HHS Public Access
Author manuscript
Intell Based Med. Author manuscript; available in PMC 2025 March 21.

Published in final edited form as:
Intell Based Med. 2021 ; 5: . doi:10.1016/j.ibmed.2021.100028.

Physician documentation matters. Using natural language processing to predict mortality in sepsis

Keaton Cooley-Rieders[a,*], **Kai Zheng**[b]

[a]School of Medicine, University of California Irvine, 1001 Health Sciences Road, Irvine, CA, 92617, USA

[b]Department of Informatics, University of California Irvine, 6095 Donald Bren Hall, Irvine, CA, 92687, USA

Abstract

Background/objective: Sepsis remains without good outcome prediction. Technological advances, specifically, natural language processing (NLP), has an opportunity to approach sepsis mortality prediction in a novel way.

Methods: Using the MIMIC III dataset, patients diagnosed with sepsis from 2008 to 2013 had physician progress notes analyzed using NLP. Researchers utilized concepts from analysis to build a model to predict for in-hospital-mortality, using notes in the first 24 hours of a patient admission. This model was retrospectively validated on septic admissions to University of California Irvine Medical Center (UCIMC) from 2013 to 2018 and compared to SOFA and qSOFA.

Results: An 80-concept model was developed and validated on 7117 admissions to UCIMC. For severe sepsis, an Area Under Curve or AUC of 0.687 (95% CI 0.618–0.748) was demonstrated which was greater than SOFA at 0.571 (0.497–0.643). Additionally, for simple sepsis the model demonstrated an AUC of 0.696 (0.649–0.738) which was greater than qSOFA at 0.590 (0.545–0.638).

Conclusions: Physician clinical judgement extracted from notes using NLP has greater performance in predicting mortality and survival in sepsis compared to structured data used in SOFA and qSOFA.

Keywords

NLP; Natural language processing; Sepsis; Mortality; Screening; Early Recognition

1. Introduction

Sepsis, or severe infections, hospitalize millions of individuals yearly in the United States [1–3]. Despite rapid advances in medicine to improve the outcomes and quality of life of

*Corresponding author. kcooleyr@uci.edu (K. Cooley-Rieders), kai.zheng@uci.edu (K. Zheng).

Declaration of competing interest
The authors declare no conflicts of interest.

individuals with numerous other pathologies, management of patients with sepsis remains a challenge. A large proportion of individuals die despite best efforts [1–6].

Extensive efforts by many disciplines have focused on changing outcomes in this pathology, including research for better treatments, earlier identification, and more accurate mortality prediction. Mortality prediction has been a topic of much interest, with hopes that early identification of individuals who may experience poor outcomes can lead to closer monitoring, assessment of tissue perfusion, and appropriate timely antimicrobial interventions [7,8]. The implementation of the Sequential Organ Failure Assessment or SOFA score and its quicker counterpart, qSOFA seen in Fig. 1, were developed and endorsed by the Society of Critical Care Medicine to do just this, and stratify patients for risk of mortality with sepsis [9].

Unfortunately, SOFA and qSOFA have proven to have either poor utility for screening for the outcomes they were intended to measure, or poor specificity [10–15]. Numerous studies have also suggested that scores from these tools are too late to offer clinical utility, and researchers have even demonstrated that significant modifications beyond the intended use of these scores are required to have performance adequate for meaningful clinical use [4,16–18].

Additionally, in the age of the electronic health record where data is more and more accessible on a large scale for analysis, technologies such as machine learning and Natural Language Processing or NLP have been leveraged to help improve early identification and more accurate mortality prediction in patients with sepsis [19,20]. Application has mainly focused on using structured data, including vital signs, lab values, and other physiological parameters to develop models [21–24]. However focus on unstructured data such as notes, has also been a focus of very recent research to reach the same end goals [19,20].

It is clear that there is much room for improvement in the clinical prediction of patient outcomes, specifically with the use of unstructured data, which has the potential to positively impact the care of many septic individuals.

2. Methods

Investigators sought out to create a novel scoring system for patients diagnosed with sepsis using unstructured data from the first 24 hours of an admission to predict in-hospital mortality.

Using the MIMIC III dataset, composed of thousands of patients admitted to critical care units at Beth Israel Deaconess Medical Center, adult patients with sepsis, with physician progress notes, were used to model a scoring system [25]. These individuals from 2008 to 2013 were selected by ICD code for sepsis, at a time period when Sepsis-2 criteria were used to make this diagnosis.

Unstructured data would be derived from clinician notes in the first 24 hours of an admission. Using CLAMP, a Natural Language Processing tool leveraging validated technologies such as OpenNLP and CRF, notes were converted into UMLS concepts

[26]. For simplicity, no features such as negation or lab-value detection were used. CLAMP utilizes best in class tools for NLP including OpenNLP, and CRF, with validated performance across multiple corpora with precision of 92–96% and recall of 89–94%. This tool also demonstrates greater performance when compared to other tools such as MetaMap and cTAKES [26].

Concepts that demonstrated high strength combined with low to moderate frequency were hypothesized to be most desireable for model use. To select concepts that had both these qualities, investigators needed a surrogate for strength and frequency. For strength each concept was statistically examined with respect to in-hospital-mortality by hazard ratios at a 0.05 significance level and confidence interval. Only statistically significant concepts with confidence intervals that did not cross 1, were used in further model development. Concepts with intervals including 1 suggest no difference with respect to outcomes and therefore were not used. Frequency was determined with the formula [(# of admissions with concept/total # of admissions)]. Concepts that were present in all admissions at less than 50-percent frequency were then used.

To select concepts, a simple decision tree was programmed to automatically evaluate performance of different combined groups of concepts as seen in Fig. 2. Concepts if present in an admission in the first 24 hours, and associated with a hazard ratio of greater 1 would be given a positive-one score. A negative-one score for the presence of similar concepts with hazard ratio less than 1 were accordingly assigned. These summative scores per admission were evaluated at a greater than or equal to 0 cut-off for sensitivity [true positive/(true positive + false negative)], specificity [true negative/(true negative + false positive)], and accuracy [sensitivity × specificity]. An enforced minimum increase of half-a-percent in accuracy per added concept was utilized to prevent the use of many very infrequent concepts, which may lead to poor generalizability.

Subsequent combinations of concepts were sorted by sensitivity and combined until a maximum number of admissions were scored while also maintaining a model with good screening characteristics. Receiver operator curves (ROC) were evaluated for end models. These end models would be compared to the ROC for SOFA, using the criteria seen in Fig. 1.

Automatically derived models were validated on 5 years of adult severe sepsis and septic admissions to University of California Irvine Medical Center (UCIMC) from 2013 to 2018 by ICD code. Of note Sepsis-3 criteria were published in early 2016, however ICD coding and billing was via Sepsis-2 criteria throughout the 5 year period. Models would be compared to SOFA for severe sepsis admissions, and qSOFA, for simple sepsis.

Study activities were reviewed by the office of research, human research protections, and provided institutional review board approval.

3. Results

Patient demographics can be seen in Table 1 for both MIMIC III and UCIMC data. Demographics stratified by simple and severe sepsis can also be seen in this table. Of 1052 admissions in MIMIC III data, all had available progress notes for evaluation in the first 24 hours. 79.3% of 8973 UCIMC admissions had available progress notes in the first 24 hours to test subsequent models.

645 concepts were found from MIMIC to have statistically significant hazard ratios with a 95% confidence interval. An 80 concept model was discovered using the automated decision tree in Fig. 2. Selected concepts with hazard ratios and frequency on MIMIC III data can be seen in Table 2.

On MIMIC III, the model demonstrated sensitivity of 93.2% (95% CI 90.2–96.1%) and specificity of 39.2% (33.6–44.8%) using the specified cutoff of greater than or equal to 0. Receiver Operating Curve or ROC can be seen in Fig. 3 and demonstrated an Area Under Curve (AUC) of 0.783 (95% CI 0.707–0.854). This AUC was statistically greater ($p < 0.001$) than that of SOFA at 0.604 (0.502–0.698) which can also be seen in Fig. 3.

Subsequently with validation on severe sepsis admissions to UCIMC, using the specified cutoff of greater than or equal to 0, sensitivity was 88.3% (95% CI 85.6–90.0%) with specificity of 29.0% (24.7–33.2%). A ROC AUC of 0.687 (95% CI 0.618–0.748) was obtained. This was statistically greater than SOFA ($p < 0.001$) which demonstrated a ROC AUC of 0.571 (0.497–0.643). Comparison of these ROC's can be seen in Fig. 4.

With application of the NLP model to simple sepsis admissions, using the specified cutoff of greater than or equal to 0, sensitivity was 86.0% (95% CI 84.0–88.1%) with specificity of 32.2% (29.6–34.9%). A ROC AUC of 0.696 (95% CI 0.649–0.738) was revealed which was statistically greater than qSOFA ($p < 0.001$) with a ROC AUC of 0.590 (0.545–0.638). With a cutoff of >2 qSOFA had a sensitivity of 55.0% (95% CI 51.4–58.7%) and specificity of 59.3% (57.2–61.3%). These ROC's can also be seen in Fig. 4.

4. Discussion

Mortality prediction tools such as SOFA and qSOFA have demonstrated poor performance for their intended use of screening and predicting mortality in sepsis [10–16,27]. This poor performance has led to clinician frustration and research to find better approaches to screen and stratify patients with severe infections [4,17,18,28].

Previous studies have demonstrated that the use of NLP extracted concepts can be utilized to develop tools to predict the presence of sepsis with good performance in combination with other structured forms of data [19,20,23]. Other models have been built and validated to perform the same task using solely structured forms of data [21–24]. However, this study was the first to utilize only unstructured concepts from physician progress notes to develop a validated in-hospital-mortality outcome screening tool with head to head comparison to SOFA and qSOFA.

Utilizing only unstructured data in the tool described previously and its comparison to existing scoring systems utilizing structured data allows three major conclusions to be made. First, attention is drawn to the power of clinical judgement in predicting patient outcomes and its relative performance to quantitative measures of clinical condition such as lab values and vital signs. Second, isolation of data types emphasizes the timeliness of unstructured data in providing prediction in patient outcomes compared to structured data. Finally, the statistically significant elements that make up clinical judgment can be elucidated, bringing insight into physicians' processes in coming to conclusions on patient condition.

Unlike other studies with unelucidated machine learning models only available to the researchers who produced them, this study presents a reproducible model. The positive-one, negative-one scoring system based on the presence of a given concept per admission derived from CLAMP, an accessible NLP tool that is easy to use, allows for application of the developed model and further research. Particular weight to different concepts was not undertaken for simplicity, however hazard ratios and frequencies outlined in Table 2 provide a foundation for further research to determine if weighting of concepts increases performance.

The observed performance of the 80 concept model may be argued as requiring further refinement before clinical use. However, existing previously mentioned tools have demonstrated much worse performance in studies which has again been elucidated in this study [10–14,16,27]. Additionally, the timely results provided by the model, within 24 hours of a patient admission, may allow for earlier identification of patients who may require more intensive intervention.

Documented labs, events, physical exam findings, diagnoses, and treatments used in the end model highlight the wealth of information in unstructured electronic health care data that can be leveraged to predict patient outcomes. Also, individual, statistically significant, observed hazard ratios from MIMIC III data were consistent with described literature on sepsis. For example, comorbidities and organ failure, known to be associated with poor outcomes, were found in this study to be associated with in-hospital-mortality [6,29–31]. The concept of fluid overload and its association with in-hospital-mortality was consistent with recent research demonstrating the importance of guided fluid resuscitation [32–35]. Concepts with reference to influenza vaccination and stress ulcer prophylaxis found to be associated with survival-to-discharge, reflect available evidence as well [36–40].

However, other identified statistically significant concepts were more surprising. Interestingly, the majority of concepts referring to physical exam findings utilized in the end model were associated with survival-todischarge while lab results were mostly associated with in-hospital-mortality. It is unclear how to interpret these findings, and further research will be required to elucidate if any conclusions can be drawn from this trend and reach a greater understanding of the important elements of clinical judgment.

It is important to note that there are a number of limitations to research findings. First, the retrospective design of the study limits generalizability. Additionally, the NLP derived model was used to predict outcomes in those already diagnosed with sepsis, not the presence of sepsis.

Use of ICD code to identify patients may have an unknown impact on findings. Literature has suggested that there may be differences between clinical sepsis and coding data [2,41]. Further, transition of Sepsis-2 definitions used during the study period to Sepsis-3 may confound results. Future application of the system above may elucidate if these changes have a meaningful impact.

It is not known the impact of including structured data in combination with unstructured data, which was not the intent of this study. The findings presented above show the utility of unstructured data, an expression of clinical judgement, and its ability to outperform structured data in mortality prediction in sepsis. Further research combining both data types is warranted to assess further improvements that can be made.

Written physician language and documentation is complex, with different words reflecting the same meaning. Accordingly, investigators hypothesize that different, yet related concepts from the identified model could also be used to attain similar performance for mortality screening in sepsis.

Because CLAMP has been validated in prior studies, performance was not assessed. Additional research will have to be conducted to determine if utilization of a larger feature set, for example negation, can further improve model performance.

Application of SOFA was intended to be used in Intensive Care Unit or ICU environments, while qSOFA outside of the ICU. Investigators did not have access to level of care data for UCIMC data, so the decision to use simple and severe sepsis definitions for application of relative scoring systems was used. This decision was unlikely to impact performance of qSOFA results, as it was likely none or few ICU admissions were in this cohort. However, it was possible that non-ICU admissions were included in patients with severe sepsis and SOFA performance could be affected.

Although the developed model may facilitate early identification of patients who may experience in-hospital-mortality, interventions to prevent this negative outcome are still being actively researched [7,8,35,42–44]. Finally, the impact of clinician knowledge on statistically significant concepts may influence documentation behaviors and impact future performance and model application.

5. Conclusions

An 80 concept validated model based only on NLP derived concepts in physician progress notes, has improved performance to SOFA and qSOFA in predicting mortality and survival in sepsis within 24 hours of patient admission. Concepts expressed in physician progress notes serve as the foundation of the proposed model, emphasizing the importance of documented clinical judgement in the electronic health record.

Funding information

This study was supported by a Summer Research Scholarship award received from the School of Medicine, University of California, Irvine; and in part by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant (UL1 TR000153). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

References

[1]. Bauer M, Gerlach H, Vogelmann T, Preissing F, Stiefel J, Adam D. Mortality in sepsis and septic shock in Europe, North America and Australia between 2009 and 2019-results from a systematic review and meta-analysis. Crit Care 2020;24(1):239. 10.1186/s13054-020-02950-2. [PubMed: 32430052]

[2]. Rhee C, Dantes R, Epstein L, et al. Incidence and trends of sepsis in US hospitals using clinical vs claims data, 2009–2014. J Am Med Assoc 2017;318(13):1241–9. 10.1001/jama.2017.13836.

[3]. Paoli CJ, Reynolds MA, Sinha M, Gitlin M, Crouser E. Epidemiology and costs of sepsis in the United States-an analysis based on timing of diagnosis and severity level. Crit Care Med 2018;46(12):1889–97. 10.1097/CCM.0000000000003342. [PubMed: 30048332]

[4]. Rhee C, Zhang Z, Kadri SS, et al. Sepsis surveillance using adult sepsis events simplified eSOFA criteria versus sepsis-3 sequential organ failure assessment criteria. Crit Care Med 2019;47(3):307–14. 10.1097/CCM.0000000000003521. [PubMed: 30768498]

[5]. Rudd KE, Johnson SC, Agesa KM, et al. Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. Lancet 2020;395(10219):200–11. 10.1016/S0140-6736(19)32989-7. [PubMed: 31954465]

[6]. Strandberg G, Walther S, Agvald O€ hman C, Lipcsey M. Mortality after severe sepsis and septic shock in Swedish intensive care units 2008–2016-A nationwide observational study. Acta Anaesthesiol Scand 2020. 10.1111/aas.13587. March.

[7]. Pruinelli L, Westra BL, Yadav P, et al. Delay within the 3-hour surviving sepsis campaign guideline on mortality for patients with severe sepsis and septic shock. Crit Care Med 2018;46(4):500–5. 10.1097/CCM.0000000000002949. [PubMed: 29298189]

[8]. Rhee C, Filbin MR, Massaro AF, et al. Compliance with the national SEP-1 quality measure and association with sepsis outcomes: a multicenter retrospective cohort study. Crit Care Med 2018;46(10):1585–91. 10.1097/CCM.0000000000003261. [PubMed: 30015667]

[9]. Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score-development, utility and challenges of accurate assessment in clinical trials. Crit Care 2019;23(1):374. 10.1186/s13054-019-2663-7. [PubMed: 31775846]

[10]. Liu Z, Meng Z, Li Y, et al. Prognostic accuracy of the serum lactate level, the SOFA score and the qSOFA score for mortality among adults with Sepsis. Scand J Trauma Resuscitation Emerg Med 2019;27(1):51. 10.1186/s13049-019-0609-3.

[11]. Serafim R, Gomes JA, Salluh J, Póvoa P. A comparison of the quick-SOFA and systemic inflammatory response syndrome criteria for the diagnosis of sepsis and prediction of mortality: a systematic review and meta-analysis. Chest 2018;153(3):646–55. 10.1016/j.chest.2017.12.015. [PubMed: 29289687]

[12]. Raith EP, Udy AA, Bailey M, et al. Prognostic accuracy of the SOFA score, SIRS criteria, and qSOFA score for in-hospital mortality among adults with suspected infection admitted to the intensive care unit. J Am Med Assoc 2017;317(3):290–300. 10.1001/jama.2016.20328.

[13]. de Grooth H-J, Geenen IL, Girbes AR, Vincent J-L, Parienti J-J, Oudemans-van Straaten HM. SOFA and mortality endpoints in randomized controlled trials: a systematic review and meta-regression analysis. Crit Care 2017;21(1):38. 10.1186/s13054-017-1609-1. [PubMed: 28231816]

[14]. Holder AL, Overton E, Lyu P, et al. Serial daily organ failure assessment beyond ICU day 5 does not independently add precision to ICU risk-of-death prediction. Crit Care Med 2017;45(12):2014–22. 10.1097/CCM.0000000000002708. [PubMed: 28906286]

[15]. Anand V, Zhang Z, Kadri SS, Klompas M, Rhee C, CDC Prevention Epicenters Program. Epidemiology of quick sequential organ failure assessment criteria in undifferentiated patients and association with suspected infection and sepsis. Chest 2019;156(2):289–97. 10.1016/j.chest.2019.03.032. [PubMed: 30978329]

[16]. Kievlan DR, Zhang LA, Chang C-CH, Angus DC, Seymour CW. Evaluation of repeated quick sepsis-related organ failure assessment measurements among patients with suspected infection. Crit Care Med 2018;46(12):1906–13. 10.1097/CCM.0000000000003360. [PubMed: 30130261]

[17]. Alkhatib A, Price LL, Esteitie R, LaCamera P. A predictive model for acute respiratory distress syndrome mortality using red cell distribution width. Crit Care Res Pract 2020;2020:3832683. 10.1155/2020/3832683. [PubMed: 32399293]

[18]. Zhang X, Wang Q, Luo X, Lei Y, Huang X. [Prognostic value of Charlson weighted index of comorbidities combined with sequential organ failure assessment score and procalcitonin in patients with sepsis]. Zhonghua Wei Zhong Bing Ji Jiu Yi Xue 2019;31(11):1335–9. 10.3760/cma.j.issn.2095-4352.2019.11.005. [PubMed: 31898562]

[19]. Marafino BJ, Park M, Davies JM, et al. Validation of prediction models for critical care outcomes using natural language processing of electronic health record data. JAMA Netw Open 2018;1(8):e185097. 10.1001/jamanetworkopen.2018.5097. [PubMed: 30646310]

[20]. Liu R, Greenstein JL, Sarma SV, Winslow RL. Natural language processing of clinical notes for improved early prediction of septic shock in the ICU. Conf Proc IEEE Eng Med Biol Soc 2019;2019:6103–8. 10.1109/EMBC.2019.8857819.

[21]. Pirracchio R, Petersen ML, Carone M, Rigon MR, Chevret S, van der Laan MJ. Mortality prediction in intensive care units with the Super ICU Learner Algorithm (SICULA): a population-based study. Lancet Respir Med 2015;3(1):42–52. 10.1016/S2213-2600(14)70239-5. [PubMed: 25466337]

[22]. Taylor RA, Pare JR, Venkatesh AK, et al. Prediction of in-hospital mortality in emergency department patients with sepsis: a local big data-driven, machine learning approach. Acad Emerg Med 2016;23(3):269–78. 10.1111/acem.12876. [PubMed: 26679719]

[23]. Perng J-W, Kao I-H, Kung C-T, Hung S-C, Lai Y-H, Su C-M. Mortality prediction of septic patients in the emergency department based on machine learning. J Clin Med 2019;8(11). 10.3390/jcm8111906.

[24]. Barnaby DP, Fernando SM, Herry CL, Scales NB, Gallagher EJ, Seely AJE. Heart rate variability, clinical and laboratory measures to predict future deterioration in patients presenting with sepsis. Shock 2019;51(4):416–22. 10.1097/SHK.0000000000001192. [PubMed: 29847498]

[25]. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. Sci Data 2016;3:160035. 10.1038/sdata.2016.35. [PubMed: 27219127]

[26]. Soysal E, Wang J, Jiang M, et al. Clamp - a toolkit for efficiently building customized clinical natural language processing pipelines. J Am Med Inf Assoc 2018;25(3):331–6. 10.1093/jamia/ocx132.

[27]. Karakike E, Kyriazopoulou E, Tsangaris I, Routsi C, Vincent J-L, Giamarellos-Bourboulis EJ. The early change of SOFA score as a prognostic marker of 28-day sepsis mortality: analysis through a derivation and a validation cohort. Crit Care 2019;23(1):387. 10.1186/s13054-019-2665-5. [PubMed: 31783881]

[28]. Austrian JS, Jamin CT, Doty GR, Blecker S. Impact of an emergency department electronic sepsis surveillance system on patient mortality and length of stay. J Am Med Inf Assoc 2018;25(5):523–9. 10.1093/jamia/ocx072.

[29]. Rhee C, Wang R, Song Y, et al. Risk adjustment for sepsis mortality to facilitate hospital comparisons using centers for disease control and prevention's adult sepsis event criteria and routine electronic clinical data. Crit Care Explor 2019;1(10):e0049. 10.1097/CCE.0000000000000049. [PubMed: 32166230]

[30]. Canora J, Moreno G, Marco J, et al. Admittances characteristics by sepsis in the Spanish internal medicine services between 2005 and 2015: mortality pattern. Postgrad Med. 2020:1–5. 10.1080/00325481.2020.1718388. January.

[31]. Mankowski RT, Anton SD, Ghita GL, et al. Older sepsis survivors suffer persistent disability burden and poor long-term survival. J Am Geriatr Soc April 2020. 10.1111/jgs.16435.

[32]. Yuan J, Yang X, Yuan Q, Li M, Chen Y, Dong C. [Systematic review of ultrasound-guided fluid resuscitation vs. early goal-directed therapy in patients with septic shock]. Zhonghua Wei Zhong Bing Ji Jiu Yi Xue 2020;32(1):56–61. 10.3760/cma.j.cn121430-20191114-00010. [PubMed: 32148232]

[33]. Douglas IS, Alapat PM, Corl KA, et al. Fluid response evaluation in sepsis hypotension and shock: a randomized clinical trial. Chest 2020;158(4):1431–45. 10.1016/j.chest.2020.04.025. [PubMed: 32353418]

[34]. Chapalain X, Vermeersch V, Egreteau P-Y, et al. Association between fluid overload and SOFA score kinetics in septic shock patients: a retrospective multicenter study. J Intensive Care 2019;7:42. 10.1186/s40560-019-0394-0. [PubMed: 31417678]

[35]. Antal O, Stefănescu E, Mlesnite M, Bălan AM, Hagău N. Initial fluid resuscitation following adjusted body weight dosing in sepsis and septic shock. J Crit Care Med (Targu Mures). 2019;5(4):130–5. 10.2478/jccm-2019-0025. [PubMed: 31915718]

[36]. Andrew MK, Shinde V, Ye L, et al. The importance of frailty in the assessment of influenza vaccine effectiveness against influenza-related hospitalization in elderly people. J Infect Dis 2017;216(4):405–14. 10.1093/infdis/jix282. [PubMed: 28931244]

[37]. Prado-Galbarro FJ, Gaminõ-Arroyo AE, Sánchez-Piedra C, Sánchez-Pájaro A, Sarría-Santamera A. Evaluation of the effect of hospitalization on mortality in patients with heart failure followed in primary care. Arch Cardiol Mex 2019;89(1):118–25. 10.24875/ACME.M19000034. [PubMed: 31702737]

[38]. Chang Y-C, Tung H-J, Huang Y-T, Lu C-T, Ernawaty E, Wu S-Y. Effect of influenza vaccination on mortality and risk of hospitalization in elderly individuals with and without disabilities: a nationwide, population-based cohort study. Vaccines (Basel). 2020;8(1). 10.3390/vaccines8010112.

[39]. Zhou X, Fang H, Xu J, et al. Stress ulcer prophylaxis with proton pump inhibitors or histamine 2 receptor antagonists in critically ill adults - a meta-analysis of randomized controlled trials with trial sequential analysis. BMC Gastroenterol 2019;19(1):193. 10.1186/s12876-019-1105-y. [PubMed: 31752703]

[40]. Barletta JF. Histamine-2-receptor antagonist administration and gastrointestinal bleeding when used for stress ulcer prophylaxis in patients with severe sepsis or septic shock. Ann Pharmacother 2014;48(10):1276–81. 10.1177/1060028014540513. [PubMed: 24982312]

[41]. Rhee C, Jentzsch MS, Kadri SS, et al. Variation in identifying sepsis and organ dysfunction using administrative versus electronic clinical data and impact on hospital outcome comparisons. Crit Care Med 2019;47(4):493–500. 10.1097/CCM.0000000000003554. [PubMed: 30431493]

[42]. Brown RM, Wang L, Coston TD, et al. Balanced crystalloids versus saline in sepsis. A secondary analysis of the smart clinical trial. Am J Respir Crit Care Med 2019;200(12):1487–95. 10.1164/rccm.201903-0557OC. [PubMed: 31454263]

[43]. Kempker JA, Kramer MR, Waller LA, Martin GS. Risk factors for septicemia deaths and disparities in a longitudinal US cohort. Open Forum Infect Dis 2018;5(12):ofy305. 10.1093/ofid/ofy305. [PubMed: 30568980]

[44]. Capan M, Hoover S, Ivy JS, Miller KE, Arnold R. S.E.P.S.I.S (Sepsis Early Prediction Support Implementation System) Collaborative. Not all organ dysfunctions are created equal - prevalence and mortality in sepsis. J Crit Care 2018;48:257–62. 10.1016/j.jcrc.2018.08.021. [PubMed: 30245367]

SOFA

	0	+1	+2	+3	+4
Respiratory (PaO2/FiO2)	>400	399 - 300	299 - 200	199 - 100 and mechanically ventilated	< 100 and mechanically ventilated
Cardiovascular	MAP > 69	MAP < 69	Dopamine < 5 mcg/kg/min OR Dobutamine any dose	Dopamine 5 - 15 mcg/kg/min OR Epi / Norepi < 0.1 mcg/kg/min	Dopamine > 15 mcg/kg/min OR Epi / Norepi > 0.1 mcg/kg/min
Nervous (CGS)	15	14 - 13	12 - 10	9 - 6	< 6
Hepatic (Bili)	< 1.2	1.2 - 1.9	2.0 - 5.9	6.0 - 11.9	> 12
Coagulation (Plts)	> 149	149 - 100	99 - 50	49 - 20	< 20
Renal (Cr)	< 1.2	1.2 - 1.9	2.0 - 3.4	3.5 - 4.9	> 5.0

qSOFA

	Value	Score
Glascow Coma Scale	≤ 13	+1
Respiratory Rate	≥ 22	+1
Systolic Blood Pressure	≤ 100	+1

Fig. 1.
SOFA and qSOFA.

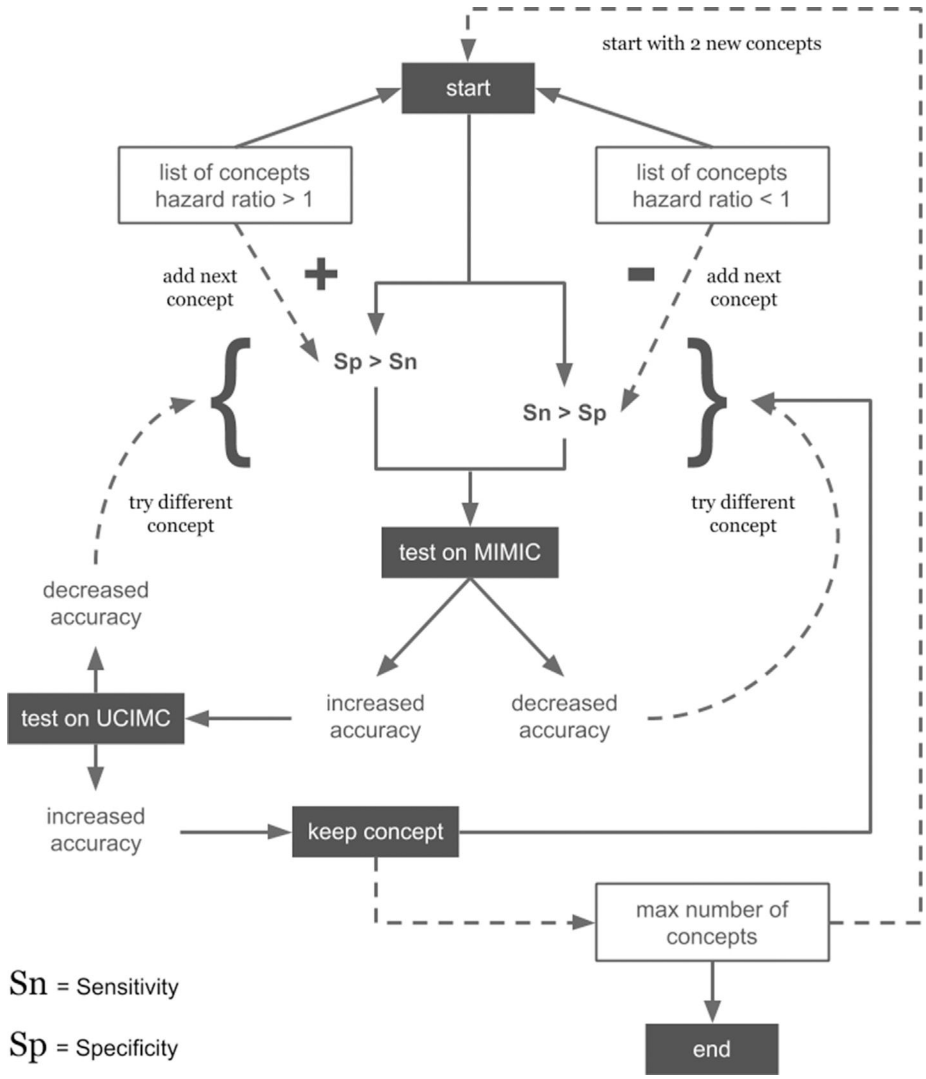

Fig. 2.

Decision tree.

Note: Concept lists were ordered by strength of hazard ratio. One concept with hazard ratio >1 and one concept with hazard ratio <1 were used at each start. Initial seeded concepts in successful end models were not reused.

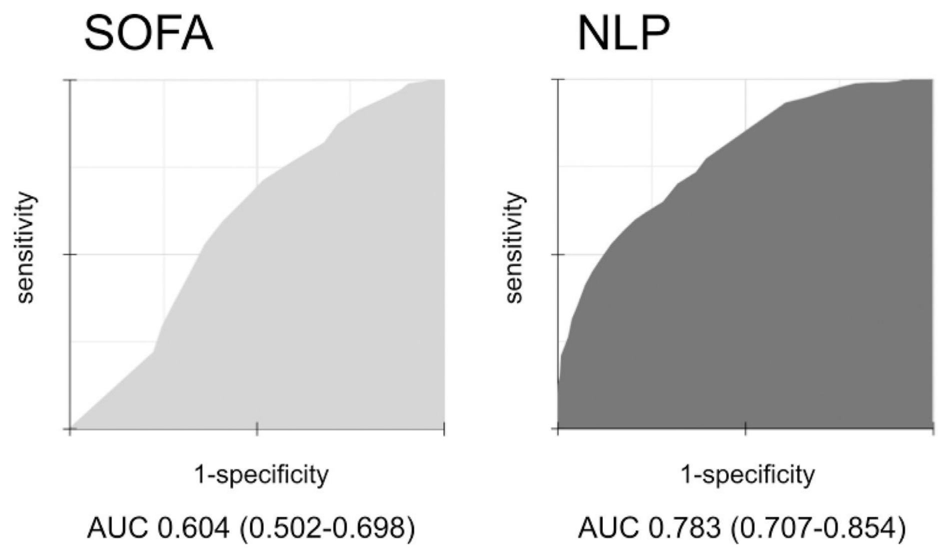

Fig. 3.
MIMIC III ROC curve - NLP and SOFA.

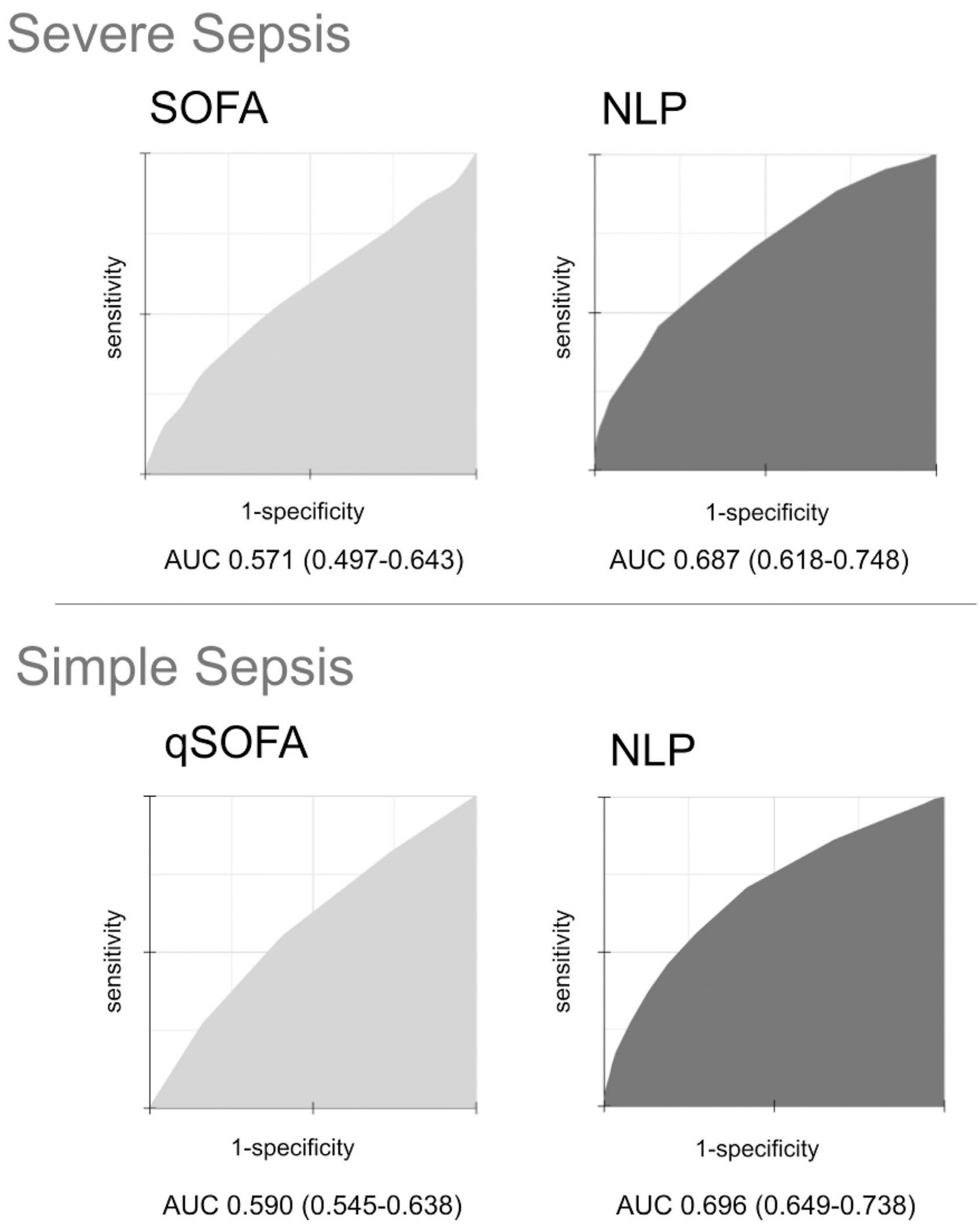

Fig. 4.
UCIMC ROC Curve - NLP vs SOFA / qSOFA for Severe and Simple Sepsis

Table 1

Patient demographics.

	MIMIC III	**UCIMC**	***Severe sepsis***	***Simple sepsis***
demographics				
Admissions (n)	1052	7117	2174	4943
males	56.8%	52.4%	52.9%	52.1%
females	43.2%	47.6%	47.1%	47.9%
age (y/o & std dev)	64.2 (17.3)	58.9 (18.4)	60.3 (18.0)	58.2 (18.6)
mortality	26.9%	27.6%	30.2%	26.5%
caucasian	76.3%	40.1%	40.4%	40.0%
latino	3.6%	33.9%	34.1%	33.7%
asian	3.0%	13.9%	13.5%	14.1%
african american	10.7%	3.1%	3.3%	3.1%
other	6.4%	9.0%	8.7%	9.1%
comorbidities				
hypertension	31.8%	56.6%	58.9%	55.6%
diabetes	20.1%	32.8%	32.9%	32.7%
hyperlipidemia	17.5%	17.0%	17.1%	17.0%
coronary artery disease	16.0%	15.1%	16.3%	14.6%
atrial fibrillation	27.3%	14.6%	15.8%	14.0%
end stage renal disease	9.2%	13.7%	13.0%	13.9%

Table 2

Concepts from MIMIC III.

CUI	Freq	HR	Concept	CUI	Freq	HR	Concept
C3815196	0.01	0.306	Sx	C0476273	0.04	1.978	Respiratory Distress
C0331794	0.06	0.674	Boots	C0036974	0.06	1.673	Shock
C0202691	0.04	1.640	Head CT	C0149758	0.03	0.656	Poor Dentition
C0428167	0.07	3.213	FiO2	C1261075	0.01	0.700	RLL
C0452837	0.02	3.095	Drip	C0175723	0.07	1.576	Band
C0043352	0.02	0.468	Dry Mouth	C0170457	0.02	0.688	PIVs
C0026267	0.01	0.555	Mitral Valve Prolapse	C0036980	0.03	1.980	Cardiogenic Shock
C0521009	0.01	0.648	Bacterial	C0019593	0.03	0.705	H2 Blocker
C0147728	0.07	0.525	UF	C0051298	0.01	0.565	N/V
C0176751	0.01	0.528	PEG	C0014792	0.02	0.681	RBC
C0012739	0.02	3.155	DIC	C0808387	0.07	3.198	Ventilator Mode
C0003862	0.03	0.421	Arthralgias	C4054315	0.02	0.492	Organomegaly
C0424790	0.01	0.358	Rigors	C0005367	0.02	2.058	HCO3
C0546817	0.04	1.749	Volume Overload	C2071425	0.01	0.665	Crackles at Bases
C0024202	0.01	0.676	Lymph	C0030305	0.01	0.652	Pancreatitis
C0876139	0.04	1.635	Pantoprazole	C0525032	0.06	1.614	INR
C0030554	0.02	0.390	Numbness/Tingling	C0025285	0.01	0.556	Meningitis
C0237284	0.03	3.100	Unresponsive	C0268494	0.03	1.796	ATN
C0018802	0.04	1.764	CHF	C0006826	0.02	1.960	Cancer
C0034161	0.01	0.517	Pus	C0443203	0.02	0.528	Distant
C0242339	0.01	0.394	Dyslipidemia	C0235195	0.07	1.744	Sedated
C0038661	0.01	0.386	Suicidal	C1522438	0.04	0.590	SC
C0006736	0.01	0.541	Stones	C0802927	0.06	3.137	Vt
C1707737	0.03	0.652	Diaphoretic	C1265292	0.03	0.669	MRSA
C0004238	0.04	1.830	Atrial Fibrillation	C0087153	0.03	2.535	Ventilation
C1513302	0.01	0.431	Mild	C1321095	0.05	2.017	Gtt
C0237795	0.04	2.302	Pressors	C0152404	0.04	0.629	VS
C0149801	0.02	0.386	Urosepsis	C0040034	0.02	1.997	Thrombocytopenia

CUI	Freq	HR	Concept	CUI	Freq	HR	Concept
C1260880	0.02	0.506	Rhinorrhea	C0228341	0.01	0.676	LP
C0021400	0.02	0.734	Influenza	C3897988	0.07	2.897	PaO2/FiO2
C0001122	0.03	2.616	Acidosis	C0013428	0.05	0.437	Dysuria
C0016006	0.04	2.939	Fibrinogen	C0020179	0.04	2.107	HD
C0231174	0.04	2.310	Failure	C0008031	0.07	0.533	Chest Pain
C1511726	0.01	0.566	Data	C0021925	0.09	1.734	Intubation
C2219848	0.01	0.421	Daytime Somnolence	C0035078	0.05	2.109	Renal Failure
C0178733	0.01	0.571	Loud	C0014591	0.02	0.667	Epistaxis
C0085619	0.03	0.583	Orthopnea	C0015230	0.08	0.647	Rash
C0001924	0.08	1.604	Albumin	C0234246	0.02	0.614	Rebound Tenderness
C0030547	0.01	0.386	Parental Nutrition	C1145670	0.05	2.616	Respiratory Failure
C0021403	0.01	0.451	Influenza Vaccine	C0020461	0.03	2.079	Hyperkalemia

Note: CUI = SNOMED CT concept, Freq = Observed Frequency, HR = Hazard Ratio.