

Article

Unstructured Text in EMR Improves Prediction of Death after Surgery in Children

Oguz Akbilgic ^{1,*} , Ramin Homayouni ^{2,3} , Kevin Heinrich ³, Max Raymond Langham Jr. ⁴ and Robert Lowell Davis ¹

¹ UTSHC-ORNL Center for Biomedical Informatics, Department of Pediatrics, Memphis, TN 38103, USA; rdavis88@uthsc.edu

² Department of Foundational Medical Sciences, Oakland University William Beaumont School of Medicine, Rochester, MI 48309, USA; rhomayouni@oakland.edu

³ Quire Inc., Memphis, TN 38103, USA; kevin@quireanalytics.com

⁴ Department of Surgery, University of Tennessee Health Science Center, Memphis, TN 38103, USA; mlangham@uthsc.edu

* Correspondence: oz@uthsc.edu or oakbilg1@uthsc.edu; Tel.: +1-901-287-5841

Received: 28 October 2018; Accepted: 5 January 2019; Published: 10 January 2019



Abstract: Text fields in electronic medical records (EMR) contain information on important factors that influence health outcomes, however, they are underutilized in clinical decision making due to their unstructured nature. We analyzed 6497 inpatient surgical cases with 719,308 free text notes from Le Bonheur Children’s Hospital EMR. We used a text mining approach on preoperative notes to obtain a text-based risk score to predict death within 30 days of surgery. In addition, we evaluated the performance of a hybrid model that included the text-based risk score along with structured data pertaining to clinical risk factors. The C-statistic of a logistic regression model with five-fold cross-validation significantly improved from 0.76 to 0.92 when text-based risk scores were included in addition to structured data. We conclude that preoperative free text notes in EMR include significant information that can predict adverse surgery outcomes.

Keywords: post-operative death; unstructured data; logistic regression; text mining; surgery outcome

1. Introduction

The National Academies of Science, Engineering, and Medicine has suggested that health information technology has the potential to improve care and health outcomes [1]. To make this vision a reality requires the capacity to predict uncommon events that are potentially preventable. Death after surgery in children is an infrequent occurrence, with an incidence rate of <1.0% [2,3]. Other adverse outcomes such as unplanned return to the operating room, reintubation after surgery, need for blood transfusions, and unplanned readmission are more common, with incidence rates ranging from 0.2% to 4.4% [4–6]. Since over 5 million operations are performed on children each year in the United States [7], even a low rate of postoperative mortality represents thousands of lives lost prematurely. The best published models predicting these events rely on structured data such as that contained in the National Surgical Quality Improvement Program-Pediatric (NSQIP-Ped) by American College of Surgeons [3,8,9]. Such models are useful in quality improvement work, risk-based payment methods, in improving surgical decision-making, and in providing accurate informed consent discussions with parents [10–13].

The value of EMR data mining for research applications and clinical care is beginning to be realized [14,15]. However, analysis of the unstructured text data found in clinical narratives, such as admission and discharge summaries, observation notes, and a variety of reports in the EMR has been

relatively underutilized. Several Natural Language Processing (NLP) methods have been specifically developed for mining EMR data, including MedLEE [16], HiTex [17], cTAKES [18], and TEPAPA [19]. A recent systematic review showed that the majority of NLP approaches have focused on information extraction tasks such as case-detection [20]. Meta-analysis of 19 case-detection studies showed that the median precision and C-statistic significantly increased when unstructured text was used in addition to structured data from the EMR [20]. Other studies have utilized unstructured text in the EMR to predict health outcomes [21,22]. Frost et al. [23], using used text fields from over 43,000 patients to, predict risk of frequent emergency department visits and high system costs with a C-statistic of 0.71 and 0.76, respectively. Weissman et al. [24] showed that inclusion of unstructured text along with structured data improved prediction of death in the ICU by using four different predictive modeling approaches.

Text-mining of clinical notes has been used to identify postoperative complications in veterans [25] but, to the best of our knowledge, has not been utilized for predicting postoperative surgery outcomes in children. In this study, we examined the use of free text notes in the EMR for preoperative prediction of death after surgery in children. We hypothesized that (1) a risk score can be created through mining unstructured free text notes in EMRs and (2) this text-based risk score will improve the performance of models that only use structured data such as those defined by NSQIP-Ped [26].

2. Materials and Methods

We analyzed a sample of children undergoing inpatient surgical procedures at Le Bonheur Children’s Hospital, Memphis, TN, USA, on or before their 19th birthday, whose medical records included preoperative free text notes and whose operation occurred between 1 January 2014 and 31 May 2017 (See Figure 1 for details). Children without preoperative text data were excluded. Those with preoperative text data were divided into two groups based on their inclusion (or not) in the NSQIP-Pediatric program. Preoperative text from non-NSQIP cohort was used to develop a model text-based risk model (training cohort) and this model was then tested on the NSQIP cohort (the testing cohort).

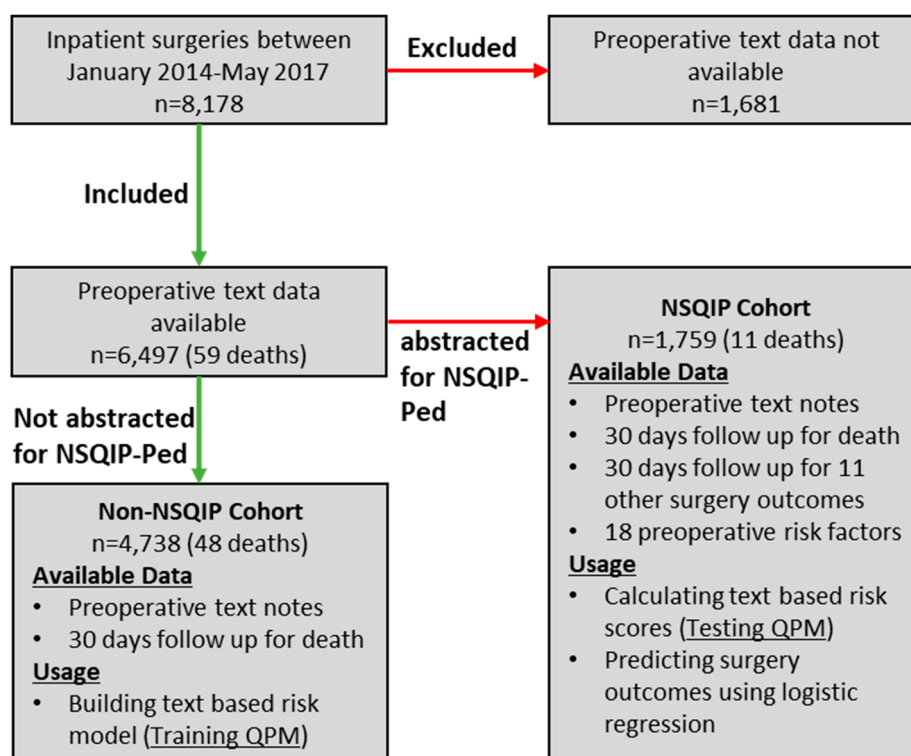


Figure 1. Summary of NSQIP (training) and non-NSQIP (testing) cohorts’ determination and methods.

2.1. NSQIP Cohort

LeBonheur Children's Hospital participates in the NSQIP-Pediatric Program, and reports data that is included in the Participant Use File of the NSQIP-Pediatric Program. A surgical case reviewer abstracts clinical data for a nonrandom sample of children undergoing operative procedures as previously published [26]. Over 300 perioperative-standardized variables were collected. For this study, death within 30 days of surgery (D30) was chosen as the main outcome variable, and other adverse events as secondary outcomes. Based on previous work by our group [3,8,9], fifteen preoperative variables were identified as risk factors of D30. Dichotomous risk factors included ventilator dependency, oxygen support, previous cardiac intervention, cerebral palsy, open wound with or without infections, neuromuscular disorder, bleeding disorder, hematologic disorder, inotropic support, blood transfusion, malignancy, do-not-resuscitate order, and neonatal status. Case type and sepsis were the two risk factors with more than two categories and were converted to multiple dichotomous risk factors. We excluded the American Society of Anesthesiology (ASA) class score since it has been shown to correlate with most of the included variables and is itself a risk score. NSQIP definitions for risk factors and outcomes were used throughout [26].

2.2. Text Mining and Development of Text-Based Risk Score

The flow diagram for the text processing and risk score calculation is shown in Figure 2. Unstructured text fields from all inpatient, outpatient and ambulatory settings prior to the date of surgery were extracted for individuals who had surgeries from January 2014 to May 2017. Records were extracted using HL7 Clinical Document Architecture (HL7/CCA) standard from the Cerner EMR system by Methodist/LeBonheur Hospital, then converted to Continuity of Care Document (CCD) in XML, and submitted to Quire Inc. (Memphis, TN, USA) for downstream processing. Each XML document represents a single patient encounter, which may potentially have multiple notes spanning multiple days at multiple locations. In an example where a patient arrived at ER then transferred to surgery and had a later follow-up visit, each of these three related interactions had an independent provider note but all notes were linked to one encounter ID (FIN). The unstructured text was UTF8 HTML encoded and extracted from the "text" element in the "Assessment and Plan" section of the XML document. The HTML tags were removed and minimal formatting such as tabs and line breaks were kept in the final plain text document. The top 10 document types and corresponding counts were as follows: Clinical Document (311,960), Depart Clinical Summary (15,496), Ambulatory Visit Summary Depart (13,905), Office Visit Note (11,251), ED Clinical Summary (11,232), Pediatric Surgery (10,915), ENT (10,125), pediatric surgery (6647), Teacher Note (6462), and Progress Note (5582). A more extensive list of document types is provided in Supplementary Materials. For each patient, all corresponding unstructured text fields were concatenated into one patient document. Patient documents were then pre-processed using a set of Python scripts to remove: (1) form letters; (2) tabulated numeric lab data; (3) vital signs; and (4) negation phrases. Negation rules for text processing were developed by Quire and were modified iteratively to achieve high precision for this specific collection. Finally, only the most recent history and physical examination prior to surgery was included in each patient-document. All text processing steps described above were performed on the entire patient cohort (including the non-training and test sets described below).

Semantic analysis and text-based risk prediction was performed using an algorithm developed by Quire Inc., which uses a vector space modeling approach called latent semantic indexing [27]. Here, patients were represented as a vector of weighted terms extracted from their medical records (Figure 2). A log-entropy term weighting scheme was used as described by Berry and Browne [27]. Once the term-by-patient document matrix was constructed, singular value decomposition (SVD) was performed to reduce the dimensionality of this matrix into lower rank approximation (concept space). We used a rank of 500 in this study based on evaluation of this and other collections. Patient similarities were calculated using the cosine of the vector angles [27].

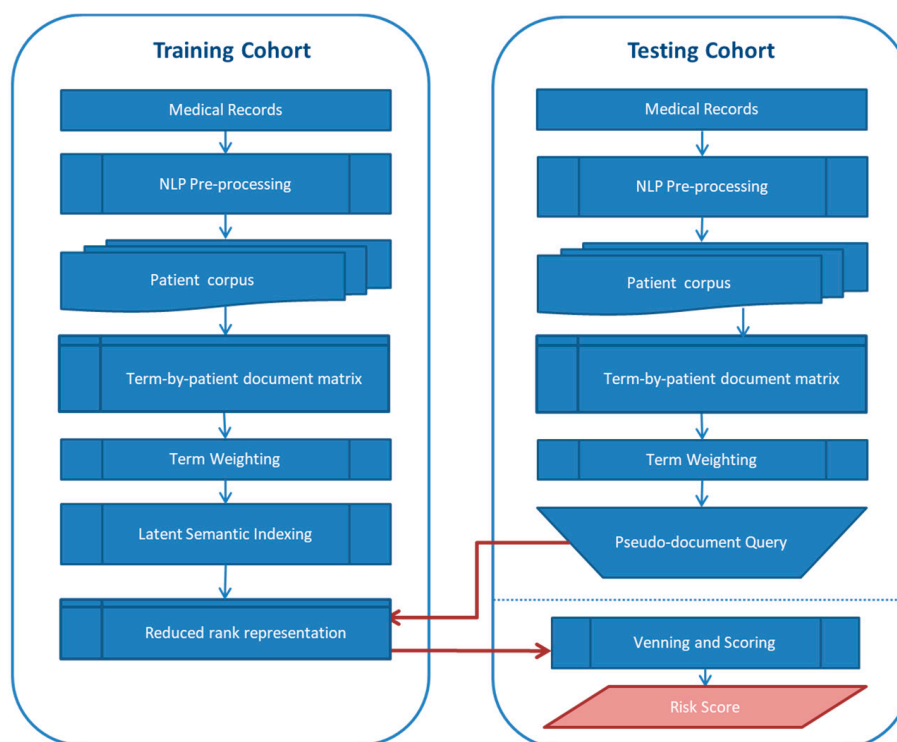


Figure 2. Flow diagram for Quire Predictive Modeling (QPM) algorithm.

The Quire Predictive Modeling (QPM) algorithm (Patent pending, Quire Inc., Memphis, TN, USA) ranks patients in a collection based on the reduced-rank vector cosine values to a set of sentinel patients who exhibit the target outcome (Figure 2). An important advantage of this approach is that the LSI model trained on one cohort can be used to determine a patient’s risk in a different cohort. A risk score is calculated for each patient in the test cohort based on the percentage of sentinels who have cosine similarities above a preset threshold of 0.55. In this study, QPM was built on the training cohort of 4738 patients, which included 48 sentinel patients who died within 30 days of surgery. A separate cohort of 1759 patients was used for testing and evaluation of QPM. Each of the testing cohort were represented by pseudo-document, a vector projected into the training space, to rank and generate individual risk scores. The rankings were normalized on a scale of 0–100, where the patient with cosine similarity (above the threshold) to the largest number of sentinels received a risk score of 100. The risk scores for each patient in the test cohort were then included in regression models as described below.

2.3. Hypothesis Testing and Prediction

We used the Kolmogorov-Smirnov test to check for the normality of data and the Mann-Whitney U test to check whether the distribution of text-based risk variable was the same for different categories of the outcome variables. We implemented stepwise logistic regression analysis with backward elimination in predicting outcomes of cases in the NSQIP cohort by (A) using only text-based risk variables from unstructured data, (B) using only 15 NSQIP risk factors, and (C) using all factors in A and B as predictors. C-statistic calculation was used for model prediction performance, and the DeLong test [28] was used to compare the models. In addition to predicting D30, the model was used to predict 11 separate secondary surgery outcomes including death within 90 days of surgery, unplanned readmission, unplanned readmission to operating room, unplanned repeat surgery related to the principle surgery, unplanned second surgery, blood transfusion within 72 h of surgery start time, postoperative unplanned intubations, postoperative systemic sepsis, septic shock, postoperative superficial incisional surgical site infections (SSI), and postoperative organ/space SSI. For all models, we implemented five-fold cross-validation to avoid and detect possible overfitting.

3. Results

There were records for 8178 operative emergent or urgent cases performed at Le Bonheur Children’s Hospital, Memphis, TN, USA between 1 January 2014 and 31 May 2017 on patients aged 18 years or younger. We excluded 1681 cases without preoperative text notes. A total of 719,308 free text notes were available for the remaining 6497 cases. A total of 4738 patients in the non-NSQIP cohort was used as the training data set to develop the text-based risk model that was then tested on the NSQIP cohort of 1759 surgical cases. We evaluated if the text-based risk score could improve the performance of models that used structured preoperative risk factors to predict surgery outcome.

Mortality rates in the testing and training cohorts were 11/1759 (0.63%) and 48/4738 (1.01%), respectively, and did not show a significant difference ($p = 0.15$). Age at operation was younger ($p < 0.001$) for the NSQIP cohort (mean \pm standard deviation of 6.4 ± 6.0 years vs. 7.1 ± 5.3 years for training). Gender (55% male for both) and race (47% white vs. 45% white, 40% black vs. 43% black for testing vs. training, respectively) were similar in the two cohorts.

3.1. Association between Free Text-Based Risk Score and Death after Surgery

The text-based risk score was trained on the non-NSQIP cohort and then used to calculate the text-based risk score for patients in the NSQIP test cohort as described in the Methods section. We used the Mann-Whitney U test to compare text-based risk scores between D30 patients and those who survived because the text-based risk scores were non-normally distributed (Kolmogorov-Smirnov test $p > 0.1$). We found that the text-based risk scores were significantly higher ($p < 0.001$, Mann-Whitney U Test) for D30 cases compared with those who survived beyond 30 days in both training set and testing set (Table 1). D30 cases were concentrated at higher risk scores, both in the training dataset (data not shown) and in the test set (see Figure 3).

Table 1. Comparison of text-based risk scores for categories of D30.

Data Sets	Categories	Count	Mean Text-Based Risk; 95% CI	Mann-Whitney U Test p Value
Non-NSQIP (Training)	D30 = No	4690	0.35; 0.34–0.36	<0.001
	D30 = Yes	48	0.64; 0.59–0.72	
NSQIP (Testing)	D30 = No	1748	0.44; 0.42–0.46	<0.001
	D30 = Yes	11	0.84; 0.78–0.90	

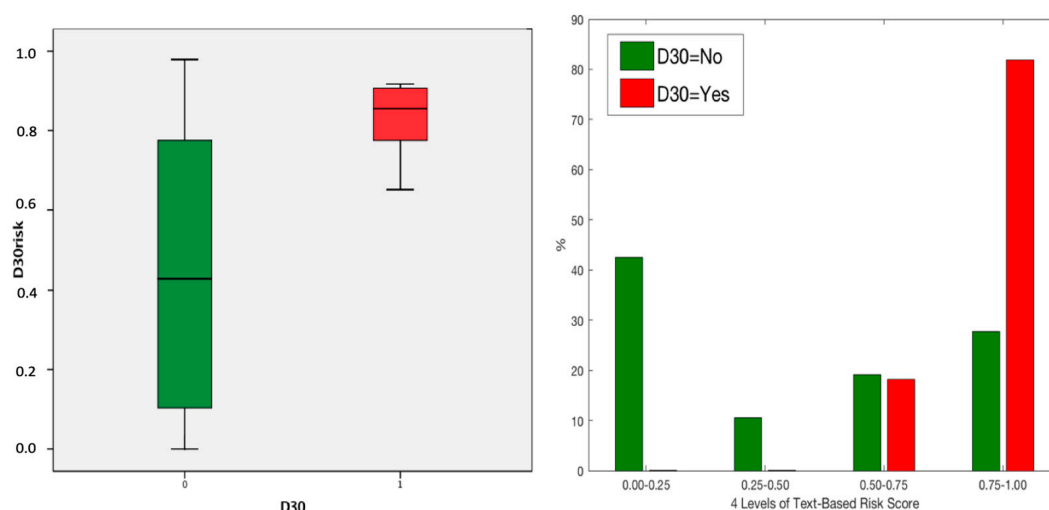


Figure 3. (Left) Text-based risk scores for training cohort for D30: 0 (No, Green) and 1 (Yes, Red). (Right) Observed D30 (percentage) for four different levels of text-based risk scores.

3.2. Sensitivity Analysis for Text-Based Risk Scores

We used the C-statistic to evaluate the performance of QPM on predicting postsurgical mortality risk. We found that using a sentinel pool of 48 (all D30 cases) in the training dataset, QPM achieved a C-statistic of 0.83 (95% confidence interval (CI) 0.77–0.89) on the independent test dataset. To evaluate the sensitivity of QPM with respect to the number of sentinels in the training dataset, we calculated the C-statistic using randomly selected smaller sets of sentinels. The average C-statistic for five randomly selected sets of 10, 20, 30, and 40 D30 sentinels were 0.80, 0.83, 0.83, 0.83 in the testing cohort, respectively (Figure 4). Although the C-statistics for all sets were similar, the variance of the C-statistic between the five random sets was larger with fewer sentinels. These results suggest that as few as 20 sentinel D30 patients in the training cohort could be effectively used to calculate the mortality risk in the test cohort.

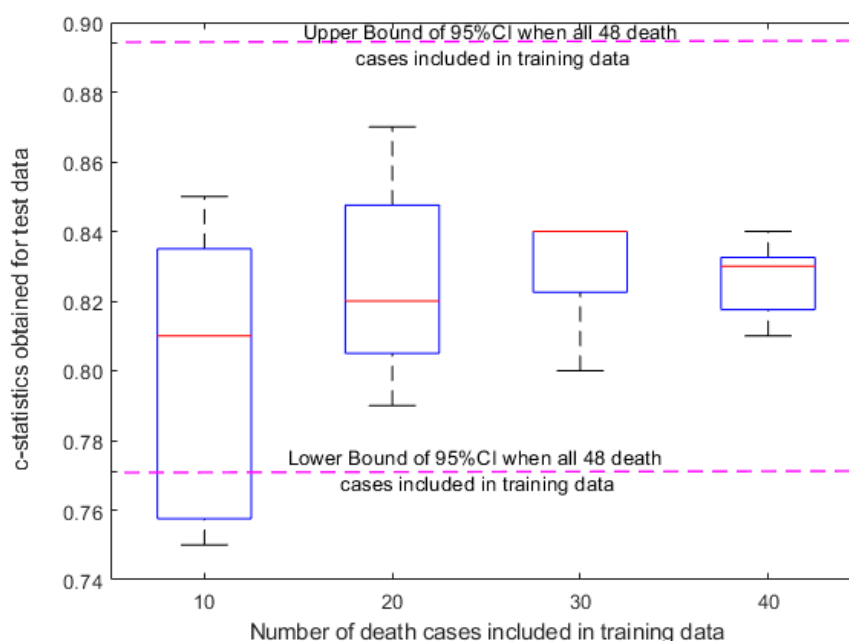


Figure 4. Sensitivity analysis for QPM model.

3.3. Prediction of Postsurgical Mortality in the NSQIP Cohort

We developed three logistic regression models predicting risk of death within 30 days of surgery; Model A using text-based risk score as a single predictor, Model B using 15 risk factors from structured data fields and analyzed with stepwise logistic regression (SWLR), and Model C using all variables from Model A and B and also analyzed with SWLR. Model A had a C-statistic of 0.83 (0.77–0.89, 95% CI) (Supplementary Materials) without cross-validation and 0.81 (0.74–0.88, 95% CI) with five-fold cross-validation. Model B identified ventilator status, bleeding disorder, and inotropic support as significant risk factors, yielding a C-statistic of 0.86 (0.69–1.00, 95% CI) (Supplementary Materials) without cross-validation and 0.76 (0.54–0.99) with five-fold cross-validation. The difference between C-statistics of models A and B was non-significant ($p > 0.1$; DeLong test).

Finally, in Model C, the final logistic regression model selected the text-based risk score, ventilator status, bleeding disorder, current receipt of inotropic support, and emergent case as significant risk factors, resulting in a C-statistic of 0.96 (0.92–1.00, 95% CI) (Supplementary Materials) without cross-validation and 0.92 (0.84–0.99) with five-fold cross-validation using the same five selected variables. The values of regression coefficients and odd ratios of these five variables obtained for each run of cross-validation were found to be within the 95% confidence intervals (Supplementary Materials). The performance of the final logistic regression model including both text-based risk scores and structured data was significantly better than the performance of models using only text-based

risk score ($p = 0.036$; DeLong test) and the model using structured data-based risk factors ($p = 0.055$; DeLong test) in terms of C-statistics.

We used a threshold value to convert predicted probabilities from the combined logistic regression model into binary class predictions for the D30 variable. We identified this threshold value as the cutoff value for cross-validated predictions where F1 score, $(2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$, was maximized. Using this threshold, the model correctly classified 7 of 11 deaths (sensitivity of 63.6%) but also produced 12 false positives (specificity of 99.3%), yielding positive predictive value (precision) of 36.8% and negative predictive value of ~100.0%. When we looked at the 12 false positive predictions, we found that there was one case each of death within 30–90 days of surgery, unplanned postoperative intubation, deep wound disruption, and surgery-related admission within 30 days of surgery, two cases with surgery-related repeat surgery, and four cases of unplanned blood transfusion. In other words, 10 of 12 of the false positives experienced either death after 30 days or other adverse surgery outcomes. We note that these results are based on the cutoff value maximizing the F1-score and higher sensitivity (at the cost of lower specificity) can be calculated for different cutoff values using the same predictive model.

3.4. Association between Free Text-Based Risk Score and Other Adverse Surgery Outcomes

The text-based risk scores for the NSQIP cohort were also significantly ($p < 0.05$, Mann-Whitney U test) associated with other outcomes such as death within 90 days of surgery, intra- or post-operative blood transfusion within 72 h of surgery, unplanned readmission within 30 days of surgery, postoperative unplanned intubation, and first unplanned return to operating room (Table 2). In contrast, the text-based risk scores were not significantly associated with post-operative deep organ space surgical site infection.

Table 2. Distribution of text-based risk scores over categories of binary outcomes for the NSQIP cohort.

Outcome		Count	Mean Text-Based Risk Value with 95%CI	p Value
Death within 30 days of surgery	No	1748	0.44; 0.42–0.45	<0.001
	Yes	11	0.84; 0.78–0.90	
Death within 90 days of surgery	No	1738	0.44; 0.42–0.45	<0.001
	Yes	21	0.82; 0.77–0.87	
Postoperative superficial (incisional) surgical site infection	No	1736	0.44; 0.42–0.45	0.015
	Yes	23	0.62; 0.52–0.74	
Intra- or post-operative blood transfusion within 72 h of surgery start time	No	1625	0.45; 0.43–0.47	<0.001
	Yes	134	0.31; 0.25–0.37	
Unplanned readmission within 30 days of surgery	No	1621	0.43; 0.41–0.45	<0.001
	Yes	138	0.57; 0.51–0.62	
Postoperative Unplanned Intubation	No	1735	0.43; 0.42–0.45	0.001
	Yes	24	0.71; 0.61–0.81	
First Unplanned Return to Operating Room	No	1690	0.44; 0.42–0.45	0.039
	Yes	69	0.52; 0.43–0.60	

3.5. The Role of Free Text-Based Risk Score in Predicting Other Adverse Surgery Outcomes

The text-based risk score (derived for predicting D30) was significantly predictive of death between 30–90 days after surgery (C-statistic 0.96, 0.92–0.99 95% CI), along with additional outcomes such as postoperative superficial incisional surgical site infection, intra- or post-operative blood transfusion within 72 h of surgery start time, and unplanned readmission within 30 days of surgery (Table 3). Table 3 also includes the five-fold cross-validation results for each logistic regression model. More details about the logistic regression models can be found in the Supplementary Materials.

Table 3. Logistic regression of outcome including text-based risk score as a predictor.

Outcome	c-Statistics with 95%CI		Selected Preoperative Risk Factors
Death within 30 days of surgery	No CV Five-fold CV	0.96; 0.92–1.00 0.92; 0.84–0.99	Text-based risk score, ventilator dependency, bleeding disorder, inotropic support, Emergent Case
Death within 90 days of surgery	No CV Five-fold CV	0.95; 0.92–0.99 0.94; 0.89–0.99	Text-based risk score, ventilator dependency, neonate, bleeding disorder, Emergent case
Postoperative superficial incisional surgical site infection	No CV Five-fold CV	0.72; 0.61–0.83 0.67; 0.55–0.79	Text-based risk score, neonate
Intra- or post-operative blood transfusion within 72 h of surgery start time	No CV Five-fold CV	0.76; 0.71–0.80 0.73; 0.69–0.78	Text-based risk score, oxygen support, neuromuscular disorder, hematologic disorder, inotropic support, malignancy, urgent case
Unplanned readmission within 30 days of surgery	No CV Five-fold CV	0.67; 0.62–0.72 0.66; 0.61–0.70	Text-based risk score, neonate, SIRS, Sepsis

4. Discussion

Our study suggests that unstructured preoperative text available in EMRs contains critical information predictive of postoperative death in children undergoing surgical procedures. Further, these data suggest that information contained in unstructured text notes can be useful even when distilled to a single risk variable developed via a text modeling approach. Finally, we found that the use of text-based risk scores combined with structured data improves the prediction accuracy of death within 30 days of surgery when compared with models using either unstructured or structured data from the NSQIP database alone.

Data from the unstructured notes is currently used in creating clinically useful risk assessments for surgical procedures [29]. An example is the ASA class that is included as a key variable in the Pediatric Risk Calculator by NSQIP [30] developed by American College of Surgeons. Automated systems utilizing algorithms such as the one used in this study have the potential to decrease bias introduced by human retrieval and interpretation of such data and may save time for clinicians.

The clinical utility of any risk assessment depends on its accuracy. US health expenditures are higher than other technically advanced countries reporting better objective health outcomes due in part to the provision of expensive care that is unlikely to provide meaningful benefit [31]. Sharing accurate risk estimates with the patient and family is a key component of informed consent and shared decision-making. This allows providers and consumers of surgical care to better weigh alternative treatments that may be less expensive and have equivalent benefit, or to forego treatment in settings where the probability of death after surgery approaches certainty. Formal studies of the impact of clinical decision support tools for surgery are limited, but are needed to determine their impact on practice and patient outcomes.

Our prediction model was created to predict the risk of death within 30 days of surgery. The finding that the text-based risk score also contributed to accurate prediction of other major surgery outcomes such as death within 90 days of surgery, postoperative surgical site infection, and unplanned blood transfusion and readmission within 30 days of surgery suggest that postoperative adverse events are interrelated. Logistic regression models for each of these outcomes performed almost equally well in five-fold cross-validation, suggesting that the logistic regression models built on the text-based risk score are robust and generalizable to a broad variety of adverse events despite the challenges of a relatively small sample size and low event rate.

Limitations

Our study has some limitations. The training set included all types of operations while the test set, NSQIP cohort, systematically excluded some operations [26]. Since the text-based risk scores are calculated based on vectorized combinations of thousands of terms extracted from patient records, the precise words that contribute to postsurgical mortality risk is difficult to deduce. The next step in our work will be to investigate specific keywords that are associated with modifiable risk factors to guide clinicians in developing interventions designed to reduce the risk of severe surgery outcomes.

Our study presents a proof of concept on the significance of unstructured text notes in predicting post-operative adverse outcomes in children. Its clinical applicability requires further work on both improving the predictive model performance and external validation.

5. Conclusions

We conclude that text data in EMRs can improve the ability of structured data tools to predict serious patient outcomes after surgery.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2227-9709/6/1/4/s1>, Supplementary files S1 and S2.

Author Contributions: O.A. and R.L.D. designed and conceptualized the research. O.A. performed the statistical analysis and drafted the manuscript. R.H. and K.H. carried out text data acquisition and text mining. M.R.L.J. provided cleaned structured data and clinical outcomes. All authors contributed to reviewing and writing of the manuscript.

Funding: This research received no external funding.

Acknowledgments: We would like to extend our appreciation to Mark McMath, Anuj Mathur, Deepthi Kurup, and Brian Marshall for making the EMR data available for this study. We also thank Robert Beck for abstracting NSQIP data, Ari Walter and Ruchi Ruchi for gathering and validating data and Shane Morrell for processing the text data and running the Quire applications. We would also like to thank Amada Preston from the LeBonheur Scientific Writing Office and Rich Redfearn and Patti Smith for editing the manuscript.

Conflicts of Interest: Authors K.H. and R.H. are affiliated with Quire Inc. (Memphis, TN, USA). Quire personnel were involved in data extraction from the EMR system at Methodist Hospital (Memphis, TN, USA) under a Business Associates Agreement. The other authors have no conflict of interest.

References

1. BS Systems. *Health IT and Patient Safety: Building Safer Systems for Better Care*; BS Systems: Zusmarshausen, Germany, 2011.
2. Van der Griend, B.F.; Lister, N.A.; McKenzie, I.M.; Martin, N.; Ragg, P.G.; Sheppard, S.J.; Davidson, A.J. Postoperative mortality in children after 101,885 anesthetics at a tertiary pediatric hospital. *Anesth. Analg.* **2011**, *112*, 1440–1447. [[CrossRef](#)] [[PubMed](#)]
3. Akbilgic, O.; Langham, M.R.; Walter, A.I.; Jones, T.L.; Huang, E.Y.; Davis, R.L. A novel risk classification system for 30-day mortality in children undergoing surgery. *PLoS ONE* **2018**, *13*, e0191176. [[CrossRef](#)] [[PubMed](#)]
4. Cheon, E.C.; Palac, H.L.; Paik, K.H.; Hajduk, J.; De Oliveira, G.S.; Jagannathan, N.; Suresh, S. Unplanned, Postoperative Intubation in Pediatric Surgical Patients: Development and Validation of a Multivariable Prediction Model. *Anesthesiology* **2016**, *125*, 914–928. [[CrossRef](#)] [[PubMed](#)]
5. Stey, A.M.; Vinocur, C.D.; Moss, R.L.; Hall, B.L.; Cohen, M.E.; Kraemer, K.; Ko, C.Y.; Kenney, B.D. Variation in intraoperative and postoperative red blood cell transfusion in pediatric surgery. *Transfusion* **2016**, *56*, 666–672. [[CrossRef](#)] [[PubMed](#)]
6. Brown, E.G.; Anderson, J.E.; Burgess, D.; Bold, R.J.; Farmer, D.L. Pediatric surgical readmissions: Are they truly preventable? *J. Pediatr. Surg.* **2017**, *52*, 161–165. [[CrossRef](#)] [[PubMed](#)]
7. Oldham, K.T. Optimal Resources for Children's Surgical Care. *J. Pediatric Surg.* **2015**, *49*, 667–677. [[CrossRef](#)] [[PubMed](#)]
8. Langham, M.R.; Walter, A.; Boswell, T.C.; Beck, R.; Jones, T.L. Identifying children at risk of death within 30 days of surgery at an NSQIP pediatric hospital. *Surgery* **2015**, *158*, 1481–1491. [[CrossRef](#)]
9. Akbilgic, O.; Langham, M.R.; Davis, R.L. Race, Preoperative Risk Factors, and Death after Surgery. *Pediatrics* **2018**, *141*, e20172221. [[CrossRef](#)]
10. Harris, A.H.S. Path from predictive analytics to improved patient outcomes. *Ann. Surg.* **2017**, *265*, 461–463. [[CrossRef](#)]
11. Bilimoria, K.Y.; Liu, Y.; Paruch, J.L.; Zhou, L.; Kmiecik, T.E.; Ko, C.Y.; Cohen, M.E. Development and evaluation of the universal ACS NSQIP surgical risk calculator: A decision aid and informed consent tool for patients and surgeons. *J. Am. Coll. Surg.* **2013**, *217*, 833–842. [[CrossRef](#)]

12. Parikh, R.B.; Kakad, M.; Bates, D.W. Integrating Predictive Analytics into High-Value Care. *JAMA* **2016**, *315*, 651–652. [[CrossRef](#)] [[PubMed](#)]
13. Amarasingham, R.; Audet, A.M.; Bates, D.W.; Cohen, I.G.; Entwistle, M.; Escobar, G.J.; Liu, V.; Etheredge, L.; Lo, B.; Ohno-Machado, L.; et al. Consensus Statement on Electronic Health Predictive Analytics: A Guiding Framework to Address Challenges. *eGEMs* **2016**, *4*, 1163. [[CrossRef](#)] [[PubMed](#)]
14. Jensen, K.; Soguero-Ruiz, C.; Mikalsen, K.O.; Lindsetmo, R.O.; Kouskoumvekaki, I.; Girolami, M.; Skrovseth, S.O.; Augestad, K.M. Analysis of free text in electronic health records for identification of cancer patient trajectories. *Sci. Rep.* **2017**, *7*, 46226. [[CrossRef](#)]
15. Lin, J.; Jiao, T.; Biskupiak, J.E.; McAdam-Marx, C. Application of electronic medical record data for health outcomes research: A review of recent literature. *Expert Rev. Pharmacoecon. Outcomes Res.* **2013**, *13*, 191–200. [[CrossRef](#)] [[PubMed](#)]
16. Friedman, C.; Shagina, L.; Lussier, Y.; Hripcsak, G. Automated encoding of clinical documents based on natural language processing. *J. Am. Med. Inform. Assoc.* **2004**, *11*, 392–402. [[CrossRef](#)]
17. Zeng, Q.T.; Goryachev, S.; Weiss, S.; Sordo, M.; Murphy, S.N.; Lazarus, R. Extracting principal diagnosis, co-morbidity and smoking status for asthma research: Evaluation of a natural language processing system. *BMC Med. Inform. Decis. Mak.* **2006**, *6*, 30. [[CrossRef](#)] [[PubMed](#)]
18. Savova, G.K.; Masanz, J.J.; Ogren, P.V.; Zheng, J.; Sohn, S.; Kipper-Schuler, K.C.; Chute, C.G. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications. *J. Am. Med. Inform. Assoc.* **2010**, *17*, 507–513. [[CrossRef](#)]
19. Lin, F.P.Y.; Pokorny, A.; Teng, C.; Epstein, R.J. TEPAPA: A novel in silico feature learning pipeline for mining prognostic and associative factors from text-based electronic medical records. *Sci. Rep.* **2017**, *7*, 6918. [[CrossRef](#)]
20. Ford, E.; Carroll, J.A.; Smith, H.E.; Scott, D.; Cassell, J.A. Extracting information from the text of electronic medical records to improve case detection: A systematic review. *J. Am. Med. Inform. Assoc.* **2016**, *23*, 1007–1015. [[CrossRef](#)]
21. Kim, Y.S.; Yoon, D.; Byun, J.; Park, H.; Lee, A.; Kim, I.H.; Lee, S.; Lim, H.S.; Park, R.W. Extracting information from free-text electronic patient records to identify practice-based evidence of the performance of coronary stents. *PLoS ONE* **2017**, *12*, e0182889. [[CrossRef](#)]
22. Thomas, A.A.; Zheng, C.; Jung, H.; Chang, A.; Kim, B.; Gelfond, J.; Slezak, J.; Porter, K.; Jacobsen, S.J.; Chien, G.W. Extracting data from electronic medical records: Validation of a natural language processing program to assess prostate biopsy results. *World J. Urol.* **2014**, *32*, 99–103. [[CrossRef](#)] [[PubMed](#)]
23. Frost, D.W.; Vembu, S.; Wang, J.; Tu, K.; Morris, Q.; Abrams, H.B. Using the Electronic Medical Record to Identify Patients at High Risk for Frequent Emergency Department Visits and High System Costs. *Am. J. Med.* **2017**, *130*, 601.e17–601.e22. [[CrossRef](#)] [[PubMed](#)]
24. Weissman, G.E.; Hubbard, R.A.; Ungar, L.H.; Harhay, M.O.; Greene, C.S.; Himes, B.E.; Halpern, S.D. Inclusion of Unstructured Clinical Text Improves Early Prediction of Death or Prolonged ICU Stay. *Crit. Care Med.* **2018**, *46*, 1125–1132. [[CrossRef](#)] [[PubMed](#)]
25. Murff, H.J.; FitzHenry, F.; Matheny, M.E.; Gentry, N.; Kotter, K.L.; Crimin, K.; Dittus, R.S.; Rosen, A.K.; Elkin, P.L.; Brown, S.H.; et al. Automated identification of postoperative complications within an electronic medical record using natural language processing. *JAMA* **2011**, *306*, 848–855. [[CrossRef](#)] [[PubMed](#)]
26. *American College of Surgeons National Surgical Quality Improvement Program-Pediatrics User Guide for the ACS NSQIP Pediatric Participant Use File*; American College of Surgeons: Chicago, IL, USA, 2013.
27. Berry, M.; Browne, M. *Understanding Search Engines: Mathematical Modeling and Text Retrieval*, 1st ed.; SIAM: Philadelphia, PA, USA, 1999.
28. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **1988**, *44*, 837–845. [[CrossRef](#)] [[PubMed](#)]
29. Merrick, S.; Shaker, M. ASA Relative Value Guide (RVG): A Defining Moment in Fair Pricing of Medical Services. *ASA Monit.* **2014**, *78*, 26–27.

30. Kraemer, K.; Cohen, M.E.; Liu, Y.; Barnhart, D.C.; Rangel, S.J.; Saito, J.M.; Bilimoria, K.Y.; Ko, C.Y.; Hall, B.L. Development and Evaluation of the American College of Surgeons NSQIP Pediatric Surgical Risk Calculator. *J. Am. Coll. Surg.* **2016**, *223*, 685–693. [[CrossRef](#)]
31. Moses, H.; Matheson, D.H.M.; Dorsey, E.R.; George, B.P.; Sadoff, D.; Yoshimura, S. The anatomy of health care in the United States. *JAMA* **2013**, *310*, 1947–1963. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).