹³; 5 of these 27 genes were also in the list of 1837 genes that predicted VAP in the current study. Three genes (NAIP, CEACAM8, and the alpha-defensins) have well-characterized roles in the host response to acute infection. NAIP is a mediator of acute inflammatory disease that activates caspase-1 and leads to IL-1 β secretion and apoptosis.⁵⁰ Ligation of membrane-bound CEACAM8 (CD66b) activates neutrophils via β 2-integrins (CD11/CD18) and regulates recruitment and adhesion to endothelial cells during acute inflammation.⁵¹ α -Defensins (also known as human neutrophil peptides) are antimicrobial peptides stored in neutrophil granules that play important roles in host defense (eg, leukocyte chemotaxis and adhesion), after neutrophil activation and degranulation.⁵² The other 2 genes are less well characterized. Neurogranin is a small, calmodulin accessory protein that regulates calcium signaling by altering calmodulin-dependent activation, critical for nitric oxide synthase activity, as an example.⁵³ Despite the importance of calmodulin to immune homeostasis, role of neurogranin has not been well studied. However, 1 group reported recently that changes in mRNA abundance of neurogranin are associated with B-cell activation.⁵⁴ Finally, ATPase, H⁺ transporting, lysosomal 42kDa, V1 subunit C1 (ATP6V1C1) is a component of the multisubunit enzyme that mediates acidification of intracellular compartments of eukaryotic cells, including receptor-mediated endocytosis.⁵⁵ ATP6V1C1 has not been associated previously with infection in published reports, but the NIH Gene Expression Omnibus repository links this gene to microarray data from *Chlamydia* and respiratory syncytial virus pneumonia experiments.⁵⁶

There are a number of important limitations to our study, the most important of which is that the TRDB data set was derived from a trauma study that was not designed to investigate pneumonia (or sepsis). As a result, data collection was relatively sparse (especially after the first 7 days) and no information regarding antibiotic treatment was available. In addition, the predictive models do not take into account that the case mix for VAP evolves over time (Table 3), and those patients who did well are insufficiently represented at later time points as are those patients who did poorly (died) after injury. We cannot rule out the influence of infections other than VAP or the confounding effect of other perturbations that alter the host systemic inflammatory response (surgery, blood transfusions, etc).

Nevertheless, further study to develop riboleukograms for the early diagnosis of occult infection is motivated by our findings, reported herein, and very recent reports by others providing additional proof-of-principle.^{31,33} In particular, the negative predictive values approaching 90% are encouraging, sparing injured patients who will not develop VAP the complications associated with inappropriate antibiotic therapy. The application of microfluidics to rapidly isolate purified circulating leukocyte populations,⁵⁷ larger patient cohorts that permit study of differences due to geographic ancestry, age, and gender,¹³ and prospective validation using rigorous classification approaches^{14,34,35} will likely provide results that are significantly more informative and robust.^{14,45} These studies should couple examination of VAP-induced changes in transcriptional profiles with changes in cellular proteomics, including injury-induced alterations in markers linked to immune competency (eg, HLA-DR).^{41,42} Finally, it is worth emphasizing that attempts to discover a single biochemical marker (eg, cytokines or DAMPS) to diagnose sepsis (including pneumonia) have been, and continue to be, unsuccessful. In conclusion, the current study provides proof-of-concept that riboleukograms are associated with the development of VAP but the variance observed using available technology precludes accurate prediction of VAP onset in individual patients.

Acknowledgments

Supported by NIH U54GM6211906 and NIH R21GM075023 (to J.P.C., Q.L., N.L., B.H.B., W.S.S., and D.J.D.), and also by NIH R01GM081524.

The authors thank their colleagues in the NIGMS Trauma Glue Grant for critical review of the results presented herein.

References

1. Safdar N, Dezfulian C, Collard HR, et al. Clinical and economic consequences of ventilator-associated pneumonia: a systematic review. *Crit Care Med*. 2005; 33:2184–2193. [PubMed: 16215368]
2. Magnotti LJ, Croce MA, Fabian TC. Is ventilator-associated pneumonia in trauma patients an epiphenomenon or a cause of death? *Surg Infect (Larchmt)*. 2004; 5:237–242. [PubMed: 15684794]
3. Centers for Disease Control and Prevention. An Overview of Ventilator-Associated Pneumonia (VAP). Atlanta, GA: Centers for Disease Control and Prevention; 2009.
4. Cocanour CS, Ostrosky-Zeichner L, Peninger M, et al. Cost of a ventilator-associated pneumonia in a shock trauma intensive care unit. *Surg Infect (Larchmt)*. 2005; 6:65–72. [PubMed: 15865552]
5. Cocanour CS, Peninger M, Domonoske BD, et al. Decreasing ventilator-associated pneumonia in a trauma ICU. *J Trauma*. 2006; 61:122–129. [PubMed: 16832259]
6. Klompas M, Platt R. Ventilator-associated pneumonia—the wrong quality measure for benchmarking. *Ann Intern Med*. 2007; 147:803–805. [PubMed: 18056666]
7. DuBose JJ, Inaba K, Shiflett A, et al. Measurable outcomes of quality improvement in the trauma intensive care unit: the impact of a daily quality rounding checklist. *J Trauma*. 2008; 64:22–27. [PubMed: 18188094]
8. Mayhall CG. Ventilator-associated pneumonia or not? Contemporary diagnosis. *Emerg Infect Dis*. 2001; 7:200–204. [PubMed: 11294706]
9. American Thoracic Society, Infectious Diseases Society of America. Guidelines for the management of adults with hospital-acquired, ventilator-associated, and healthcare-associated pneumonia. *Am J Respir Crit Care Med*. 2005; 171:388–416. [PubMed: 15699079]
10. Torres A, Ewig S. Diagnosing ventilator-associated pneumonia. *N Engl J Med*. 2004; 350:433–435. [PubMed: 14749448]
11. Kumar A, Roberts D, Wood KE, et al. Duration of hypotension before initiation of effective antimicrobial therapy is the critical determinant of survival in human septic shock. *Crit Care Med*. 2006; 34:1589–1596. [PubMed: 16625125]
12. Kollef KE, Schramm GE, Wills AR, et al. Predictors of 30-day mortality and hospital costs in patients with ventilator-associated pneumonia attributed to potentially antibiotic-resistant gram-negative bacteria. *Chest*. 2008; 134:281–287. [PubMed: 18682456]
13. McDunn JE, Husain KD, Polpitiya AD, et al. Plasticity of the systemic inflammatory response to acute infection during critical illness: development of the riboleukogram. *PLoS One*. 2008; 3:e1564. [PubMed: 18270561]
14. Polpitiya AD, McDunn JE, Burykin A, et al. Using systems biology to simplify complex disease: immune cartography. *Crit Care Med*. 2009; 37(suppl 1):S16–S21. [PubMed: 19104218]
15. Calvano SE, Xiao W, Richards DR, et al. A network-based analysis of systemic inflammation in humans. *Nature*. 2005; 437:1032–1037. [PubMed: 16136080]
16. Laudanski K, Miller-Graziano C, Xiao W, et al. Cell-specific expression and pathway analyses reveal novel alterations in trauma-related human T-cell and monocyte pathways. *Proc Natl Acad Sci USA*. 2006; 103:15564–15569. [PubMed: 17032758]
17. Lederer JA, Brownstein BH, Lopez MC, et al. Comparison of longitudinal leukocyte gene expression after burn injury or trauma-hemorrhage in mice. *Physiol Genomics*. 2008; 32:299–310. [PubMed: 17986522]
18. Trauma-related Database (TRDB). Massachusetts General Hospital; Boston, MA: 2009.

19. National Institute of General Medical Science. Inflammation and the Host Response to Injury Program (Trauma Glue Grant). Bethesda, MD: National Institute of General Medical Science; 2009.
20. Feezor RJ, Baker HV, Mindrinos M, et al. Whole blood and leukocyte RNA isolation for gene expression analyses. *Physiol Genomics*. 2004; 19:247–254. [PubMed: 15548831]
21. Cobb JP, Mindrinos M, Miller-Graziano C, et al. Application of genome-wide expression analysis to human health and disease. *Proc Natl Acad Sci USA*. 2005; 102:4801–4806. [PubMed: 15781863]
22. Warren HS, Elson CM, Hayden DL, et al. A genomic score prognostic of outcome in trauma patients. *Mol Med*. 2009; 15:220–227. [PubMed: 19593405]
23. Evans HL, West MA, Cuschieri J, et al. Inflammation and the host response to injury, a large-scale collaborative project: patient-oriented research core standard operating procedures for clinical care IX. Definitions for complications of clinical care of critically injured patients. *J Trauma*. 2009 In press.
24. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. *Proc Natl Acad Sci USA*. 2001; 98:31–36. [PubMed: 11134512]
25. Radmacher MD, McShane LM, Simon R. A paradigm for class prediction using gene expression profiles. *J Comput Biol*. 2002; 9:505–511. [PubMed: 12162889]
26. Pugin J, Auckenthaler R, Mili N, et al. Diagnosis of ventilator-associated pneumonia by bacteriologic analysis of bronchoscopic and nonbronchoscopic “blind” bronchoalveolar lavage fluid. *Am Rev Respir Dis*. 1991; 143:1121–1129. [PubMed: 2024824]
27. Ramilo O, Allman W, Chung W, et al. Gene expression patterns in blood leukocytes discriminate patients with acute infections. *Blood*. 2007; 109:2066–2077. [PubMed: 17105821]
28. Johnson SB, Lissauer M, Bochicchio GV, et al. Gene expression profiles differentiate between sterile SIRS and early sepsis. *Ann Surg*. 2007; 245:611–621. [PubMed: 17414611]
29. Tang BM, McLean AS, Dawes IW, et al. The use of gene-expression profiling to identify candidate genes in human sepsis. *Am J Respir Crit Care Med*. 2007; 176:676–684. [PubMed: 17575094]
30. Tang BM, McLean AS, Dawes IW, et al. Gene-expression profiling of gram-positive and gram-negative sepsis in critically ill patients. *Crit Care Med*. 2008; 36:1125–1128. [PubMed: 18379237]
31. Tang BM, McLean AS, Dawes IW, et al. Gene-expression profiling of peripheral blood mononuclear cells in sepsis. *Crit Care Med*. 2009; 37:882–888. [PubMed: 19237892]
32. Lissauer ME, Johnson SB, Bochicchio GV, et al. Differential expression of toll-like receptor genes: sepsis compared with sterile inflammation 1 day before sepsis diagnosis. *Shock*. 2009; 31:238–244. [PubMed: 18665047]
33. Howrylak JA, Dolinay T, Lucht L, et al. Discovery of the gene signature for acute lung injury in patients with sepsis. *Physiol Genomics*. 2009; 37:133–139. [PubMed: 19174476]
34. Michiels S, Koscielny S, Hill C. Prediction of cancer outcome with microarrays: a multiple random validation strategy. *Lancet*. 2005; 365:488–492. [PubMed: 15705458]
35. Ioannidis JP. Microarrays and molecular research: noise discovery? *Lancet*. 2005; 365:454–455. [PubMed: 15705441]
36. Deng MC, Eisen HJ, Mehra MR, et al. Noninvasive discrimination of rejection in cardiac allograft recipients using gene expression profiling. *Am J Transplant*. 2006; 6:150–160. [PubMed: 16433769]
37. Mehra MR, Kobashigawa JA, Deng MC, et al. Clinical implications and longitudinal alteration of peripheral blood transcriptional signals indicative of future cardiac allograft rejection. *J Heart Lung Transplant*. 2008; 27:297–301. [PubMed: 18342752]
38. XDX Expression Diagnostics. 2009.
39. Genomic Health, Inc. 2009.
40. Agendia. 2009.
41. Naziri W, Cheadle WG, Pietsch JD, et al. Pneumonia in the surgical intensive care unit. Immunologic keys to the silent epidemic. *Ann Surg*. 1994; 219:632–640. [PubMed: 8203972]

42. Heinzelmann M, Mercer-Jones M, Cheadle WG, et al. CD14 expression in injured patients correlates with outcome. *Ann Surg.* 1996; 224:91–96. [PubMed: 8678624]
43. Tarlowe MH, Duffy A, Kannan KB, et al. Prospective study of neutrophil chemokine responses in trauma patients at risk for pneumonia. *Am J Respir Crit Care Med.* 2005; 171:753–759. [PubMed: 15618463]
44. Turina M, Dickinson A, Gardner S, et al. Monocyte HLA-DR and interferon-gamma treatment in severely injured patients—a critical reappraisal more than a decade later. *J Am Coll Surg.* 2006; 203:73–81. [PubMed: 16798489]
45. Jenner RG, Young RA. Insights into host responses against pathogens from transcriptional profiling. *Nat Rev Microbiol.* 2005; 3:281–294. [PubMed: 15806094]
46. Wagner TH, Drewry AM, MacMillan SK, et al. Surviving sepsis: *bcl-2* overexpression modulates splenocyte transcriptional responses in vivo. *Am J Physiol Regul Integr Comp Physiol.* 2007; 292:R1751–R1759. [PubMed: 17234957]
47. Chang KC, Unsinger J, Davis CG, et al. Multiple triggers of cell death in sepsis: death receptor and mitochondrial-mediated apoptosis. *FASEB J.* 2007; 21:708–719. [PubMed: 17307841]
48. Martin GS, Mannino DM, Eaton S, et al. The epidemiology of sepsis in the United States from 1979 through 2000. *N Engl J Med.* 2003; 348:1546–1554. [PubMed: 12700374]
49. Barnato AE, Alexander SL, Linde-Zwirble WT, et al. Racial variation in the incidence, care, and outcomes of severe sepsis: analysis of population, patient, and hospital characteristics. *Am J Respir Crit Care Med.* 2008; 177:279–284. [PubMed: 17975201]
50. Wilmanski JM, Petnicki-Ocwieja T, Kobayashi KS. NLR proteins: integral members of innate immunity and mediators of inflammatory diseases. *J Leukoc Biol.* 2008; 83:13–30. [PubMed: 17875812]
51. Skubitz KM, Skubitz AP. Interdependency of CEACAM-1, -3, -6, and -8 induced human neutrophil adhesion to endothelial cells. *J Transl Med.* 2008; 6:78. [PubMed: 19077207]
52. Quinn K, Henriques M, Parker T, et al. Human neutrophil peptides: a novel potential mediator of inflammatory cardiovascular diseases. *Am J Physiol Heart Circ Physiol.* 2008; 295:H1817–H1824. [PubMed: 18805897]
53. Slemmon JR, Feng B, Erhardt JA. Small proteins that modulate calmodulin-dependent signal transduction: effects of PEP-19, neuromodulin, and neurogranin on enzyme activation and cellular homeostasis. *Mol Neurobiol.* 2000; 22:99–113. [PubMed: 11414283]
54. Glynne R, Ghandour G, Rayner J, et al. B-lymphocyte quiescence, tolerance and activation as viewed by global gene expression profiling on microarrays. *Immunol Rev.* 2000; 176:216–246. [PubMed: 11043780]
55. Entrez Gene. Novus Biologicals. 2009. ATP6V1C1 ATPase, H⁺ transporting, lysosomal 42kDa, V1 subunit C1.
56. National Center for Biotechnology Information. Gene Expression Omnibus. 2009.
57. Russom A, Sethu P, Irimia D, et al. Microfluidic leukocyte isolation for gene expression analysis in critically ill hospitalized patients. *Clin Chem.* 2008; 54:891–900. [PubMed: 18375483]

FIGURE 1.

Frequency and etiology of 57 cases of ventilator-associated pneumonia (VAP) after severe injury. Frequency of VAP on postinjury day is depicted, as well as data broken down by nature of the bacteria responsible for infection. See Table 3 for changes over time in the VAP case mix.

FIGURE 2.

Change in Clinical Pulmonary Infection Score variables used commonly to make the diagnosis of pneumonia.²⁶ The 3 objective variables of the Clinical Pulmonary Infection Score are plotted for time points prior to the average day of diagnosis, 6.9 days. All 3 variables showed significant changes; see text for details.

FIGURE 3.

Riboleukogram principal component plots of (A) 85 genes and (B) 1837 genes used to qualitatively track the immunoinflammatory trajectories of patients for 28 days after severe injury. In each graph of the second and third principal components, the black solid curves are derived from patients without VAP and the blue dotted curves are derived from patients with VAP. The number of patients in each cohort and the evolving VAP case mix at each time point is shown in Table 3; the data are the same for panel B. Red and green denote the beginning (0.5 days) and end (28 days) of the study, respectively. The yellow triangles depict gene expression values in normal volunteers. Note that the graphs separate maximally and significantly days prior to the average day of VAP diagnosis (6.9 days), depicted as “VAP.” Coincident with initiation of antibiotic therapy, the trajectories converge subsequently (days 21 and 28) and appear to head towards a “normal” state (immune homeostasis).

TABLE 1**Models Used to Test the Predictive (Diagnostic) Potential of 85 Leukocyte Genes***

Stepwise Logistic Regression (SAS)
Linear and Quadratic Diagonal Discriminant Analysis (SAS)
K-Nearest Neighbor (SAS)
Classification and Regression Trees (R)
Support Vector Machines (R)
1-Nearest Neighbor (BRB-Array Tools)
3-Nearest Neighbors (BRB-Array Tools)
Diagonal Linear Discriminant Analysis (BRB-Array Tools)
Nearest Centroid Predictor (BRB-Array Tools)
Compound Covariate Predictor (BRB-Array Tools)

* The software program used to perform the test is included in parentheses.

TABLE 2

Patient Characteristics

	VAP n = 57	No VAP n = 101	P
Age mean (SD) - Wilcoxon*	35.3 (11.1)	33.6 (11.1)	0.335
AIS-Thoracic mean (SD) [†]	2.8 (1.7)	2.0 (1.8)	0.017
Male number (%)	45 (79%)	58 (57%)	0.006
APACHE II SCORE (SD)	29.3 (5.6)	26.9 (5.4)	0.023
Deaths number (%)	4 (7%)	3 (3%)	0.235
VAP day mean (SD)	6.9 (4.5)		
Gram-positive infection number (% of VAP)	20 (35%)		
Gram-negative infection number (% of VAP)	33 (58%)		
Both types of infection number (% of VAP)	4 (7%)		

* 2 Sample Test.

[†] Wilcoxon 2 Sample Test.

TABLE 3

Number of Patients Remaining at Each Time Point in the VAP and non-VAP Cohorts

Time After Injury (d)	Non-VAP (n)	VAP (n)	VAP Proportion Post-VAP Onset
0.5	101	57	0.00
1	96	56	0.02
4	86	52	0.42
7	71	46	0.48
14	39	40	0.88
21	25	26	1.00
28	14	18	1.00

TABLE 4

Validated 27 Genes*

Probe Set ID	Gene and Symbol	Gene Title
200996_at	<i>ACTR3</i>	ARP3 actin-related protein 3 homolog (yeast)
202187_s_at	<i>PPP2R5A</i>	Protein phosphatase 2, regulatory subunit B', alpha isoform
202381_at	<i>ADAM9</i>	ADAM metallopeptidase domain 9 (meltrin gamma)
202530_at	<i>MAPK14</i>	Mitogen-activated protein kinase 14
202872_at	<i>ATP6V1C1</i> *	ATPase, H ⁺ transporting, lysosomal 42kDa, V1 subunit C1
202974_at	<i>MPP1</i>	Membrane protein, palmitoylated 1, 55kDa
203200_s_at	<i>MTRR</i>	5-methyltetrahydrofolate-homocysteine methyltransferase reductase
203757_s_at	<i>CEACAM6</i>	Carcinoembryonic antigen-related cell adhesion molecule 6 (non-specific cross reacting antigen)
204081_at	<i>NRGN</i> *	Neurogranin (protein kinase C substrate, RC3)
204099_at	<i>SMARCD3</i>	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily d, member 3
204860_s_at	<i>NAIP</i> /// <i>LOC652755</i> *	NLR family, apoptosis inhibitory protein /// similar to baculoviral IAP repeat-containing protein 1 (Neuronal apoptosis inhibitory protein)
205033_s_at	<i>DEFA1</i> /// <i>DEFA3</i> /// <i>LOC728358</i> *	Defensin, alpha 1 /// defensin, alpha 3, neutrophil-specific /// defensin, alpha 1
205513_at	<i>TCN1</i>	Transcobalamin I (vitamin B12 binding protein, R binder family)
205557_at	<i>BPI</i>	Bactericidal/permeability-increasing protein
206493_at	<i>ITGA2B</i>	Integrin, alpha 2b (platelet glycoprotein IIb of IIb/IIIa complex, antigen CD41)
206676_at	<i>CEACAM8</i> *	Carcinoembryonic antigen-related cell adhesion molecule 8
206838_at	<i>TBX19</i>	T-box 19
207269_at	<i>DEFA4</i>	Defensin, alpha 4, corticostatin
209288_s_at	<i>CDC42EP3</i>	CDC42 effector protein (Rho GTPase binding) 3
209369_at	<i>ANXA3</i>	Annexin A3
210140_at	<i>CST7</i>	Cystatin F (leukocystatin)
210254_at	<i>MS4A3</i>	Membrane-spanning 4-domains, subfamily A, member 3 (hematopoietic cell-specific)
211963_s_at	<i>ARPC5</i>	Actin related protein 2/3 complex, subunit 5, 16kDa
212063_at	<i>CD44</i>	CD44 molecule (Indian blood group)
214146_s_at	<i>PPBP</i>	Pro-platelet basic protein (chemokine (C-X-C motif) ligand 7)
214177_s_at	<i>PBXIP1</i>	Pre-B-cell leukemia homeobox interacting protein 1
39402_at	<i>IL1B</i>	Interleukin 1, beta

* The row indicates that genes are also in the list of 1837 gene derived from 158 patients.