

ORIGINAL RESEARCH

The Practice of Emergency Medicine

Prediction of emergency department resource requirements during triage: An application of current natural language processing techniques

Nicholas W. Sterling MD, PhD¹ | Felix Brann BSc² | Rachel E. Patzer PhD, MPH^{3,4} |
Mengyu Di MSPH⁵ | Megan Koebbe RN¹ | Madalyn Burke RN¹ |
Justin D. Schrager MD, MPH¹

¹ Department of Emergency Medicine, Emory University School of Medicine, Atlanta, Georgia, USA

² Vital Software Inc., Auckland, New Zealand

³ Department of Surgery, Emory University School of Medicine, Atlanta, Georgia, USA

⁴ Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA

⁵ Department of Biostatistics and Bioinformatics, Emory University, Atlanta, Georgia, USA

Correspondence

Justin D. Schrager, MD, MPH, Department of Emergency Medicine, Emory University School of Medicine, 531 Asbury Circle, Annex Building N340, Atlanta, GA 30322, USA.
Email: jschrag@emory.edu

Funding and support: By *JACEP Open* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article as per ICMJE conflict of interest guidelines (see www.icmje.org).

Nicholas W. Sterling and Felix Brann should be considered joint first author.

Abstract

Objective: Accurate triage in the emergency department (ED) is critical for medical safety and operational efficiency. We aimed to predict the number of future required ED resources, as defined by the Emergency Severity Index (ESI) triage protocol, using natural language processing of nursing triage notes.

Methods: We constructed a retrospective cohort of all 265,572 consecutive ED encounters from 2015 to 2016 from 3 separate clinically heterogeneous academically affiliated EDs. We excluded encounters missing relevant information, leaving 226,317 encounters. We calculated the number of resources used by patients in the ED retrospectively and based outcome categories on criteria defined in the ESI algorithm: 0 (30,604 encounters), 1 (49,315 encounters), and 2 or more (146,398 encounters). A neural network model was trained on a training subset to predict the number of resources using triage notes and clinical variables at triage. Model performance was evaluated using the test subset and was compared with human ratings.

Results: Overall model accuracy and macro F1 score for number of resources were 66.5% and 0.601, respectively. The model had similar macro F1 (0.589 vs 0.592) and overall accuracy (65.9% vs 69.0%) compared to human raters. Model predictions had slightly higher F1 scores and accuracy for 0 resources and were less accurate for 2 or more resources.

Conclusions: Machine learning of nursing triage notes, combined with clinical data available at ED presentation, can be used to predict the number of required future ED

Supervising Editor: Christian A. Tomaszewski, MD, MBA.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *JACEP Open* published by Wiley Periodicals LLC on behalf of the American College of Emergency Physicians.

resources. These findings suggest that machine learning may be a valuable adjunct tool in the initial triage of ED patients.

KEYWORDS

emergency department, machine learning, resources, triage, natural language processing

1 | BACKGROUND

The number of emergency department (ED) visits in the United States is continuing to rise. Given this influx of patients into the ED, crowding and overall ED efficiency have become focal points of discussion. Indeed, crowding in the ED has been associated with negative outcomes, including increased morbidity and mortality, longer wait times, increased length of stay, increased hospital expenses, and poorer patient perceptions of care.¹⁻⁴ Accordingly, a substantial emphasis has been placed on optimizing ED efficiency by appropriately allocating departmental resources and patient distributions. To predict and potentially prevent downstream bottlenecks in the ED, past studies have demonstrated that it is possible to use the information available at the time of patient triage to make predictions regarding the severity and final ED disposition of patients.⁵⁻⁹

Triage documentation is typically written by a nurse who interviews and examines the patient briefly in a small examination area adjacent to the ED waiting room to determine severity and likely number of resources required. The nurse records various clinical data points (eg, vital signs, weight, comorbidities). The nurse then typically documents a short 1 to 3 sentence triage note summarizing the patient's reason for presentation and then assigns a triage acuity level from 1 to 5 based on the Emergency Severity Index (ESI) algorithm.¹⁰ The ESI is used to stratify patients into triage groups from 1 (most urgent) to 5 (least urgent) based on acuity of illness and resource needs and has been validated multiple times in a variety of clinical settings. It is the most common triage algorithm in the United States.¹⁰ Resources are broadly grouped into 0, 1, or 2 or more. Examples of an ED resource, according to ESI, include a laboratory test (blood or urine) or a laceration repair. Some typical ED activities, such as providing oral medications or placing a splint, are not considered resources. Without training, this list is not necessarily intuitive.

1.1 | Importance

Nursing triage documentation is the first available medical assessment in the electronic medical record and is typically written by experienced ED nurses and may contain higher level information (ie, sentiment) regarding the patient's condition. We recently demonstrated that natural language processing (NLP) coupled with machine learning using nursing triage notes can be used to predict discrete outcomes such as hospital admission.¹¹ Machine learning excels at predicting quantifiable metrics. An advantage of using information at the time of triage is that it is present in the electronic medical record hours before physician documentation is available.

The process of assigning an ESI score requires the triage nurse to forecast the number of resources that a patient will require during his or her ED stay. A limitation of the ESI algorithm is that the information regarding the estimated number of required resources is usually lost after it is combined with the clinical severity of illness to generate the final ESI. Furthermore, the process of estimating number of resources is prone to human error and may vary from one nurse to another. Furthermore, ESI is a proprietary system owned by the Emergency Nurses Association, for which training costs money and takes time. These are potential hardships and disadvantages for EDs that are underresourced. Accordingly, there may be a role for machine learning to assist in the triage of patients in the ED, especially when it comes to the standardization of triage decisionmaking.

1.2 | Goals of this investigation

We hypothesized that NLP coupled with deep learning of triage notes and other clinical data would be able to produce a model that is predictive of the number of resources required for a given patient in the ED. To our knowledge, the current study is the first to use NLP and machine learning of clinical data to predict the number of resources required for patients in the ED.

2 | METHODS

2.1 | Study design and setting

A retrospective cohort was constructed using electronic health record data from all 260,842 consecutive ED encounters from 3 separate clinically heterogeneous academically affiliated EDs from 2015 to 2016. Hospital A is a community hospital in an urban setting with a patient volume of $\approx 75,000$ ED visits per year. Hospital B is a community hospital in a suburban setting with a volume of $\approx 27,000$ visits per year. Hospital C is a quaternary care academic medical setting in a major metropolitan area with an ED patient volume of $\approx 40,000$ visits per year. The research study protocol and procedures were reviewed and approved by the institutional review board.

2.2 | Selection of participants

We excluded encounters based on frequent use (defined as ≥ 20 visits during the study years, to minimize training bias), disposition other

than admission or discharge, and arrival time unavailable. The final cohort of interest included 226,317 unique clinical encounters from 144,421 unique patients.

2.3 | Interventions

The study was retrospective in nature. Accordingly, no interventions were performed on the study subjects.

2.4 | Measurements

For this analysis, we included the nursing triage notes in word token format and clinical variables that were available at the time of triage, including demographics, age, sex, vital signs, number of ED visits within the past 12 months, number of ED visits within past 7 days, pain score, disposition from most recent encounter, arrival mode, whether the patient had altered mental status, one-hot encoded chief complaint, past medical problems, most recent medical problem, and home medications.

2.5 | Outcomes

The outcome of interest was model performance, which was measured by accuracy (percent correct predictions over all predictions) and F1 score.

2.6 | Analysis

To use free-text natural language data as input for training and testing the model, it was first necessary to convert each triage note to a fixed-length vector to serve as numerical representations of the free text in the notes. Briefly, the pretrained word embeddings were used to transform the free text from the triage notes into word vectors of 200 length. First, common stop words were removed. Second, each word was tokenized to correspond to a row in the pretrained embedding matrix. As opposed to randomly initializing word vectors, we used pretrained embeddings published by,¹² which were generated from PubMed abstracts and a large text corpus from Wikipedia using the *word2vec* program to facilitate the learning process.¹³

Long short-term memory models have been used extensively in deep learning of natural language data and are variants of recurrent neural networks that use memory gates to update the model's internal representation using a single input word at a time.¹⁴ Tokenized triage notes were used as an input to an embedding layer with weights from the embedding matrix in the aforementioned step. This embedding layer was subsequently used as the input to a long short-term memory layer with output dimensionality of 128. This long short-term memory output was concatenated with 9,650 clinical variables that were described previously.

We defined 3 possible categories for number of resources: 0 (30,604 encounters), 1 (49,315 encounters), and 2 or more (146,398

The Bottom Line

Resource use in emergency departments can be predicted by machine learning using data available at the time of triage.

encounters). The data set was divided into training, validation, and test subsets (70%, 15%, and 15%, respectively) using randomization for row selection. The training set was resampled for equal representation of each number of resources category to avoid training bias. The model was trained to optimize categorical cross entropy over 100 epochs using a learning rate of 0.001, L2 regularization of 0.05, dropout of 0.5, and batch size of 4,096. Model training and validation were performed using Keras with TensorFlow 2.0 via Amazon Web Services Sagemaker.¹⁵ Performance characteristics were characterized using the predicted versus actual values of number of resources category of all encounters in the validation data set.

To provide a real-world comparison of predictive performance, 2 experienced ED nurses were blinded to the number of resources and were tasked with predicting number of resources (by categories of 0, 1, or 2 or more). A total of 1,000 encounters were randomly selected without replacement from the test data set and were rated by both nurses. The nurses were given typically available clinical information at the time of triage (Table 1) as well as comorbidities, the ED nursing triage note, past medical history, whether the patient had altered mental status, ED visits within the past 7 days, ED visits within the past 12 months, arrival mode, and most recent problem. These were provided in an attempt to approximate the data they would have at hand in an ideal scenario to assist with triage. Model predictions were obtained for each of these encounters and were compared against human rater predictions. Interrater reliability for human raters was calculated using the κ statistic.

3 | RESULTS

3.1 | Characteristics of study subjects

Overall, there were 265,572 consecutive ED encounters from 2015 to 2016. After the exclusion of 39,255 encounters (14.8%) based on data completeness (triage note and/or vital signs available), frequent use (defined as ≥ 20 visits during the study years to minimize training bias), disposition other than admission or discharge, and arrival time unavailable we conducted our analysis on the 226,317 remaining encounters. There were 30,604 (13.5%) encounters resulting in 0 resources; 49,315 (21.8%) encounters resulting in 1 resource; and 146,398 (64.7%) encounters resulting in 2 or more resources. There were 144,421 unique patients in the cohort. The average length of the triage notes was 174.7 (SD = 92.9) characters. Demographic and clinical characteristics of the patient sample are described in Table 1.

TABLE 1 Demographic and clinical characteristics of patients across encounters (2015–2016)

	Total	Hospital A	Hospital B	Hospital C
N (%)	226,522	100,188 (44.23)	50,130 (22.13)	76,204 (33.64)
Age, N (%)				
<18 y	8,404 (3.7)	1,871 (22.26)	5,818 (69.23)	715 (8.51)
18–24 y	22,827 (10.1)	11,434 (50.09)	3,855 (16.89)	7,538 (33.02)
25–44 y	73,997 (32.7)	36,272 (49.02)	14,196 (19.18)	23,529 (31.8)
45–64 y	68,688 (30.3)	30,531 (44.45)	14,527 (21.15)	23,630 (34.4)
65–74 y	26,422 (11.7)	10,667 (40.37)	5,394 (20.41)	10,361 (39.21)
≥75 y	26,184 (11.6)	9,413 (35.95)	6,340 (24.21)	10,431 (39.84)
Sex, N (%)				
Male	91,998 (40.6)	41,171 (44.75)	20,426 (22.2)	30,401 (33.05)
Female	134,523 (59.4)	59,017 (43.87)	29,703 (22.08)	45,803 (34.05)
Race, N (%)				
White	78,799 (35.8)	15,314 (19.43)	32,033 (40.65)	31,452 (39.91)
Black	132,781 (60.3)	82,097 (61.83)	10,756 (8.1)	39,928 (30.07)
Other	8,761 (4.0)	1,241 (14.17)	4,635 (52.9)	2,885 (32.93)
Ethnicity, N (%)				
Non-Hispanic or Latino	210,774 (97.0)	95,398 (45.26)	44,593 (21.16)	70,783 (33.58)
Hispanic or Latino	6,417 (3.0)	1,725 (26.88)	2,508 (39.08)	2,184 (34.03)
Pain score, N (%)				
0–2	67,304 (30.9)	27,586 (40.99)	13,315 (19.78)	26,403 (39.23)
3–6	47,048 (21.6)	19,990 (42.49)	11,336 (24.09)	15,722 (33.42)
7–10	103,567 (47.5)	50,305 (48.57)	19,520 (18.85)	33,742 (32.58)
Heart rate, N (%)				
<60	9,262 (4.3)	3,495 (37.73)	2,096 (22.63)	3,671 (39.64)
60–100	169,554 (77.8)	77,021 (45.43)	34,783 (20.51)	57,750 (34.06)
>100	39,031 (17.9)	17,342 (44.43)	7,284 (18.66)	14,405 (36.91)
Temperature, N (%)				
<36°C	15,862 (7.3)	12,406 (78.21)	1,159 (7.31)	2,297 (14.48)
36°C–38°C	196,727 (90.3)	82,840 (42.11)	42,265 (21.48)	71,622 (36.41)
>38°C	5,328 (2.4)	2,633 (49.42)	747 (14.02)	1,948 (36.56)
DBP, N (%)				
<60	17,975 (8.3)	7,230 (40.22)	5,009 (27.87)	5,736 (31.91)
60–80	104,872 (48.2)	48,360 (46.11)	21,667 (20.66)	34,845 (33.23)
>80	94,785 (43.6)	42,255 (44.58)	17,481 (18.44)	35,049 (36.98)
SBP, N (%)				
<80	789 (0.4)	351 (44.49)	95 (12.04)	343 (43.47)
80–120	48,856 (22.4)	21,776 (44.57)	9,469 (19.38)	17,611 (36.05)
>120	167,991 (77.2)	75,718 (45.07)	34,592 (20.59)	57,681 (34.34)
SPO ₂ , N (%)				
≤90%	2,887 (1.3)	1,061 (36.75)	776 (26.88)	1,050 (36.37)
>90%	214,927 (98.7)	96,801 (45.04)	43,384 (20.19)	74,742 (34.78)
Respiratory rate, mean ± SD	17.91 ± 3.17	17.79 ± 2.83	18.04 ± 3.2	18 ± 3.54

DBP, diastolic blood pressure; SD, standard deviation; SBP, systolic blood pressure; SPO₂, oxygen saturation.

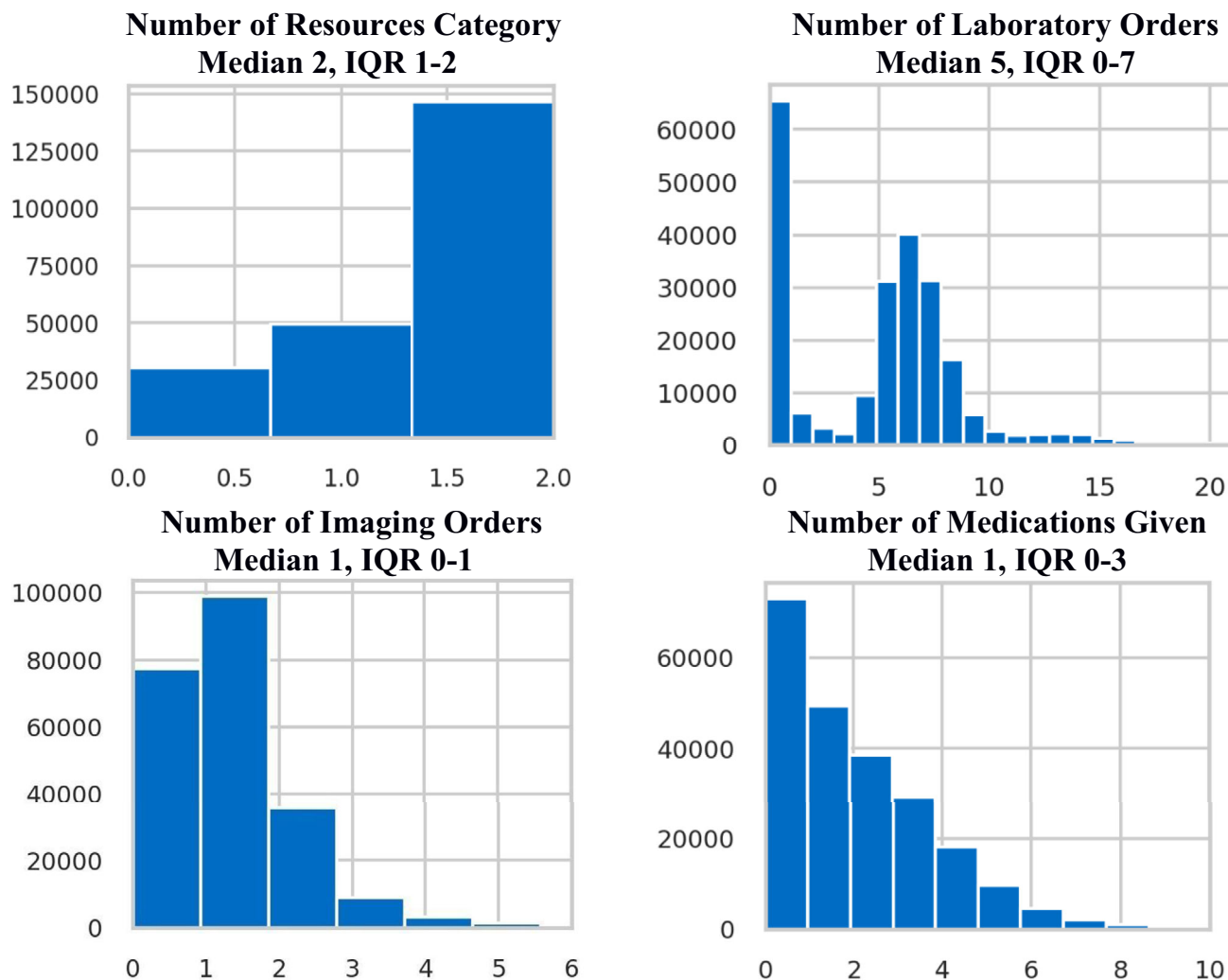


FIGURE 1 Distributions of “number of resources” category, laboratory orders, imaging orders, and medications administered during emergency department encounter. IQR, interquartile range

Frequencies of number of resources category, number of imaging orders, number of laboratory orders, and number of medications given are illustrated in Figure 1.

3.2 | Main results

Metrics were computed to gauge overall model performance, and a confusion matrix was constructed to determine accurate classification versus misclassification of model predictions (Table 2). For the model performance using the test data set, macro and weighted F1 scores were 0.601 and 0.684, respectively. Accuracy was 66.6% overall, 71.7% (987/1,377) for 0 resources, 52.7% (1,142/2,166) for 1 resource, and 70.1% (4,527/6,457) for 2 or more resources (Figure 2A). For the model performance using the validation data set, macro and weighted F1 scores were 0.645 and 0.645, respectively. Accuracy was 64.5% overall, 70.6% (2,408/3,410) for 0 resources, 54.4% (1,799/3,308) for 1 resource, and 68.4% (2,246/3,282) for 2 or more resources (Figure 2B).

The accuracies of the first and second raters for the prediction of number of resources category were 67.8% and 70.1%, respectively. The combined accuracy of both raters was 69.0%, and agreement (κ statistic) was 0.588 (Figure 3). The model had similar macro F1 (0.589 vs 0.592) and overall accuracy (65.9% vs 69.0%) compared to human raters. The model had slightly higher macro F1 score (0.515 vs 0.484) and higher accuracy (69.0% vs 45.2%) of predicting 0 number of resources when compared to human raters. However, the model was less accurate (71.2% vs 80.3%) and had slightly lower F1 score (0.791 vs 0.817) for the prediction of 2 or more resources. For the prediction of 1 resource, the model had similar accuracy (50.0% vs 51.3%) and F1 score (0.461 vs 0.473) compared to human raters (Table 2).

3.3 | Limitations

The current study had several limitations. First, the design of this study was retrospective in nature. Although we excluded encounters that resulted in dispositions other than admission or discharge, it is

TABLE 2 F1 scores across each “number of resources” category based on predictions of the trained model using all test set data, the trained model using all validation data set encounters, and the trained model and human raters using 1,000 encounters randomly selected from the test set

Model predictions (all test set encounters)				
Macro F1 = 0.601, overall accuracy = 66.6%				
	Precision	Recall	Accuracies	F1
0	0.447	0.717	0.717	0.551
1	0.419	0.527	0.527	0.467
2 or more	0.894	0.701	0.701	0.786
Model predictions (all validation set encounters)				
Macro F1 = 0.645, overall accuracy = 64.5%				
	Precision	Recall	Accuracies	F1
0	0.664	0.706	0.706	0.684
1	0.554	0.544	0.544	0.549
2 or more	0.719	0.684	0.684	0.701
Model predictions (N = 1,000, for comparison with nurse predictions)				
Macro F1 = 0.589, overall accuracy = 65.9%				
	Precision	Recall	Accuracies	F1
0	0.410	0.690	0.690	0.515
1	0.428	0.500	0.500	0.461
2 or more	0.888	0.712	0.712	0.791
Nurse predictions (N = 1,000, for comparison with model predictions)				
Macro F1 = 0.592, overall accuracy = 69.0%				
	Precision	Recall	Accuracies	F1
0	0.521	0.452	0.452	0.484
1	0.440	0.513	0.513	0.473
2 or more	0.833	0.803	0.803	0.817

possible that some encounters may have been incomplete, making it difficult to determine the actual number of resources used. Second, although data available at the time of triage may carry particular relevance to determine the number of resources required for patients in the ED, the general clinical impression of triage nurses may not be carried through to the triage note in some instances, especially when the ED is busy. Third, although separate training and testing data sets were used to train and test the model, respectively, prospective studies are needed to evaluate the effects of such a tool on clinical practice. Fourth, it is possible that embedded clinical decision support tools could be used to increase human accuracy in predicting number of resources. Although this is beyond the scope of the present study, future investigations may address this as a possibility. Fifth, although the overall model performance may be useful to approximate the likelihood of obtaining accurate model predictions, there was considerable variation of model performance between low and high resources subcategories. This should be considered when interpreting the study results. Finally, although the model was able to predict the number of resources, as defined by the ESI algorithm, the actual resources used were not specified. Future studies may focus on predicting the precise resources that are likely to be used in the ED.

4 | DISCUSSION

In this study, we used triage notes from 226,317 patient encounters across 3 large metropolitan EDs to predict the number of resources (0, 1, or 2 or more) required for patients during their ED visits. A particular advantage of using triage notes, vital signs, and known clinical history is that this information is typically available at the time of triage. Accordingly, these data are able to be used early in the encounter to make predictions about likely events during the ED visit. In the current study, we demonstrated that the free text from nursing triage notes can be converted into numerical representations of natural language and may then be combined with other quantitative and categorical clinical data to predict the number of resources required by a given patient during an ED visit. These findings suggest that free-text data contained in patient charts, such as the triage notes used in this study, carry information that should be considered and may be useful when developing predictive models of an anticipated course in the ED.

To our knowledge, the current study is the first to use free text from nursing triage notes with other quantitative clinical data available at the time of triage to predict the number of resources, defined according to ESI criteria,¹⁰ required for patients during their ED visits. A strength

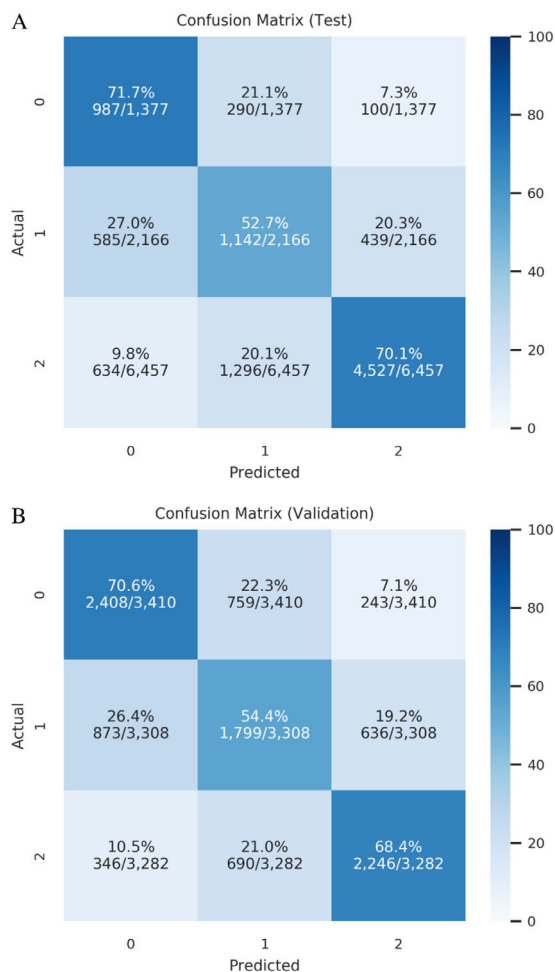


FIGURE 2 (A) Test data set confusion matrix for number of resources category prediction using natural language processing of nursing triage notes and current and past clinical data. (B) Validation data set confusion matrix for number of resources category prediction using natural language processing of nursing triage notes and current and past clinical data

of this study was the use of clinically diverse samples from 3 different hospital settings. Moreover, an advantage of the current study is that the model performances were compared to those of human raters. A notable finding of this study was that the overall predictive performance of the model was comparable to that of trained ED nurses. In addition, model prediction performance was strongest for encounters requiring 0 resources, which are likely to represent less medically complex visits. Although our results may suggest that machine learning outperforms human raters for encounters requiring fewer resources, further studies are needed to replicate this finding.

Previous research has suggested that ESI may correlate inversely with number of resources category. For example, a patient having an ESI of 1 is likely to have greater resource requirements than another having an ESI of 5.¹⁶ In the ED setting, triage nurses are required to make predictions about disease severity and the number of resources needed during the ED visit to compute the final ESI. This approach has important limitations. Notably, although the interrater reliability of the ESI has been shown to be acceptable,¹⁷ it is reasonable to sus-

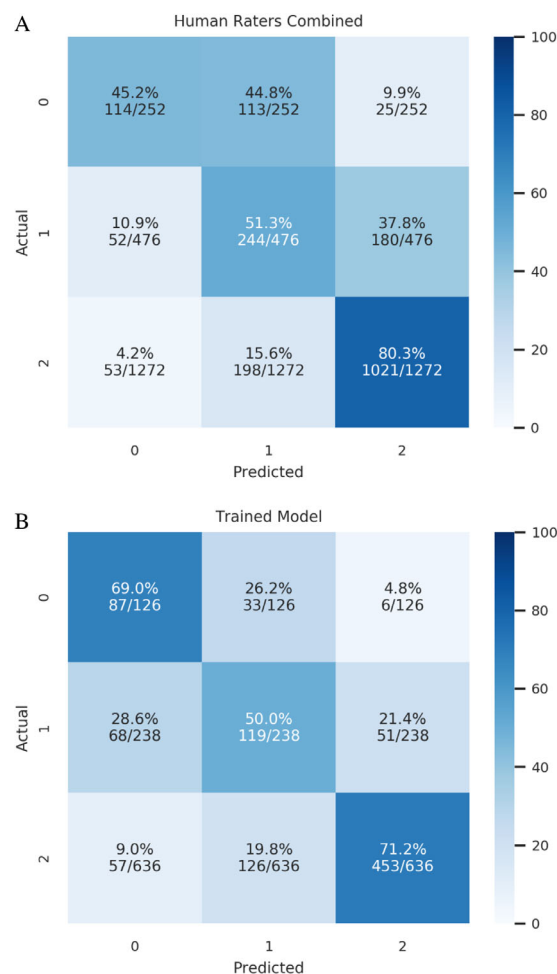


FIGURE 3 (A) Confusion matrix for “number of resources” category using predictions of 2 experienced emergency department nurses over N = 1,000 patient encounters selected randomly from the test data set. (B) Confusion matrix for “number of resources” category using predictions of the trained model over these same N = 1,000 patient encounters

pect that reliability may be lower among less experienced clinicians and in EDs where there are fewer resources to provide training regarding the proprietary ESI algorithm. In addition, although recent patient history is often available at the time of triage, it is not always feasible to access the entirety of a patient’s clinical history. Moreover, this process is prone to human error and important risk factors may be overlooked. The process of ED triage, in a busy ED, is a balance between expediency, appropriateness, and completeness. Combined with human oversight, machine learning may have an important role facilitating accurate, consistent, and expeditious triage of patients in the ED.

The fact that it is possible to generate reliable and meaningful representations of free-text data has far-reaching implications for health research. Indeed, written data within patient charts was previously considered to be qualitative in nature, and it largely still is considered as such. This paradigm is challenged by NLP methodologies, which have demonstrated consistent improvement in the way that it is feasible to convert free text to quantitative data in a formulaic and repeatable process.

The context of NLP and modeltraining may influence the final numerical representations of natural language. For example, triage notes contain specific language and symbols that are relatively uncommon in everyday English language. In addition, many words and phrases that are used in triage notes carry different meanings in the healthcare setting (ie, “complains of,” “endorses,” and “stable”). The current study used word embeddings generated from a large collection of PubMed abstracts and Wikipedia articles to capture language in medical and general contexts.⁹

Although the current study is limited to the use of machine learning and NLP for prediction at an individual patient level, it is worth noting that these technologies may have particular relevance to the operation of health systems. For example, modern health system design may incorporate concepts from resilience engineering,¹⁸ wherein systems are designed to be able to proactively adjust to anticipate a stressor, absorb the stressor, and return to a baseline state of functioning. Machinelearning tools, in such systems, might be used to anticipate impending stressors and assist in the allocation of resources.

In this study, we used machine learning to predict the number of resources needed for patients during their ED visits via NLP of nursing triage notes in combination with other current and past clinical data. The results suggest that the number of resources is predictable via machine learning using these data, which are typically available at the time of triage. This has important clinical implications, as machine learning may hold potential to augment triage capabilities for triage staff, inform allocation of ED clinical and staff resources, and improve flow in the ED. External prospective studies will be required to validate these findings and further explore the role of machine learning in the ED triage process.

CONFLICT OF INTEREST

J.S. is the cofounder of Vital Software, Inc, a company engaged in developing artificial intelligence clinical decision support products for the emergency department. The other authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

NS, FB, RP, and JS were involved in designing the paper. NS, FB, MD, MK, MB, and JS helped in the execution of the work. All the authors contributed in writing the paper.

REFERENCES

- Chalfin DB, Trzeciak S, Likourezos A, Baumann BM, Dellinger RP. Impact of delayed transfer of critically ill patients from the emergency department to the intensive care unit. *Crit Care Med*. 2007;35(6):1477-1483.
- Rabin E, Kocher K, McClelland M, et al. Solutions to emergency department “boarding” and crowding are underused and may need to be legislated. *Health Aff*. 2012;31(8):1757-1766.
- Forero R, McCarthy S, Hillman K. Access block and emergency department overcrowding. *Crit Care*. 2011;15(2):216.
- Sun R, Karaca Z, Wong HS. Trends in Hospital Emergency Department Visits by Age and Payer, 2006–2015: Statistical Brief #238. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs. Rockville (MD): Agency for Healthcare Research and Quality (US); 2018.
- Camargo CA, Hasegawa K, Faridi MK, Freishtat RJ, Goto T. Machine learning-based prediction of clinical outcomes for children during emergency department triage. *JAMA Netw Open*. 2019;2(1):e186937.
- Barak-Corren Y, Israelit SH, Reis BY. Progressive prediction of hospitalisation in the emergency department: uncovering hidden patterns to improve patient flow. *Emerg Med J*. 2017;34(5):308-314.
- Sun Y, Heng BH, Tay SY, Seow E. Predicting hospital admissions at emergency department triage using routine administrative data. *Acad Emerg Med*. 2011;18(8):844-850.
- Dugas AF, Kirsch TD, Toerper M, et al. An electronic emergency triage system to improve patient distribution by critical outcomes. *J Emerg Med*. 2016;50(6):910-918.
- Levin S, Toerper M, Hamrock E, et al. Machine-learning-based electronic triage more accurately differentiates patients with respect to clinical outcomes compared with the emergency severity index. *Ann Emerg Med*. 2018;71(5):565-574.e2.
- Gilboy N, Tanabe P, Travers D, Rosenau AM. *Emergency Severity Index (ESI): A Triage Tool for Emergency Department Care, Version 4*. Rockville, MD: Implementation Handbook. 2012 Edition; 2012.
- Sterling NW, Patzer RE, Di M, Schrage JD. Prediction of emergency department patient disposition based on natural language processing of triage notes. *Int J Med Inform*. 2019;129:184-188.
- Pyysalo S, Ginter F, Moen H, Salakoski T, Ananiadou S. Distributional semantics resources for biomedical text processing. *Proceedings of LBM*. 2013; 39-4.
- Mikolov T, Sutskever I, Chen K, Corrado GJD. Distributed representations of words and phrases and their compositionality. *Lit Viajes Y Traducción*. 2007. arXiv:1310.4546.
- Gers FA, Schmidhuber J, Cummins F. Learning to forget: continual prediction with LSTM. *Neural Comput*. 2000;12(10):850-855. <https://doi.org/10.1162/089976600300015015>
- Eclipse DeepLearning4j Development Team. DeepLearning4j: Open-source distributed deep learning for the JVM. <http://deeplearning4j.org>. Published 2018.
- Tanabe P, Gimbel R, Yarnold PR, Adams JG. The Emergency Severity Index (version 3) 5-level triage system scores predict ED resource consumption. *J Emerg Nurs*. 2004;30(1):22-29.
- Mirhaghi A, Heydari A, Mazlom R, Hasanzadeh F. Reliability of the emergency severity index: Meta-analysis. *Sultan Qaboos Univ Med J*. 2015;15(1):e71-e77.
- Righi AW, Saurin TA, Wachs P. A systematic literature review of resilience engineering: Research areas and a research agenda proposal. *Reliab Eng Syst Saf*. 2015;141:142-152.

AUTHOR BIOGRAPHY



Nicholas Sterling, MD, PhD, is an emergency resident in the Emory University School of Medicine, Atlanta, GA.

How to cite this article: Sterling NW, Brann F, Patzer RE, et al. Prediction of emergency department resource requirements during triage: An application of current natural language processing techniques. *JACEP Open*. 2020;1:1676-1683. <https://doi.org/10.1002/emp212253>