HHS Public Access

Author manuscript

IEEE EMBS Int Conf Biomed Health Inform. Author manuscript; available in PMC 2017 February 01.

Published in final edited form as:

IEEE EMBS Int Conf Biomed Health Inform. 2016 February ; 2016: 86–89. doi:10.1109/BHI.2016.7455841.

PEPCOR — A Risk Prediction Model for Pediatric Intensive Care Units Utilizing Ventilator Days and Length of Stay

Theruni Pethiyagoda [Student Member, IEEE],

Wallace H. Coulter Department of Biomedical Engineering of Georgia Institute of Technology, Atlanta, GA 30332, USA

Nikhil Chanani,

Sibley Heart Center Cardiology, Children's Healthcare of Atlanta and Department of Pediatrics, Emory University School of Medicine, Atlanta, GA 31141, USA

Chihwen Cheng [Student Member], and

Electrical and Computer Engineering Department of Georgia Institute of Technology, Atlanta, GA 30332, USA

May D. Wang [Senior Member, IEEE]

Bio-MIBLab (www.bio-miblab.org), the Wallace H. Coulter Department of Biomedical Engineering Department, Georgia Institute of Technology and Emory University, Atlanta, GA 30332 USA

Abstract

Great sources of concern for pediatric and neonatal intensive care units are the total resource utilization and cost of caring for very sick children. This paper attempts to create a usable clinical decision support system (PEPCOR) that would help improve personalized health care and avoid unnecessary secondary, and costly, procedures that do not improve the care of a critically ill patient. The system uses the Children's Healthcare of Atlanta ICU database to analyze the effects on the risk, (quantified as the risk of invasive ventilation and associated complications) when combining two procedures versus administering one procedure without the other. The risk is calculated by computing the ratio of ventilator days to the length of stay, with a higher risk score indicating the second procedure did not improve the overall health of the patient. 82/90 of the procedure combinations showed a statistically significant difference with a p-value < 0.05 when the stand-alone procedure was compared to the combination of two procedures.

Section I Introduction

The importance of cost reduction analysis for pediatric patients stems from the fact that there is a significant concern about the ongoing rise in medical costs. According to the U.S. Census Bureau, 7.3% of children under the age of 18 are uninsured medically by either private health insurance or government health insurance. Of all children with health insurance, the government provides coverage for 41.1 % of them. These statistics

demonstrate the extent of government funding and the importance of reducing unnecessary hospital bills on society at large [1].

The rising concern about the increase in medical costs has led to the need to develop a means of predicting the severity of illness of ICU patients and the risks associated with future stay. Such models are used to ensure proper resource utilization and also predict the different aspects of the stay of a patient that may add to any unnecessary financial cost such as tests and other medical procedures. The most frequently cited models to predict the risk of death for pediatric patients in the ICU setting are the Pediatric Risk of Mortality (PRISM) III score and the Pediatric Index of Mortality (PIM) score [2]. The PRISM III score uses minimum systolic blood pressure, abnormal pupillary reflexes, and stupor/coma along with two chronic diagnoses and four other risk factors to calculate the final score [3]. The PRISM III score is an improved model based on PRISM I, which calculates the severity of illness based on previously recorded data of the institution of choice. The PIM score, on the other hand, is calculated using nine variables that are recorded at the time of admission to the ICU (including mechanical ventilation during the first hour of admission and Systolic Blood Pressure). The score is an indication of the child's health at the start of an ICU stay [4].

Although the performance of the modified prediction models has been credited to be successful in mortality risk prediction, the PIM and PRISM III models tend to over-predict death by 16% and 30% respectively, when comparing the observed deaths to expected deaths [5]. The PRISM III system can only be applied after the period required to gather variables, as the prediction model can only be used 24 hours after PICU admission [3] [6]. The PIM score is calculated at the time of PICU admission and does not require a 24 hour wait period. However, the usefulness of the concept of a mortality prediction score in assessing PICU performance is a relatively minor one as mortality only occurs in less than 5% of patients and may not reflect true resource utilization [7]. These existing models for risk prediction in the pediatric ICU fail to address additional relevant outcomes such as length of stay and risks associated regarding medical procedures to be performed, which may be more beneficial features to be considered to help increase ICU performance by better-personalized healthcare.

To incorporate more commonly seen features into a prediction model we propose a model for the prediction of risk of invasive ventilation and associated complications. This paper develops the Pediatric Procedure Combination Risk, **PEPCOR**, clinical decision support system that allows ICU clinicians to predict the potential ICU risk of combining a medical procedure-of-interest given another existing procedure. We defined the risk of invasive ventilation as the ratio of ventilator days to total length of stay, as a higher number of ventilator days per unit length of stay indicates a patient more prone to higher mortality and related complications. As some of the main causes of unnecessarily high ICU costs are the administration of excess tests, increased length of stay and increased duration of mechanical ventilation [8], PEPCOR aims to utilize data mining techniques to facilitate personalized healthcare and eventually lead to reducing unnecessary administration of procedures that may thereby reduce the associated cost of care.

Section II Methods

A. Dataset

The CHOA database contains 4,023 Pediatric ICU (PICU) patient records, 941 Neonatal ICU (NICU) patient records, and 774 Cardiac ICU (CICU) patient records. These records contain multiple categories of information per patient as listed in the “Patient/Visit Information.” measure examples of Table I. The CHOA database also contains 416,520 combined records of patient procedures for the ICU units mentioned above. The PEPCOR prediction model was developed and tested using this database

B. Variable Selection

Following the selection of the length of stay (LOS) as the independent variable, we used literature review to determine the criteria for selection of the dependent variable of the study. The most notable reviews were the work of [9] which suggested that the severity of the illness was a major factor in determining LOS resource utilization and cost, and the study in [10] which noted ventilator management as one of the four possible treatments that most affected cost in the ICU setting. Ventilator-days was thus selected as the dependent variable for the PEPCOR system as it was the feature present in the CHOA database that was most indicative of the severity of the patient’s illness (as identified by the clinician) and most compatible with LOS data, as they shared the same units.

C. Data Organization

Once the preliminary feature was chosen, we interrogated the CHOA database and extracted only the relevant information namely the *visit ID*, LOS, ventilator days and procedure IDs. The *procedure ID* number and *procedure name* along with all the patient *visitID numbers* of the patients that had each procedure performed on them was extracted, and a new table was created. The *visitIDs* were not unique in that the extracted data included multiple of the same *visitID* if the same procedure had been performed multiple times on that patient. After organizing the list of non-unique *visitID* numbers against each procedure, the corresponding ventilator days of each patient on that list and the corresponding LOS days was also included separately in the table.

Section III Case study and Hypothesis

A. Setting

To demonstrate the feasibility of our system, we focused on patients who were diagnosed with “acute respiratory failure,” as these patients were more likely to be associated with invasive ventilation, which is a key component of this study as the risk ratio is calculated using ventilator-days. 1640 of the 5738 patients were diagnosed with “acute respiratory failure” which is 28.6% of the total patients. This could be extended to include 49.8% of the patients if patients diagnosed with ‘respiratory distress syndrome’, ‘acute upper respiratory infections’, ‘acute bronchitis’ and other respiratory classifications were included. We hypothesize that the results of such an inclusion may be similar as such patients would have ventilator-days data, however such classifications were not included for the study in this paper, as further clarification is needed about such an inclusion. The ten most repeated

procedures (i.e., the ten procedures with the highest number of *visitID* counts that were recorded to have “acute respiratory failure”) were then extracted along with their corresponding *visitIDs*, ventilator days and LOS data. Each procedure was then combined with each of the other nine procedures producing 90 (N×9) combinations (order was considered important, and procedure A+B for example, and procedure B+A were both included in the list of combinations).

Patients were separated into two groups for each combination. The first group consisted of ventilator and LOS data from visits with *both* procedures. The second group consisted of ventilator and LOS data from visits with the first procedure but without the second. With the LOS data and ventilator day data from the two groups, a graph was created with the x -axis being the length of stay (days) plotted against the y-axis of ventilator days. The slope of these graphs was expected to resemble the risk ratio as a higher number of ventilator days to unit LOS days would indicate a more severely ill patient prone to more invasive ventilation and associated complications thus leading to more staff service, resource utilization, and longer ICU stay. To calculate the risk ratio, the slope of the graphs was calculated using the equation

$$y=mx \quad (1)$$

Using this formula forced a zero intercept as a non-zero intercept resulted in extreme intercepts and extreme gradients that were not comparable across procedure combinations. The zero intercept instead forced the line of best fit to adjust its slope so as to pass through the origin while also having the best fit through the data of interest.

A separate p-value was also calculated to compare the risk ratios of the inclusion and exclusion groups. To do this, each ventilator-days value of each patient in the inclusion group was divided by the corresponding LOS-days value to obtain a risk ratio for each patient. This process was repeated for the exclusion group, and the p-value of the comparison of these two sets of risk ratios was calculated using a two-sample t-test. The hypothesis for the study was that there is a significant difference in the risk of administering the second procedure given an administered first procedure (inclusion) versus not administering the second procedure given a first procedure (exclusion).

Section IV Results

The statistical test showed that the majority (82 out of 90) of the procedure combinations show a statistically significant difference (p-value < 0.05) between the inclusion and exclusion groups. The top ten combinations of the results sorted by the lowest p-values are shown in Table II. The effect size of the inclusion group against the exclusion group in the table was $r = 0.93$ using Cohen's d (which is a large effect size) and the p-value was $p < 0.001$ which shows a statistically significant difference between the two groups.

From a clinical perspective, it could be understood that a high-frequency procedure such as a pulse oximetry (as coded by “Stat-pulse ox assessment” in Table II) in combination with

another high-frequency procedure such as frequent blood gas measurement (as indicated by “Poc cap bl gas”) may be redundant information. In contrast, pulse oximetry in combination with a less frequently ordered procedure such as “lipase” measurement which is a serum lipase test that measures the level of lipase in the body or pulse oximetry in combination with “antithrombin III assays” which is a test to determine antithrombin levels to help detect clotting disorders, may be more useful findings as they provide unique information about the patient’s condition.

Section V User Interface

A graphical user interface (GUI) was designed using MAT LAB GUI layout toolbox to input the patient’s current procedure (i.e. the first procedure), and the procedure of interest (i.e. the second procedure). The GUI was designed to then output a graph of the risk ratio comparison of risks from the two (i.e., the inclusion and the exclusion) groups.

As seen in Figure 1 above, the layout of PEPCOR was created to include a drop-down menu of the patient’s unit (PICU, NICU or CICU) so that the CROA databases could be narrowed down further to predict more accurate personalized predictions. An automatic check is also applied as the physician enters the age of the patient in days. The current procedure being applied or already executed on the patient can then be selected by entering the *procedure ID number*, the *procedure name* or the *procedure category* that will refine the table seen below it for easier navigation and selection. The second procedure that the physician is hoping to carry out on the patient can then be selected via the same procedure, and the “Compute” button initiates the analysis and produces the graphical and statistical results. These results include the group of “Intersection IDs data points” in red stars, the group of “Exclusion IDs data points” in blue stars, the estimated “Intersection risk” shown by the red line and the estimated “Exclusion risk” shown by the blue line. The numerical values of the intersection risk and exclusion risk are also provided for ease of comparison.

This application is intended to increase the practical and clinical application of retrospective ICU data to help in personalized health care and risk prediction that would reduce the cost of a patient’s ICU stay.

Section IV Discussion and Conclusion

The increase/decrease in the risk of invasive ventilation and associated complications was analyzed in this paper through risk assessment. The variable “ventilator days” was selected with the aid of a literature review and clinicians’ findings [8], [9]. This variable was compared with the length of stay by extracting patient data from the CHOA database. In this study, we propose the PEPCOR system, which calculates the risk ratio of administering the first procedure of choice and the second procedure of choice. A case study was conducted using the CHOA database in which the ten procedures with the most number of patient records were analyzed. The results revealed risk ratios and the respective p-values of the comparison of the inclusion data (first procedure followed by second procedure) against the exclusion data (first procedure not followed by second procedure), which showed significance ($p\text{-value} < 0.05$) for 82 of 90 procedure combinations. This study also

developed a GUI to demonstrate the clinical usability, which required two procedures of choice as input and provided the corresponding risk ratios as output, with a graphical visualization of the data. There are many situations in the ICU setting in which unnecessary tests may be administered. For example nurses may anticipate doctors' test orders or doctors may order tests similar to those already administered. A physician could use the PEP COR system to evaluate whether the to-be-administered procedure may be statistically known to help improve the overall health of the patient once a previous procedure has already been administered (by a comparison of previous patients' data). This could aid the physician in reducing the administration of unnecessary tests and thereby improve personalized healthcare and reduce overall cost. This project has the potential to be developed into a more versatile application that could be applicable in developing countries, as one of the main limitations with already existing models such as PRISM III was the lack of usability across institutions in developing nations due to the rigid structure of the application [6]. The results of the PEPCOR model could also be further improved by outputting a predicted LOS trajectory of the patients given the procedure combinations of choice, and also possibly the option for a medical practitioner to input multiple procedures instead of two. While only higher frequency procedures were considered in this study, further research is needed on low-frequency-high-cost procedure combinations that were not evaluated in this study due to limited access to cost data.

Pediatric ICU care is a more specialized field compared to adult ICU care and has a greater need for personalized health care as critical conditions depend on a greater number of factors than adult ICU cases that may be more easily generalized [10].

Invasive ventilation and associated medical complications are significant factors causing an increase in risk to patients and leading to extended hospital stays, higher resource utilization, and higher medical costs. Analysis of the effects of *procedure combination* on the increase/decrease of the risk of more invasive ventilation was examined in this study using a risk ratio assessment. The risk ratio was calculated using ventilator days to length of stay and the **PEPCOR** (Pediatric Procedure Combination Risk) clinical decision support system was created which showed statistical significance for a majority of high-frequency procedure combinations. As there is much new research being conducted on the analysis of high pediatric medical costs and finding new ways to reduce such costs, this study has great potential for improvement and collaboration.

Acknowledgments

The authors are grateful to Matthew Miller, Michael Thompson, Sherry Farrugia, and Tod Davis for helping the export and development of the CHOA ICU database used in this study. We also extend our thanks to Chanchala Kaddi, Janani Venugopalan, Po-Yen Wu, Ryan Hoffman and Sameer Mishra for their valuable comments and suggestions.

References

1. Smith, C., Medalia, C. Health Insurance Coverage in the United States: 2013 Current Population Reports. U.S. Government Printing Office; 2014.
2. Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated Pediatric Risk of Mortality score. *Critical care medicine*. 1996; 24:743–752. [PubMed: 8706448]

3. Gemke RJ, van Vught JA. Scoring systems in pediatric intensive care: PRISM III versus PIM. *Intensive care medicine*. 2002; 28:204–207. [PubMed: 11907665]
4. Shann F, Pearson G, Slater A, Wilkinson K. Pediatric index of mortality (PIM): a mortality prediction model for children in intensive care. *Intensive care medicine*. 1997; 23:201–207. [PubMed: 9069007]
5. Slater A, Shann F, Group APS. The suitability of the Pediatric Index of Mortality (PIM), PIM2, the Pediatric Risk of Mortality (PRISM), and PRISM III for monitoring the quality of pediatric intensive care in Australia and New Zealand. *Pediatric Critical Care Medicine*. 2004; 5:447–453. [PubMed: 15329160]
6. Randolph AG. Pediatric index of mortality (PIM): do we need another pediatric mortality prediction score? *Intensive care medicine*. 1997; 23:141–142. [PubMed: 9068997]
7. Dasta JF, McLaughlin TP, Mody SH, Piech CT. Daily cost of an intensive care unit day: The contribution of mechanical ventilation*. *Critical care medicine*. 2005; 33:1266–1271. [PubMed: 15942342]
8. Gruenberg D, Shelton W, Rose L, Rutter AE, Socaris S, McGee G. Factors influencing length of stay in the Intensive Care Unit. 2006; 15:502–509.
9. Bueno H, Ross JS, Wang Y, Chen J, Vidán MT, Normand SLT, et al. Trends in length of stay and short-term outcomes among Medicare patients hospitalized for heart failure, 1993–2006. *Jama*. 2010; 303:2141–2147. [PubMed: 20516414]
10. Coetzee M. Are children really different from adults in critical care settings. *Southern African Journal of Critical Care*. 2005; 21:70, 72–74, 76.

Figure 1.
User Interface of PEPCOR

Table I

Categories and examples of CHOA ICU database

Category	Measure Examples	# of records
Patient/Visit Info.	Demographics, admission/discharge time, birth weight/length, discharge destination, APGAR 1, 5, ID minutes score, ventilator days, financial class, PICU, NICU, C ICU flags	5,738
Diagnoses	Acute respiration failure, esophageal reflux, hypoxemia, unspecified asthma, with status asthmaticus, obstructive hydrocephalus, patent ductus arteriosus, dehydration	59,832
Procedures	Oxygen supply, aerosol treatment, oxygen per shift, PH probe, pulse ox assessment, gastric pressure, suction, reactive protein, tobramycin peak	416,520
Lab Testing	Glucose, arterial PCO ₂ /pH/PO ₂ , oxygen saturation, calcium ionized, HCO ₃ , creatinine, platelet count, potassium, prolactin, salicylates	3,348,924
Microbiology Testing	Culture, specimen description, specimen source	87,843

Table II

Results for top 10 combinations

Procedure Name		Number of Patients		Risk Ratio		P-value of Change
First Procedure	Second Procedure	Proc 1 without Proc 2 (Exclusion)	Proc 1 with Proc 2 (Inclusion)	Exclusion	Inclusion	
Chest IV	Stat oxygen per shift	6062	616	0.760	0.391	3.175e ⁻¹⁷⁶
Poc cap bl gas	Stat oxygen per shift	4125	396	0.795	0.406	7.894e ⁻¹⁶⁶
Poc glucose	Stat oxygen per shift	4533	581	0.656	0.394	1.514e ⁻¹⁵¹
CBC auto w/auto diff	Stat oxygen per shift	6211	673	0.728	0.390	5.730e ⁻¹⁰⁴
Basic metabolic panel	Stat oxygen per shift	5431	548	0.659	0.394	1.439e ⁻⁷³
Stat-pulse ox assessment	Poc cap bl gas	16428	543	0.635	0.533	1.828e ⁻⁷³
MDI treatment	Stat oxygen per shift	4643	270	0.754	0.439	2.969e ⁻⁷³
Stat-pulse ox assessment	Stat oxygen per shift	16191	780	0.636	0.389	1.252e ⁻⁷²
Stat-pulse ox assessment	Chest IV	16134	837	0.635	0.508	6.621e ⁻⁵⁹
Basic metabolic panel	Poc cap bl gas	5547	432	0.658	0.535	5.669e ⁻⁵⁷

Table sorted by lowest P-value