

An Evaluation of Antimicrobial Prescribing and Risk-adjusted Mortality

Jonathan H. Pelletier, MD*; Alicia K. Au, MD, MS*; Dana Fuhrman, DO, MS*; James Zullo, MS†; Ann E. Thompson, MD, MHCPM*; Robert S.B. Clark, MD*; Christopher Horvat, MD, MHA*

Abstract

Introduction: The Centers for Disease Control and Prevention recommends tracking risk-adjusted antimicrobial prescribing. Prior studies have used prescribing variation to drive quality improvement initiatives without adjusting for severity of illness. The present study aimed to determine the relationship between antimicrobial prescribing and risk-adjusted ICU mortality in the Pediatric Health Information Systems (PHIS) database, assessed by IBM-Watson risk of mortality. A nested analysis sought to assess an alternative risk model incorporating laboratory data from federated electronic health records. **Methods:** Retrospective cohort study of pediatric ICU patients in PHIS between 1/1/2010 and 12/31/2019, excluding patients admitted to a neonatal ICU, and a nested study of PHIS+ from 1/1/2010 to 12/31/2012. Hospital antimicrobial prescription volumes were assessed for association with risk-adjusted mortality. **Results:** The cohort included 953,821 ICU encounters (23,851 [2.7%] nonsurvivors). There was 4-fold center-level variability in antimicrobial use. ICU antimicrobial use was not correlated with risk-adjusted mortality assessed using IBM-Watson. A risk model incorporating laboratory data available in PHIS+ significantly outperformed IBM-Watson (c-statistic 0.940 [95% confidence interval 0.933–0.947] versus 0.891 [0.881–0.901]; $P < 0.001$, area under the precision recall curve 0.561 versus 0.297). Risk-adjusted mortality was inversely associated with antimicrobial prescribing in this smaller cohort using both the PHIS+ and Watson models ($P = 0.05$ and $P < 0.01$, respectively). **Conclusions:** Antimicrobial prescribing among pediatric ICUs in the PHIS database is variable and not associated with risk-adjusted mortality as assessed by IBM-Watson. Expanding existing administrative databases to include laboratory data can achieve more meaningful insights when assessing multicenter antibiotic prescribing practices. (*Pediatr Qual Saf* 2021;6:e481; doi: 10.1097/pq9.000000000000481; Published online December 15, 2021.)

INTRODUCTION

Antimicrobial use in pediatric intensive care units (ICUs) is common; they are prescribed to 43%–79% of ICU patients.^{1–4} The Centers for Disease Control and



Prevention (CDC) cited improvement in antimicrobial prescribing practices as “perhaps the single most important action needed to greatly slow down the development and spread of antibiotic-resistant infections.”⁵

Multi-center databases such as the Pediatric Health Information System (PHIS) are increasingly used to evaluate antimicrobial use.^{2,6} Previous PHIS studies identified variability in antimicrobial prescribing in pediatric hospital and ICU settings not explained by demographic characteristics.^{2,3} These publications are being used to guide antimicrobial stewardship efforts at multiple centers.^{7–9} As an administrative database, PHIS was not designed to assess antimicrobial prescribing. It does not contain microbiological data, or documentation of clinical decision-making pertaining antibiotic prescribing.¹⁰

These publications are being used to guide antimicrobial stewardship efforts at multiple centers.^{7–9} As an administrative database, PHIS was not designed to assess antimicrobial prescribing. It does not contain microbiological data, or documentation of clinical decision-making pertaining antibiotic prescribing.¹⁰

Skillfully refining antimicrobial prescribing in the ICU setting is possible, but challenges exist when treating patients with suspicion of sepsis.¹¹ Antimicrobial courses are often prescribed empirically; approximately 40% of sepsis in the ICU is culture-negative.¹² Even when a pathogen is identified, current culture methods may not detect co-infection.¹³ The CDC recently released an “Antimicrobial Use and Resistance (AUR) Module,” in which it recommends risk-adjusted antimicrobial benchmarking.¹⁴ Including risk-adjusted mortality as a

From the *Division of Pediatric Critical Care Medicine, Department of Critical Care Medicine, Pittsburgh, Pa. UPMC Children’s Hospital of Pittsburgh, Pittsburgh, Pa.; and †Data Warehouse, UPMC Children’s Hospital of Pittsburgh.

Supplemental digital content is available for this article. Clickable URL citations appear in the text.

Presented (in abstract format) at the 2020 Society of Critical Care Conference.

*Corresponding author. Address: Christopher Horvat, Department of Pediatric Critical Care Medicine, UPMC Children’s Hospital of Pittsburgh, 4401 Penn Ave, Pittsburgh, PA 15224
PH: 412-692-5298
Email: christopher.horvat@chp.edu

Copyright © 2021 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

To cite: Pelletier JH, Au AK, Fuhrman D, Zullo J, Thompson AE, Clark RSB, Horvat C. An Evaluation of Antimicrobial Prescribing and Risk-adjusted Mortality. *Pediatr Qual Saf* 2021;6:e481.

Received for publication August 10, 2020; Accepted June 9, 2021.

Published online December 15, 2021

DOI: 10.1097/pq9.000000000000481

balancing metric in stewardship projects would improve compliance with these recommendations and ensure optimal targeting of antimicrobials to patients most likely to benefit. Prior studies demonstrating variability in antimicrobial prescribing in pediatric ICUs have been substantially limited by a lack of adjustment for illness severity.²

The PHIS database does not contain previously-validated severity-of-illness scores, nor does it contain the variables necessary to manually derive these scores.¹⁵⁻¹⁸ However, it does contain a proprietary pediatric risk of mortality score developed by IBM Watson (Watson score) that could potentially facilitate the assessment of risk-adjusted mortality.¹⁰ Additionally, the federated electronic health record data contained in PHIS+, an expansion of the PHIS database, provides an opportunity to construct a robust risk score reflecting severity of illness.¹⁹ PHIS+ was developed to add robust laboratory data from 6 large children's hospitals to the existing administrative data harbored by PHIS. A key aim of PHIS+ was to support development of "severity adjustment tools to study inpatient quality (eg, comparing hospitals, providers, systems) across multiple institutions using large administrative and clinical databases."²⁰

The present study sought to evaluate the discrimination and calibration of the Watson score in mortality prediction in PHIS over the past decade and compare hospital-level antimicrobial prescribing with risk-adjusted mortality. Because the Watson score is proprietary, its performance was evaluated against a novel severity of illness score developed using PHIS+.

METHODS

Data Source

Clinical and pharmacy data were provided by the PHIS database, an online, anonymized, quality controlled data warehouse containing clinical and resource utilization information maintained by the Children's Hospital Association (CHA).^{10,21} We obtained laboratory data from PHIS+.¹⁹

Study Design and Patient Population

This was a retrospective cohort study of patients discharged from each of the 51 PHIS participating centers between 1/1/2010 and 12/31/2019 (Total Study Cohort). We included patients admitted to a pediatric or cardiac ICU during the study period. We excluded patients admitted to a neonatal ICU. The primary exposure was prescription of 1 of the 20 common antibacterials (amoxicillin, ampicillin, ampicillin/sulbactam, azithromycin, cefazolin, cefepime, cefotaxime, cefoxitin, ceftazidime, ceftriaxone, ciprofloxacin, clindamycin, erythromycin, gentamicin, meropenem, metronidazole, piperacillin/tazobactam, sulfamethoxazole/trimethoprim, tobramycin, vancomycin) or 8 common antifungals (amphotericin, anidulafungin, caspofungin, fluconazole, itraconazole, micafungin, posaconazole, voriconazole). These medications were

chosen because they represented the great majority of total antimicrobials prescribed during the study period. We defined antimicrobial exposure as prescription of any of these antibacterials or antifungals during the intensive care unit stay. We further divided antibacterials into "antipseudomonal" (cefepime, ceftazidime, ciprofloxacin, meropenem, and piperacillin-tazobactam) and "anti-methicillin resistant *Staphylococcus aureus* (anti-MRSA)" (clindamycin, sulfamethoxazole/trimethoprim, vancomycin) modified from Brogan et al and in compliance with the Surviving Sepsis Campaign.^{2,22} The primary outcome was in-hospital mortality. The local Institutional Review Board approved the study.

Statistical Analysis

We extracted clinical and administrative data from PHIS (age, sex, race, insurance status, admission priority, primary admission diagnosis, and billing charges for mechanical ventilation, extracorporeal membrane oxygenation, dobutamine, dopamine, epinephrine, milrinone, norepinephrine, vasopressin, discharge mortality, and Watson score). We also extracted complex chronic conditions based on Feudtner and renal replacement therapy usage based on ICD codes from Waikar.^{23,24} The cohort was described using summary statistics. Variation in antimicrobial prescribing practices between centers was analyzed by chi-square test of proportions or 1-way analysis of variance (ANOVA).

We assessed Watson score discrimination using receiver operating characteristic and precision recall curves,²⁵ and calibration using the GiViTI (Italian Group for the Evaluation of Interventions in Intensive Care Medicine) belt.^{26,27} The need to recalibrate risk scores for external datasets or subpopulations of interest is widely recognized.²⁸ Because the Watson score was poorly calibrated, we randomly divided the cohort into training (50%) and testing (50%) subsets and recalibrated the Watson score using isotonic regression.²⁹ We compared the performance of the recalibrated score against the original model. We performed linear regression to analyze the relationship between antimicrobial days per 1000 ICU days with differences in center-level actual and predicted mortality. All statistical analyses were performed using R (versions 3.5.2 & 3.6, R Foundation for Statistical Computing, Vienna, Austria) and RStudio (RStudio, Boston, Mass., versions 1.1.463 & 1.2.5033). An α value of 0.05 was considered significant.

PHIS+ Nested Subset Analysis

The PHIS+ Subset included all patients in the total study cohort discharged from 1 of the 6 PHIS+ participating ICUs between January 1, 2010 and December 31, 2012 (the entire duration of PHIS+). The PHIS+ subset was described with summary statistics as above. We trained a multivariable logistic regression mortality prediction model, using laboratory data from PHIS+ (arterial partial pressure of oxygen, arterial partial pressure of

carbon dioxide, arterial pH, venous pH, lactate, creatinine, bicarbonate, potassium, white blood cell count, platelet count, bilirubin, activated partial thromboplastin time, international normalized ratio, c-reactive protein, and ferritin). Candidate variables were determined a priori based on similarity to previously validated models.^{16–18,30,31} Missing data were imputed with normal values, as is the convention for other mortality prediction scores.^{16–18,30} We generated diagnostic risk groups based on Appendix 1 of Straney et al, based on primary encounter diagnoses.¹⁸ We combined arterial and venous pH values. Creatinine values were normalized for age, based on Leteurtre et al.¹⁶ We randomly divided the cohort into training (80%) and testing (20%) subsets because of smaller numbers than in the main analysis. A multivariable, fixed-effects logistic regression model (without interaction terms) to predict discharge mortality was generated using the variables above using the training subset of the cohort. We considered variables significant if they were associated with discharge mortality with $P < 0.1$ (variables removed: age, potassium, c reactive protein, milrinone use, and renal replacement therapy). To reduce overfitting, we conducted automated backward stepwise analysis using Akaike Information Criteria, and removed additional variables based on best performance (variables removed: total parenteral nutrition use, bicarbonate, bilirubin, activated partial thromboplastin time, ferritin).³² We evaluated model discrimination and calibration on the testing subset as described above.

RESULTS

Cohort Demographics

The total study cohort included 953,821 encounters, and the PHIS+ subset included 45,706 encounters. Both cohort demographics are listed in Table 1. Briefly, the total study cohort median age was 4.4 years (interquartile range 0.9–12.6 years), with 55.1% men, and 59.0% white. Of the total study cohort, 67.2% had at least 1 complex chronic condition, and 29.0% of children were technology-dependent. An estimated 35.5% of patients received mechanical ventilation, and 1.0% were supported with ECMO. Overall survival to discharge was 97.3%.

Antimicrobial Prescriptions

The antimicrobials included in the study represented 86.3% of the total ICU antimicrobial prescriptions during the study period; their relative frequencies are shown in Supplemental Digital Content 1. (See figure 1, Supplemental Digital Content 1, which shows antimicrobial prescriptions by drug. Antimicrobials are listed in descending order of frequency. Bar heights represent antimicrobial days of therapy per 1,000 pediatric ICU days, <http://links.lww.com/PQ9/A317>.) Overall, 71.0% (individual center range 49.7% – 83.8%, $P < 0.001$) of the

Table 1. Cohort Demographics

	Total Study Cohort (%)	PHIS+ Subset (%)
No. Admissions	953821 (100%)	45706 (100%)
Median Age, y (IQR)	4.4 (0.9–12.6)	4.2 (0.8–12.3)
Men	525430 (55.1%)	25379 (55.5%)
White	562939 (59.0%)	30731 (67.2%)
Black	186006 (19.5%)	6584 (14.4%)
Asian	30660 (3.2%)	1201 (2.6%)
Pacific Islander	4869 (0.5%)	407 (0.9%)
American Indian	6176 (0.7%)	190 (0.4%)
Other race	127285 (13.3%)	5481 (12.0%)
Hispanic	174627 (18.3%)	4138 (9.0%)
Commercial insurance	362297 (38.0%)	22725 (49.7%)
Government insurance	548471 (57.5%)	20136 (44.1%)
Other payor	43053 (4.5%)	2845 (6.2%)
Any complex chronic condition	640904 (67.2%)	33794 (73.9%)
Cardiovascular condition	261524 (27.4%)	13868 (30.3%)
Gastrointestinal condition	187406 (19.7%)	9810 (21.5%)
Hematologic or immunologic condition	58956 (6.2%)	2880 (6.3%)
Malignancy	78116 (8.2%)	4163 (9.1%)
Metabolic condition	99054 (10.4%)	3737 (8.2%)
Neurologic or neuromuscular condition	214816 (22.5%)	10978 (24.0%)
Congenital or genetic defect	133828 (14.0%)	8498 (18.6%)
Renal or urologic condition	69524 (7.3%)	3621 (7.9%)
Respiratory condition	107079 (11.2%)	6344 (13.9%)
Technology dependent	276843 (29.0%)	14831 (32.5%)
Transplant recipient	26991 (2.8%)	2510 (5.5%)
Median ICU length of service, d (IQR)	2 (1–5)	2 (1–5)
Received mechanical ventilation	338933 (35.5%)	18703 (40.9%)
ECMO	10027 (1.0%)	491 (1.1%)
Survived to discharge	927970 (97.3%)	44352 (97.0%)

cohort received antimicrobials while admitted to the ICU, with 18.2% (7.0–31.7%, $P < 0.001$) receiving antipseudomonal antibacterials and 30.2% (16.5–56.1%, $P < 0.001$) receiving anti-MRSA antibacterials. Antimicrobial duration normalized to total number of pediatric ICU days at each center was heterogeneous. (See figure 2, Supplemental Digital Content 2, which shows antimicrobial prescriptions by ICU. ICUs are listed in rank order of total days of antimicrobials per 1,000 ICU days. Bar heights represent the sum of the days of therapy of the study antimicrobial days divided by the total ICU length of stay at a given hospital, <http://links.lww.com/PQ9/A318>.) Vancomycin was the most used antimicrobial across all centers (mean 181 days of therapy per 1000 ICU days). The use of anti-MRSA agents by center varied by > 4-fold (range 131–635 days of therapy per 1,000 ICU days, $P < 0.001$). The use of antipseudomonal agents by center varied by >7-fold (range 99–705 days of therapy per 1000 ICU days, $P < 0.001$). The most used antifungals were fluconazole (mean 50 days of therapy per 1000 ICU days) and echinocandins (caspofungin and micafungin; mean 33 days of therapy per 1000 ICU days). Fluconazole use varied by 6-fold (range 21–131 days of therapy per 1000 ICU days, $P < 0.001$) and echinocandin use varied by 58-fold (range 2–117 days of therapy per 1000 ICU days, $P < 0.001$).

Watson Score Discrimination and Calibration

The Watson score discrimination and calibration are illustrated in Figure 1. Overall, before and after

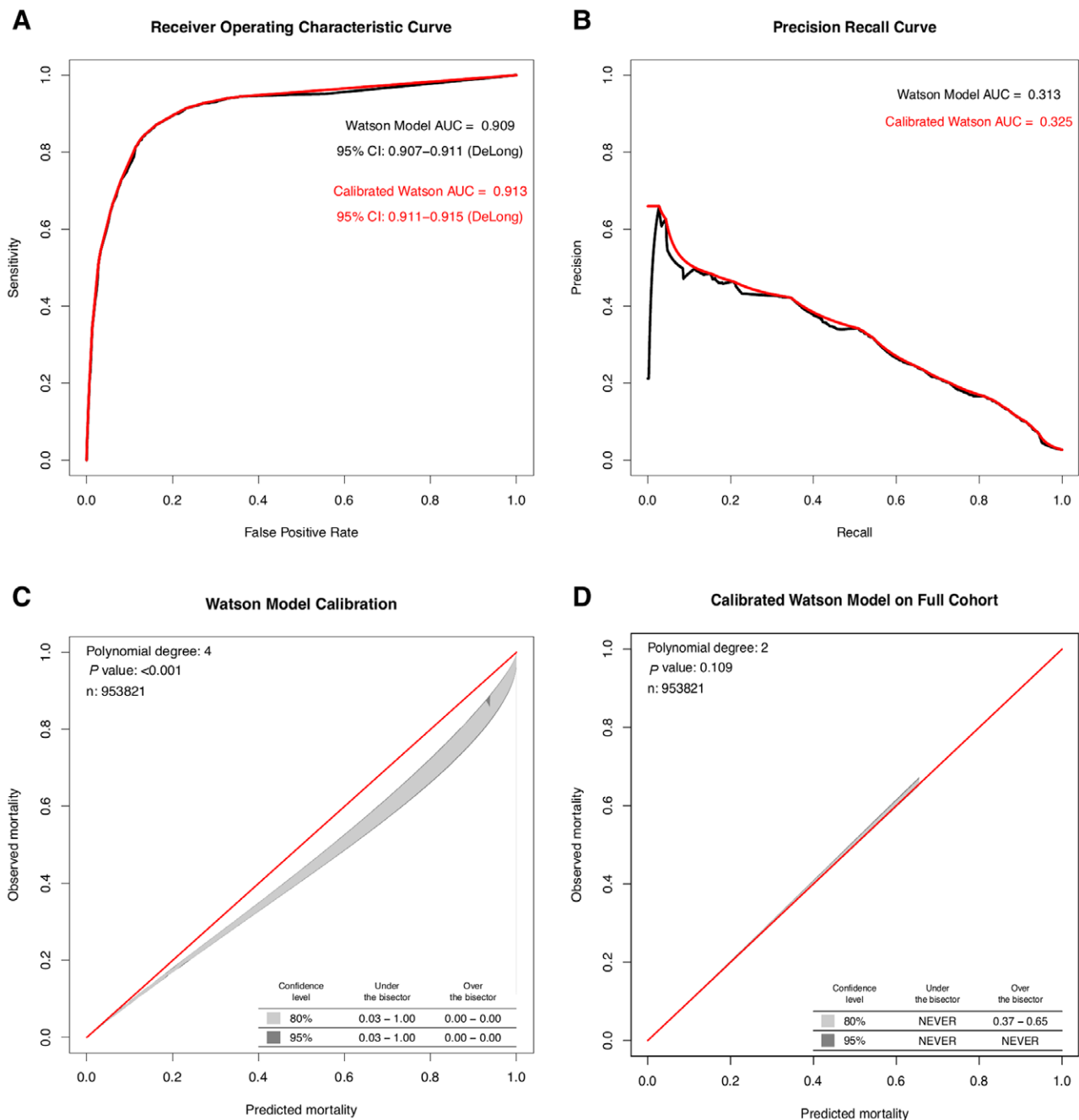


Fig. 1. Watson score performance. A, Receiver operating characteristic curve. B, Precision recall curve. The x axis shows recall (sensitivity). The y axis shows precision (positive predictive value). C, D, GiViTI calibration belt before and after isotonic regression, respectively. D, Calibration belt for full cohort, with more stringent internal calibration methodology applied. The *P* value is a likelihood ratio test evaluating the deviation of the model from the observed result; a nonsignificant *P* value is desirable.

calibration, the Watson score showed a c-statistic of 0.909 (95% CI 0.907–0.911) and 0.913 (95% CI 0.911–0.915) and an area under the precision recall curve of 0.313 and 0.325, respectively. As shown in Figure 1, before isotonic regression, calibration was poor ($P < 0.001$), with overprediction of mortality throughout the cohort. After calibration, the model showed acceptable calibration on both the full cohort ($P = 0.109$) and the test subset ($P = 0.261$).

Correlation of Antimicrobial Prescribing and Risk-adjusted Mortality

As shown in Figure 2, there was no correlation between antimicrobials per 1000 ICU days and risk-adjusted mortality according to the uncalibrated (adjusted $r^2 = -0.019$, $P = 0.83$) or calibrated Watson score (adjusted $r^2 = -0.020$, $P = 0.87$). One ICU was an outlier in antimicrobial prescriptions; removal of this ICU did not change the result (adjusted $r^2 = -0.005$, $P = 0.386$,

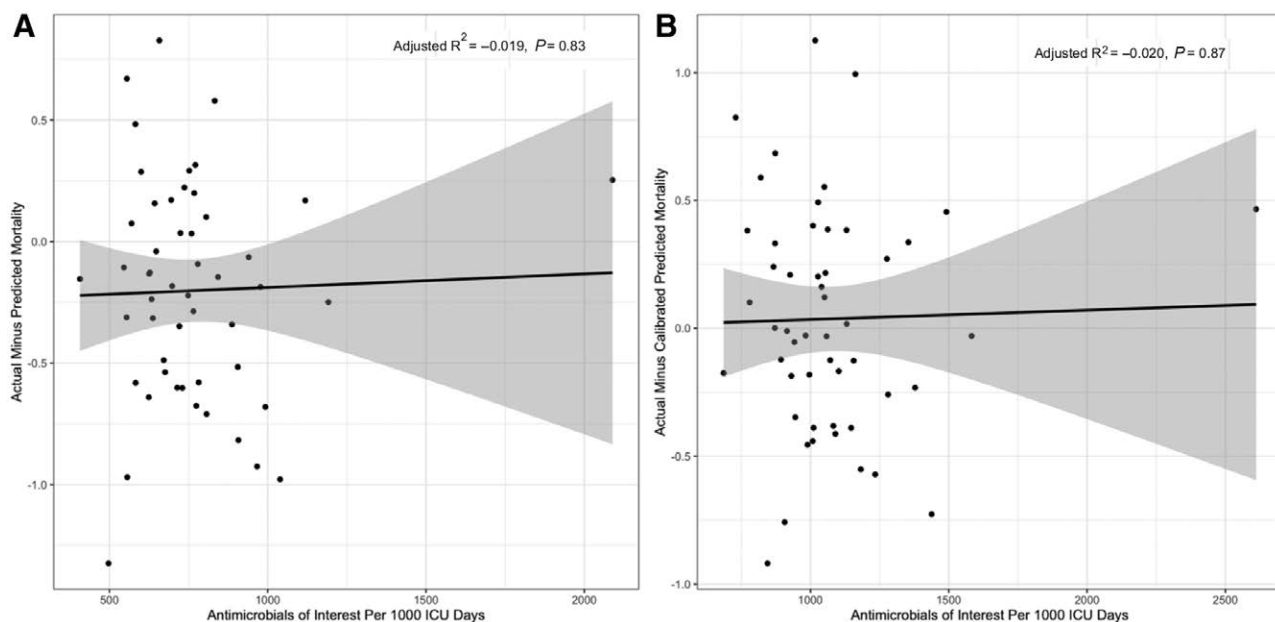


Fig. 2. Correlation between antimicrobial prescriptions and risk-adjusted pediatric ICU mortality in PHIS. A, Antimicrobial days of therapy per 1000 pediatric ICU days versus differences from Watson score predicted mortality. B, Antimicrobial days of therapy per 1000 pediatric ICU days versus differences from Watson score predicted mortality after recalibration with isotonic regression. For both panels, each point represents a single hospital, with antimicrobial prescriptions per 1000 ICU days plotted on the x axis, and difference from predicted mortality (% mortality – % predicted mortality) on the y axis. For both panels, the line represents linear regression with 95% confidence interval.

ICU Antimicrobial Use of Interest Over Time and Risk-adjusted Mortality

There has been an overall change of -42.5 days of therapy per 1000 ICU days per year. $P < 0.001$.

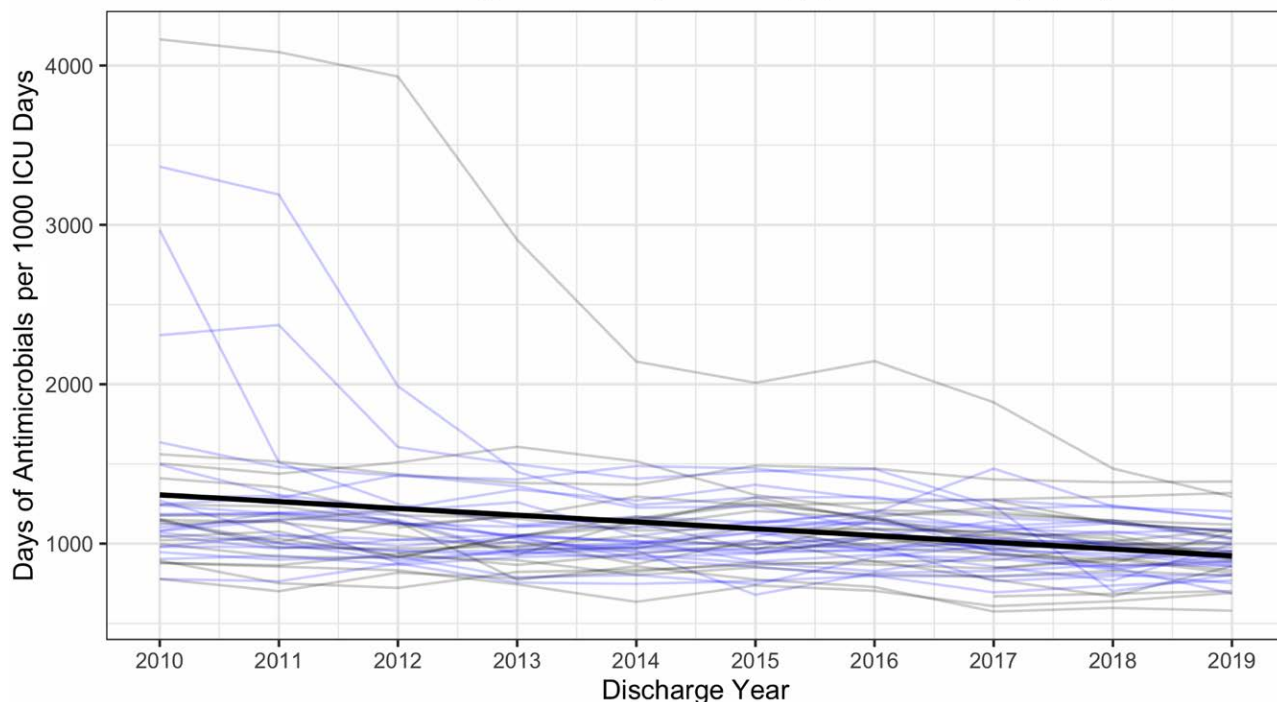


Fig. 3. Antimicrobial prescriptions over time and risk-adjusted pediatric ICU mortality. Each thin line represents 1 PHIS pediatric ICU over time. Antimicrobial days of therapy per 1000 pediatric ICU days are graphed on the y axis against patients' discharge year on the x axis. Line color is representative of changes in risk-adjusted mortality over time as assessed by linear regression. ICUs with statistically significant reductions in risk-adjusted mortality are graphed in blue. Those with nonsignificant changes are graphed in black. There were no ICUs with significant increases in risk-adjusted mortality. The thick black line represents the line of best fit for the data, or average among the ICUs.

data not shown). A sensitivity analysis excluding amoxicillin and erythromycin did not significantly change the results (adjusted $r^2 = -0.020$, $P = 0.87$, data not shown). As shown in Figure 3, overall antimicrobial prescriptions decreased by 42.5 days per 1000 ICU days over the study period ($P < 0.001$). During the same time period, 25 of 51 ICUs had a statistically significant trend of improvement in risk-adjusted mortality by linear regression ($P < 0.05$). No ICU had a significant trend toward worsening risk-adjusted mortality, and 26 of 51 ICUs had nonsignificant trends.

PHIS+ Nested Subset Analysis

The PHIS+ nested subset contained 45,706/953,821 (4.8%) of the patients in the total study cohort. The demographics of the PHIS+ nested subset are listed in Table 1. The novel multivariable logistic regression mortality prediction model coefficients are included as supplemental table. (See table, Supplemental Digital Content 3, which shows multivariable model for mortality prediction, <http://links.lww.com/PQ9/A320>.) On the testing subset, the model had a c-statistic of 0.943 (95% CI 0.930–0.957; See figure 3, Supplemental Digital Content 4, which shows PHIS+ novel mortality score performance. Panel S3A shows the receiver operating characteristic curve. Panel S3B shows the precision recall curve. The x-axis shows recall (sensitivity). The y-axis shows precision (positive predictive value). Panel S3C shows the GiViTI calibration belt. The P -value is a likelihood ratio test evaluating the deviation of the model from the observed result, a non-significant P -value is desirable. All panels show the model performance on the testing

subset, <http://links.lww.com/PQ9/A319>) and area under the precision recall curve of 0.524. Model calibration was adequate ($P = 0.175$; See figure 3, Supplemental Digital Content 4, which shows PHIS+ novel mortality score performance. Panel S3A shows the receiver operating characteristic curve. Panel S3B shows the precision recall curve. The x-axis shows recall (sensitivity). The y-axis shows precision (positive predictive value). Panel S3C shows the GiViTI calibration belt. The P -value is a likelihood ratio test evaluating the deviation of the model from the observed result, a non-significant P -value is desirable. All panels show the model performance on the testing subset, <http://links.lww.com/PQ9/A319>). The discrimination of the PHIS+ model is compared against the Watson score on the entire PHIS+ cohort in Figure 4. The PHIS+ model had significantly better discrimination than the Watson score (c-statistic 0.940 [95% CI 0.933–0.947] versus 0.891 [95% CI 0.881–0.901], respectively $P < 0.001$ and area under the precision recall curve 0.561 versus 0.297, respectively). Both the PHIS+ model and the Watson score showed a significant inverse association between antimicrobials per 1000 ICU days and risk-adjusted mortality (adjusted $r^2 = 0.57$, $P = 0.05$, and adjusted $r^2 = 0.88$, $P < 0.01$, respectively, Figure 5).

DISCUSSION

Pediatric database analyses are becoming increasingly influential in the design and implementation of antimicrobial stewardship projects.^{7,8} They have the potential to provide baseline data that can be used to drive change in antibiotic prescribing practices. However, previous

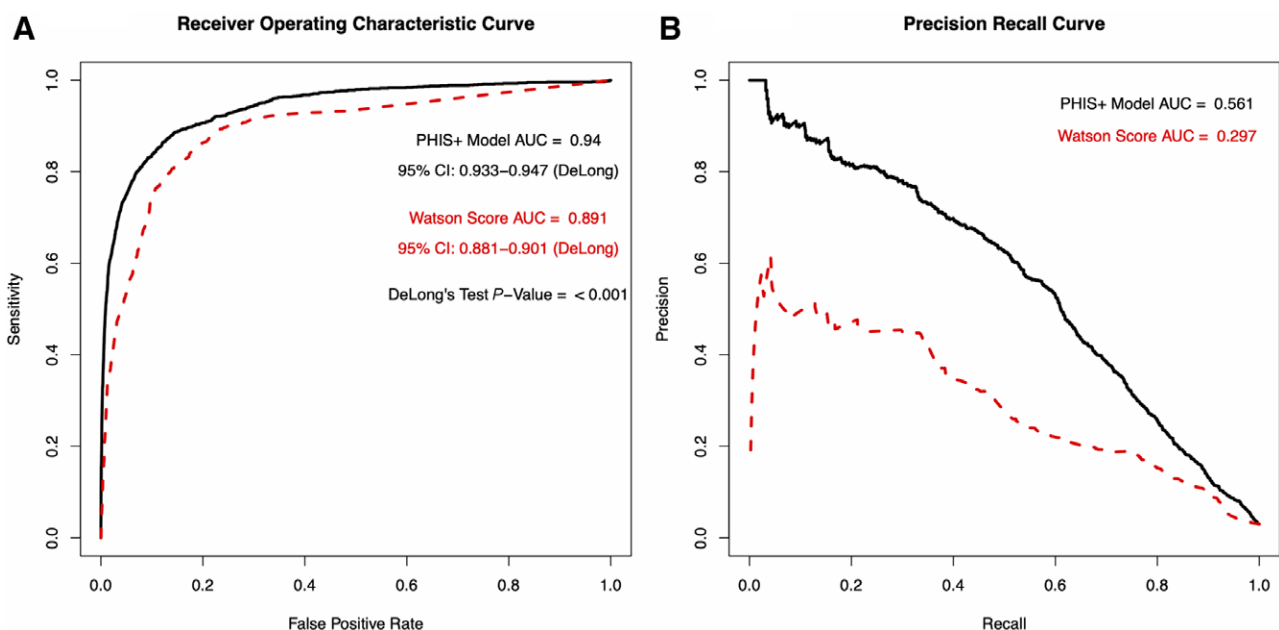


Fig. 4. Comparison of Watson score and Novel Mortality Model Performance on the PHIS+ Subset. A, Receiver operating characteristic curve. The P value is DeLong's test for related receiver operating characteristic curves. B, Precision recall curve. The x axis shows recall (sensitivity).

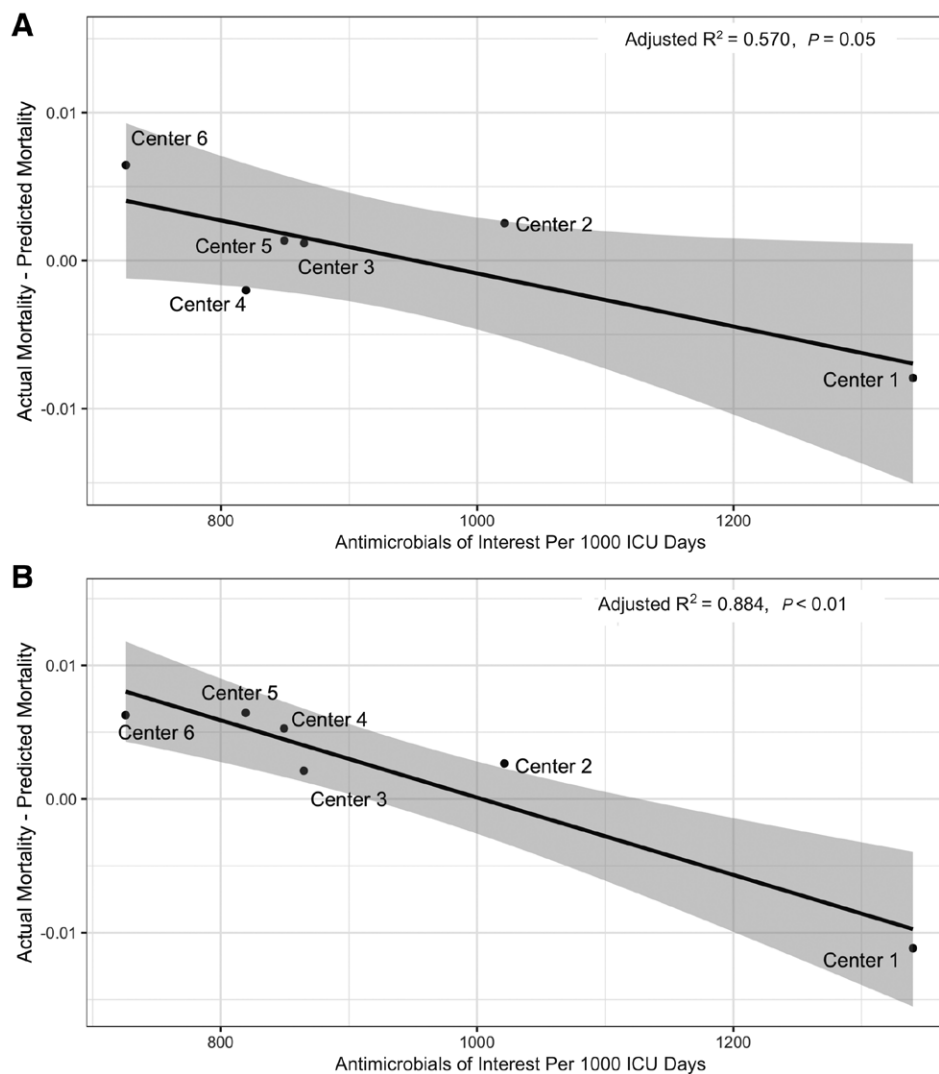


Fig. 5. Correlation between antimicrobial prescriptions and risk-adjusted pediatric ICU mortality in the PHIS+ Subset. A, Antimicrobial days of therapy per 1000 pediatric ICU days versus differences from novel PHIS+ score predicted mortality. B, Antimicrobial days of therapy per 1000 pediatric ICU days versus differences from Watson score predicted mortality. For both panels, each point represents a single hospital, with antimicrobial prescriptions per 1000 ICU days plotted on the x axis, and difference from predicted mortality (% mortality – % predicted mortality) on the y axis. For both panels, the line represents linear regression with 95% confidence interval.

analyses have been constrained by limited ability to adjust for severity of illness, as recommended by the CDC.^{2,14} In this large, decade-long, multi-center, retrospective study of the PHIS database, we report substantial heterogeneity in antimicrobial use among pediatric ICUs. Our findings are similar to those previously published, with 4-fold and 7-fold differences in anti-MRSA and antipseudomonal antibacterial use, respectively.² For the first time, we report that this variation in antimicrobial use is not associated with hospital-level risk adjusted mortality determined by the Watson score.

We also report that the PHIS ICUs reduced overall antimicrobial use over the study period with no significant worsening of risk-adjusted mortality. Notably, as shown in Figure 3, ICUs with the highest antimicrobial use at the start of the study period experienced the greatest reduction in overall use. These trends may suggest that

targeted reduction in antimicrobial use is not associated with harm. At our institution, antimicrobial stewardship efforts have been coupled with the implementation of electronic sepsis screening tools that trigger rapid response by physicians. Such tools have been shown to reduce mortality.³³ This work has been part of a larger effort instituted on a national level through the Improving Pediatric Sepsis Outcomes (IPSO) initiative, which has led to a 19% reduction in sepsis related mortality.³⁴ The approach of rapidly identifying sepsis, tailoring therapies based on culture results, and discontinuing antimicrobials at the earliest possible time point can ideally result in substantial reductions in antimicrobial use while also leading to reductions in mortality. However, there are many potential confounders that can influence patient-centered outcomes among children receiving antibiotics during an inpatient stay. Accordingly, developing robust risk

adjustment models that can be calculated using multi-center data would allow for more accurate assessments of antimicrobial stewardship initiatives when constructing process control charts. Risk adjustment has been a key component of other database analyses that have identified best practices in the management of pediatric sepsis.²⁷ Stewardship efforts at curbing antimicrobial use must be carefully monitored to ensure that the impact on patient outcomes is favorable.

The above trends suggest that current antimicrobial stewardship efforts have effectively reduced overall antimicrobial use without negatively impacting population-level mortality. However, contradictory findings in the PHIS+ subset analysis highlight the vulnerabilities in using administrative databases to drive antimicrobial stewardship projects.⁷⁻⁹ Viewed in isolation, Figures 2 and 5 would lead to contradictory conclusions. Figure 2 suggests that antimicrobial use is not associated with risk-adjusted mortality, whereas Figure 5 suggests that it may be. The reason for these disparate findings requires further investigation. Though the PHIS+ cohort in Figure 5 represented over 45,000 pediatric encounters, only 3 years of data from 6 large academic centers are included. Analysis of a small number of centers is prone to confounding by other center-specific practices. The inverse association between antimicrobial prescribing and mortality was largely mediated by a single center. This association appears related to specific centers, rather than risk-adjustment methods because it was present when using either the IBM-Watson model or the novel PHIS+ model.

The differing findings of the main analysis (Fig. 2) and the subset analysis (Fig. 5) highlight the need for specifically dedicated quality improvement databases. Ideally, databases used for antimicrobial stewardship projects would contain validated severity of illness metrics, culture data (including antimicrobial susceptibilities), data on antimicrobial choice before and after culture results, clinical diagnosis, rationale for treatment choice, and *C. difficile* rates (common metric sometimes associated with antibiotic overuse).³⁵ Given the remarkable success of the IPSO initiative,³⁴ the addition of such metrics to PHIS (or similar databases) is likely to offer a powerful platform to develop multi-center QI projects with an aim to safely reduce unnecessary antimicrobial use.

Because pediatric ICUs have heterogenous case-mix and low overall mortality, risk-adjustment is necessary to accurately compare performance between centers, or to assess performance of a single center over time. While the Watson score demonstrated very good discrimination in our cohort (*c* statistic 0.909 [95% CI 0.907–0.911]), this was likely overestimated by the overall low incidence of pediatric ICU mortality, as demonstrated by the area under the precision recall curve of 0.313. Additionally, without additional calibration, the model overpredicted mortality in most of the cohort, as shown in Figure 1C, where the calibration belt falls below the line of equivalence.

Proprietary algorithms such as the Watson score are also hampered by the lack of freely available derivation and performance data. To build an unambiguous model, we conducted a nested subset analysis of the PHIS+ database, which contains federated EHR data, and developed a mortality prediction model with improved discrimination compared with the Watson score. Our model had a *c*-statistic of 0.94 in this subset, making its performance comparable to the Paediatric Index of Mortality 3 score (0.88), Pediatric Logistic Organ Dysfunction Score (0.94), and Pediatric Risk of Mortality Score IV (0.90).^{16-18,31} This improvement in performance highlights the value of capturing granular EHR data into a multi-center repository.¹⁹

This study has important limitations. First, while consistent with prior literature, antimicrobial days of therapy per 1000 ICU days is a general metric and does not account for differences in specific antimicrobial choices.² The PHIS database does not include provider notes or microbiological data. Specific culture data and resistance patterns likely informed many antimicrobial decisions. Additionally, as addressed in the nested subset analysis, the Watson score is proprietary, and the addition of transparent severity of illness metrics would substantially improve future database analyses.^{16-18,31} PHIS+ has limited reliability of time-stamps on laboratory values, preventing our ability to censor laboratory values from within the last few hours of life. Lastly, all institutions were in the United States, limiting the generalizability out of this region.

Antimicrobial prescribing among the pediatric ICUs in the PHIS database is highly variable, and not associated with risk-adjusted mortality as assessed by the Watson score. At the ICU level, reductions in antimicrobial use have occurred simultaneously with improvement or no change in risk-adjusted mortality over time. We demonstrate that the addition of EHR-derived laboratory data to administrative registries offers a potentially powerful platform for targeted improvement in antimicrobial prescribing by adjusting for important confounders related to mortality, an important balancing measure for inpatient antibiotic stewardship initiatives that naturally intersect with the treatment of life-threatening serious bacterial infections and sepsis. In the future, sharing open-source code to perform risk adjustment at individual sites leveraging data from multicenter registries could allow sites to calibrate models to their own data while still performing standardized risk adjustment to monitor for improvement in antibiotic use over time. These approaches would allow sites to track the impact of quality improvement initiatives while adjusting for differences in patient population that may otherwise confound the relationship between antimicrobial prescribing and patient-centered outcomes such as mortality, an approach recommended by the CDC.

DISCLOSURE

The authors have no financial interest to declare in relation to the content of this article.

ACKNOWLEDGMENTS

The authors thank Erin Weslander, PharmD, for her assistance in selecting antimicrobials of interest, and access to hospital-level antimicrobial data. This study was supported by 5T32HD040686-20 (JHP), 5K23NS104133 (AKA), and 1K23HD099331-01A1 (CH).

REFERENCES

- Blinova E, Lau E, Bitnun A, et al. Point prevalence survey of antimicrobial utilization in the cardiac and pediatric critical care unit. *Pediatr Crit Care Med*. 2013;14:e280–e288.
- Brogan TV, Thurm C, Hersh AL, et al. Variability in antibiotic use across PICUs. *Pediatr Crit Care Med*. 2018;19:519–527.
- Gerber JS, Newland JG, Coffin SE, et al. Variability in antibiotic use at children's hospitals. *Pediatrics*. 2010;126:1067–1073.
- Grohskopf LA, Huskins WC, Sinkowitz-Cochran RL, et al; Pediatric Prevention Network. Use of antimicrobial agents in United States neonatal and pediatric intensive care patients. *Pediatr Infect Dis J*. 2005;24:766–773.
- Centers for Disease Control and Prevention. *Antibiotic Resistance Threats in the United States*. Atlanta, GA: U.S. Department of Health and Human Services; 2013.
- Burns JP, Sellers DE, Meyer EC, et al. Epidemiology of death in the PICU at five U.S. teaching hospitals. *Crit Care Med*. 2014;42:2101–2108.
- Children's Hospital Association. Data helps create the burning platform: key to successful antibiotic stewardship. Available at <https://www.childrenshospitals.org/-/media/Files/Groups/PHIS/Case-Studies/PHIS-Data-Comparison--Childrens-Memorial-Hermann.pdf?la=en&hash=D53B33D29C64E6F3FC5F76C140B1E-D9E0BB2F6B9>. Published 2019. Accessed 2/14/2020, 2020.
- Parker SK, Hurst AL, Thurm C, et al. Anti-infective acquisition costs for a stewardship program: getting to the bottom line. *Clin Infect Dis*. 2017;65:1632–1637.
- Children's Hospital Association. Addressing the threat of antibiotic resistance. Available at <https://www.childrenshospitals.org/Newsroom/Childrens-Hospitals-Today/Articles/2016/11/Antibiotic-Stewardship-Program>. Published 2016. Accessed November 11, 2020.
- Children's Hospital Association. PHIS. Available at <https://www.childrenshospitals.org/Programs-and-Services/Data-Analytics-and-Research/Pediatric-Analytic-Solutions/Pediatric-Health-Information-System>. Accessed July 29, 2019.
- Leis JA. Advancing infection prevention and antimicrobial stewardship through improvement science. *BMJ Qual Saf*. 2018;27:163–165.
- Phua J, Ngerng W, See K, et al. Characteristics and outcomes of culture-negative versus culture-positive severe sepsis. *Crit Care*. 2013;17:R202.
- Wilson MR, Sample HA, Zorn KC, et al. Clinical metagenomic sequencing for diagnosis of meningitis and encephalitis. *N Engl J Med*. 2019;380:2327–2340.
- Centers for Disease Control and Prevention. Antimicrobial use and resistance (AUR) module. Available at <https://www.cdc.gov/nhsn/pdfs/pscmanual/11pscscurrent.pdf>. Published 2020. Accessed February 14, 2020.
- 3M Health Information Systems. 3M All Patient Refined Diagnosis Related Groups (3M APR DRG). Available at https://www.3m.com/3M/en_US/health-information-systems-us/providers/grouping-and-classification/apr-drgs/. Published 2020. Accessed February 14, 2020.
- Leteurre S, Duhamel A, Salleron J, et al; Groupe Francophone de Réanimation et d'Urgences Pédiatriques (GFRUP). PELOD-2: an update of the Pediatric logistic organ dysfunction score. *Crit Care Med*. 2013;41:1761–1773.
- Pollack MM, Holubkov R, Funai T, et al; Eunice Kennedy Shriver National Institute of Child Health and Human Development Collaborative Pediatric Critical Care Research Network. The pediatric risk of mortality score: update 2015. *Pediatr Crit Care Med*. 2016;17:2–9.
- Straney L, Clements A, Parslow RC, et al; ANZICS Paediatric Study Group and the Paediatric Intensive Care Audit Network. Paediatric index of mortality 3: an updated model for predicting mortality in pediatric intensive care. *Pediatr Crit Care Med*. 2013;14:673–681.
- Narus SP, Srivastava R, Gouripeddi R, et al. Federating clinical data from six pediatric hospitals: process and initial results from the PHIS+ Consortium. *AMIA Annu Symp Proc*. 2011;2011:994–1003.
- Keren R. PHIS+: augmenting the Pediatric Health Information System with clinical data. Available at <https://federalreporter.nih.gov/Projects/Details/?projectId=540648&itemNum=889944&totalItems=892504&searchId=b850241613a74a58962c0bd1a1edd5d4&searchMode=Smart&page=8900&pageSize=100&sortField=Ic&sortOrder=desc&filters=&navigateon=True>. Published 2012. Accessed May 16, 2020.
- Mongelluzzo J, Mohamad Z, Ten Have TR, et al. Corticosteroids and mortality in children with bacterial meningitis. *JAMA*. 2008;299:2048–2055.
- Weiss SL, Peters MJ, Alhazzani W, et al. Surviving sepsis campaign international guidelines for the management of septic shock and sepsis-associated organ dysfunction in children. *Pediatr Crit Care Med*. 2020;21:e52–e106.
- Feudtner C, Feinstein JA, Zhong W, et al. Pediatric complex chronic conditions classification system version 2: updated for ICD-10 and complex medical technology dependence and transplantation. *BMC Pediatr*. 2014;14:199.
- Waikar SS, Wald R, Chertow GM, et al. Validity of international classification of diseases, ninth revision, clinical modification codes for acute renal failure. *J Am Soc Nephrol*. 2006;17:1688–1694.
- Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One*. 2015;10:e0118432.
- Finazzi S, Poole D, Luciani D, et al. Calibration belt for quality-of-care assessment based on dichotomous outcomes. *PLoS One*. 2011;6:e16110.
- Evans IVR, Phillips GS, Alpern ER, et al. Association between the New York sepsis care mandate and in-hospital mortality for pediatric sepsis. *JAMA*. 2018;320:358–367.
- Brady AR, Harrison D, Black S, et al; UK PICOS Study Group. Assessment and optimization of mortality prediction tools for admissions to pediatric intensive care in the United Kingdom. *Pediatrics*. 2006;117:e733–e742.
- Huang Y, Li W, Macheret F, et al. A tutorial on calibration measurements and calibration models for clinical prediction models. *J Am Med Inform Assoc*. 2020;27:621–633.
- Matics TJ, Sanchez-Pinto LN. Adaptation and validation of a pediatric sequential organ failure assessment score and evaluation of the sepsis-3 definitions in critically ill children. *JAMA Pediatr*. 2017;171:e172352.
- Horvat CM, Ogoe H, Kantawala S, et al. Development and performance of electronic pediatric risk of mortality and pediatric logistic organ dysfunction-2 automated acuity scores. *Pediatr Crit Care Med*. 2019;20:e372–e379.
- Akaike H. Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, eds. *Selected Papers of Hirotugu Akaike*. Springer, N.Y.; 1998:199–213.
- Guirgis FW, Jones L, Esma R, et al. Managing sepsis: Electronic recognition, rapid response teams, and standardized care save lives. *J Crit Care*. 2017;40:296–302.
- Children's Hospital Association. Sepsis Collaborative. Available at <https://www.childrenshospitals.org/programs-and-services/quality-improvement-and-measurement/collaboratives/sepsis/>. Accessed May 27, 2021.
- Kalil AC, Metersky ML, Klompas M, et al. Management of adults with hospital-acquired and ventilator-associated pneumonia: 2016 clinical practice guidelines by the Infectious Diseases Society of America and the American Thoracic Society. *Clin Infect Dis*. 2016;63:e61–e111.