

Capturing Concerns about Patient Deterioration in Narrative Documentation in Home Healthcare

Mollie Hobensack, BSN, RN¹, Jiyoun Song, PhD, RN¹, Sena Chae, PhD, RN³, Erin Kennedy, BSN, RN⁴, Maryam Zolnoori, PhD¹, Kathryn H. Bowles, PhD, RN^{2,4}, Margaret V. McDonald, MSW², Lauren Evans, DrPH², Maxim Topaz, PhD, RN^{1,2}

¹Columbia University School of Nursing, New York, NY, USA; ²Center for Home Care Policy & Research, Visiting Nurse Service of New York, New York, NY, USA; ³University of Iowa College of Nursing, Iowa City, IA, USA; ⁴University of Pennsylvania School of Nursing, Philadelphia, PA, USA

Abstract

Home healthcare (HHC) agencies provide care to more than 3.4 million adults per year. There is value in studying HHC narrative notes to identify patients at risk for deterioration. This study aimed to build machine learning algorithms to identify “concerning” narrative notes of HHC patients and identify emerging themes. Six algorithms were applied to narrative notes (n = 4,000) from a HHC agency to classify notes as either “concerning” or “not concerning.” Topic modeling using Latent Dirichlet Allocation bag of words was conducted to identify emerging themes from the concerning notes. Gradient Boosted Trees demonstrated the best performance with a F-score = 0.74 and AUC = 0.96. Emerging themes were related to patient-clinician communication, HHC services provided, gait challenges, mobility concerns, wounds, and caregivers. Most themes have been cited by previous literature as increasing risk for adverse events. In the future, such algorithms can support early identification of patients at risk for deterioration.

Introduction

In the United States, home healthcare (HHC) agencies provide care to more than 3.4 million adults per year¹. One objective of HHC is to identify early signs of clinical deterioration to activate prompt interventions to reduce the risk of adverse events such as hospitalization or an emergency department (ED) visit². One in five patients entering HHC are hospitalized during their homecare episode (i.e., period when patients receive HHC services)^{3,4}. Despite recent efforts, these numbers have not improved over the last several years³ and in fact, recent estimates show that approximately 30% of hospitalizations and ED visits are preventable and can be avoided with timely and patient tailored care⁵.

There is emerging evidence that warning systems embedded in the EHR can support early recognition of patient deterioration⁶. Within EHR documentation, researchers have leveraged narrative notes to identify patients at risk for clinical deterioration in HHC^{7,8}. To analyze narrative notes, novel methods such as machine learning-based natural language processing (NLP) can be applied to process the unstructured text of narrative notes⁹. Previous studies have demonstrated the value of NLP to extract concerning information in narrative notes to identify patients at higher risk for hospitalization or an ED visit^{7,8,10-14}. A recent study interviewing HHC clinicians found that clinicians often document patient deterioration in narrative notes rather than in the structured data¹¹. While past studies have focused on specific disease related factors such as symptoms^{7,10}, fewer studies have examined deterioration taking into consideration a broader array of factors. This expanded approach can help further identify concern related to clinician intuition that may not be directly related to a specific disease process¹².

In recent years, topic modeling, Latent Dirichlet Allocation (LDA), has gained popularity in biomedical literature to identify emerging themes in a subset of documents^{13,14}. This bag of words approach uses probability to identify sets of words that cluster together in a collection of documents¹⁵. Similar to previous studies that applied this technique to narrative notes, this approach can be leveraged to support the development of future prediction models by helping to characterize key themes related to patient deterioration^{13,16}. Thus, this method supports further understanding and exploration of narrative notes describing clinician concern.

In this study we created algorithms that identified concerning multidisciplinary narrative notes of HHC patients. For the purposes of this study, “concerning” was defined as the presence of factors associated with increased risk for deterioration such as hospitalization or an ED visit¹⁷⁻¹⁹. The aims of this study sought to 1) identify the best performing machine learning algorithm to identify concerning narrative notes, and 2) describe the emerging themes

of the concerning notes. Ultimately, both aims support our goal to develop a comprehensive picture of patients who are at risk for hospitalization or an ED visit and how this is reflected in narrative notes.

Methods

STUDY DATA

This secondary analysis used retrospective EHR data from a not-for-profit HHC agency located in the Northeastern United States. Data included narrative notes from the years 2015 to 2017. Narrative notes were completed by multidisciplinary HHC clinicians (registered nurses, physical therapists, occupational therapists, and social workers). Our study evaluated two types of HHC narrative notes: 1) visit notes and 2) care coordination notes. Visit notes describe care provided and the patient’s status during a HHC visit. Care coordination notes describe communication between clinicians (e.g., calling a physician) and other care-related activities (e.g., ordering wound care supplies). All parts of this study were approved by the Institutional Review Board of the participating institutions.

NARRATIVE NOTE ANNOTATION

During the annotation stage registered nurses with a Bachelor’s degree or higher reviewed a subset of 4,000 randomly selected narrative notes and binary coded each narrative note as either “concerning” or “not concerning.” The annotators were asked to use their clinical judgment in making annotation decisions, guided by the general question: “Please indicate if this narrative note describes a concerning patient situation that could result in deterioration, such as a hospitalization or ED visit.”

Three annotators (MH, JS, and MT) independently reviewed 300 notes and marked notes with a 0, “not concerning” or a 1, “concerning.” Fleiss’s Kappa, a statistic for measuring inter-rater reliability agreement²⁰, was calculated between the three annotators and weak agreement was achieved ($\kappa = 0.579$). Rationale for weak agreement was related to one annotator being more conservative in coding concerning notes. Full agreement was achieved after discussion between the three annotators and the decision to be more inclusive in coding concerning notes to capture a broad spectrum of concern reflected in the narrative notes. Table 1 includes examples of narrative notes that were classified as “concerning” or “not concerning.” Given the strong eventual inter-annotator agreement, an additional 3,700 notes were reviewed independently by two of the annotators (MH and JS). Fleiss’s Kappa was re-calculated and substantial agreement was achieved ($\kappa = 0.8$) with 93% pairwise agreement. Full agreement was achieved following discussion. The final sample included 4,000 narrative notes.

Table 1. Examples of narrative notes identified as “Concerning” or “Not concerning”

Narrative notes <i>with</i> concerning language (Concerning)	Narrative notes <i>without</i> concerning language (Not concerning)
Patient in bed reported 'too much pain.' Reviewed home exercise plan and pain management techniques with patient/home health aide.	Patient compliant with home exercise program. Patient encouraged to increase daily activities to increase ambulation with rolling walker. Patient verbalized understanding.
Patient lives alone in private home with 3 daughters who are supportive but live out of state, so cannot visit regularly. Patient having increased difficulty with ambulation and activities of daily living due to pain, weakness, and decreased sensation in hands.	Patient independent with activities of daily living, ambulation, and tasks. Patient currently has 24/7 care, daughter in same building who also provides support.
Patient has a history of multiple falls due to syncopal episodes/dizziness. Home safety/fall prevention strategies discussed. Patient refused to do transfers today due to high pain level in left lower extremity. Patient uses rolling chair to navigate apartment. Patient scheduled to get hospital bed.	Patient present with no signs of distress. Plan for next visit, wound observation and further teaching.

Table 1 Continued. Examples of narrative notes identified as “Concerning” or “Not concerning”

<p>Patient seen for wound care, patient sounds very druggy with pain meds, tolerated treatment well and was asked to follow up with vascular medical doctor because of moderate drainage to right leg ulcers, patient retaining fluids to lower extremities. Continue with care.</p>	<p>Decreased swelling noted in right hand, patient reporting less discomfort. Increased independence for dressing and transfers. Occupational therapy discharge plan scheduled for next week. Patient and spouse notified and in agreement.</p>
--	---

DATA ANALYSIS

We applied six machine learning analytic methods to develop algorithms that would identify concerning narrative notes using the software KNIME²¹. KNIME is an open-source software that allows researchers to apply data science methods such as NLP using “nodes”²¹. The machine learning algorithms utilized in this study were: Multilayer Perceptron (i.e., Neural Network), Decision Tree, Random Forest, Support Vector Machine (SVM), Gradient Boosted Tree, and Logistic Regression. These algorithms were selected based on their common use in biomedical literature in HHC^{15,16}. To identify concerning narrative notes using machine learning algorithms, an analytical pipeline was created with the following components: 1. Text Preprocessing, 2. Model Development (Algorithm Training), and 3. Evaluating Model Performance. We then conducted topic modeling to identify emerging themes of the concerning notes.

Component 1: Text Preprocessing

We used KNIME, an open-source data analytic platform²¹, to preprocess the narrative notes. Preprocessing improved the identification of concerning notes by filtering the notes down to their core elements to facilitate analysis by the machine learning algorithms^{22,23}. Classification tasks conducted in the text preprocessing stage included omitting punctuations, snowball stemming, omitting numbers, omitting common English stop words, and converting words to lowercase⁸.

Component 2: Model Development (Algorithm Training)

The labeled dataset was randomly split: 70% for training the machine learning algorithms, and 30% for validating the performance. Algorithms were applied and fine-tuned using KNIME’s “Parameter Optimization Loop” node (settings: stop value = 15, step size = 0.1, search strategy = Hill climbing). The following machine learning algorithms were applied using KNIME: (a) Multilayer Perceptron, applied with maximum iterations = 50, hidden layer = 1, and number of hidden neurons per layer = 10 (b) Decision Tree, pruned with minimum description length using gain ratio, number of records per node = 2, records to store for view = 10,000, and number of threads = 12 (c) Random Forest, applied with number of models = 100, minimum number of instances = 1, and number of levels = 10 (d) SVM, applied with overlapping penalty = 1, and using the polynomial kernel (e) Gradient Boosted Tree, applied with tree depth = 4, number of models = 100, and a learning rate = 0.1 and (f) Logistic Regression applied with Stochastic average gradient, a technique used to find model parameters that are most appropriate for fitting the predicted and actual outcomes.

Component 3: Evaluating Model Performance

A portion of labeled data (30%) were held for validating the machine learning algorithm performance. Performance was measured on the test dataset using the standard metrics of *precision*- a measure of correct instances out of retrieved instances; *recall*- a measure of correct instances retrieved out of all correct instances; and *F-score*- the harmonic mean of precision and recall²⁴. In addition, *Area Under the Curve (AUC)* was generated using the “Received Operating Curve” KNIME node. AUC is an informative metric of performance that accounts for both positive and negative cases²⁵.

Topic Modeling

To understand emerging themes among the concerning notes, an exploratory analysis was conducted using the topic modeling method LDA¹⁵. LDA uses a bag of words approach that converts text to vectors to identify how many times a word appears in a document²⁶. To identify the appropriate number of topics, a Scree Plot was generated using the “Principal Component Analysis” KNIME node. The elbow in the Scree Plot helps to identify where eigenvalues level, indicating that additional topics past that point are less significant. From the Scree Plot, we were able to identify six as the appropriate number of topics. LDA was applied to only the concerning notes (n = 795) to identify emerging themes. The LDA parameters included the following: number of topics = 6, alpha = 0.01, number of iterations = 1,000, number of words per topic = 10, beta = 0.01, number of threads = 8.

We used an iterative process to name each topic in the context of concerning narrative notes. First, three members of the team independently named each grouping of ten words with a topic name (MH, SC, EK). Topic names were presented to all authors for discussion on the most appropriate topic name to describe the set of words. Authors participating in the conversation have expertise in HHC, informatics, and nursing. Authors were asked to consider topic names in the context of how it relates to documentation describing concern about patient deterioration in HHC. Final topic names were identified through consensus among authors.

Results

Six machine learning algorithms were applied to 4,000 narrative notes and the results are presented below in Table 2. We found that visit notes had an average note length of 272.3 characters and care coordination notes had an average note length of 106.6 characters. Of the 4,000 narrative notes, clinicians documented concern in approximately 20% (n = 795). Apart from Decision Trees, all machine learning algorithms demonstrated acceptable AUC performance (0.7 - 0.8) with Logistic Regression demonstrating excellent performance (0.8 – 0.9), and Gradient Boosted Trees demonstrating outstanding performance (> 0.9)²⁷. Best performance across all machine learning algorithms was achieved using Gradient Boosted Trees with a F-score = 0.74 and AUC = 0.96.

Table 2. Machine learning algorithms' performance

Model type	Precision	Recall	Weighted F-score	AUC
Multilayer Perceptron	0.48	0.45	0.46	0.75
Decision Trees	0.53	0.25	0.34	0.64
Random Forests	0.54	0.31	0.39	0.80
Support Vector Machines	0.56	0.36	0.44	0.80
Gradient Boosted Trees	0.99	0.59	0.74	0.96
Logistic Regression	0.59	0.19	0.29	0.81

* AUC: Area Under the Curve

Based on the Scree Plot, six topics were identified as appropriate with 10 words in each topic (Table 3). Percent weights were assigned to each word, with higher weights indicating stronger association with each topic. Following group consensus, emerging themes or topics from concerning notes related to patient-clinician communication, HHC services provided, gait challenges, mobility concerns, wounds, and caregivers.

Table 3. Topic Modeling Results using Latent Dirichlet Allocation

Topic #	Topic Label	Top 10 Keywords and Weights (%)
Topic 1	Patient-Clinician Communication	Pain (18.6), patient (18), stated (8.1), instructed (7.4), visit (7.2), reports (7.0), reported (6.4), appointment (6.1), meds (5.7), home (5.6)
Topic 2	HHC Services Provided	Patient (11.6), due (8.3), pain (2.7), visit (5.2), unable (5.1), reports (4.3), able (4.0), hha (3.5), management (3.5), status (3.2)
Topic 3	Gait Challenges	Patient (7.5), pain (6.3), visit (4.1), refused (2.9), aware (2.8), fall (2.7), hospital (2.6), difficulty (2.4), gait (2.4), informed (2.3)
Topic 4	Mobility Concerns	Patient (24.5), lives (6.7), home (5.5), gait (5.3), due (4.8), balance (4.6), visit (4.6), private (4.4), weakness (3.9), difficulty (3.4)
Topic 5	Wounds	Wound (20.2), care (19.5), patient (13.7), skin (7.7), left (6.6), noted (5.9), visit (5.3), hha (5.0), instructed (4.8), called (4.7)
Topic 6	Caregivers	Lives (11.0), patient (9.9), home (6.8), female (5.6), htn (5.0), pmh (4.8), spouse (4.5), son (4.3), alert (4.2), supportive (4.0)

* HHC: home healthcare, hha: home health aide, htn: hypertension, pmh: past medical history

Discussion

In this study, we evaluated multidisciplinary narrative notes in the EHR to identify HHC patients that clinicians may be concerned about for deterioration such as a hospitalization or an ED visit. While other studies have used narrative notes to identify patient deterioration in the inpatient setting^{27,28}, this study was one of the first in HHC to focus on identifying concern expressed broadly in narrative notes rather than focusing on disease-specific symptoms. One reason this approach is valuable is because it can potentially capture clinician intuition describing concerning observations that may not be directly related to a specific disease process such as socio-behavioral factors¹².

This study compared six machine learning algorithms and found that the best performing algorithm to be Gradient Boosted Trees (Weighted F-score = 0.74; AUC = 0.96). Tree-based algorithms are a reliable method in biomedical literature²⁹. Specifically, Gradient Boosted Trees has demonstrated strong performance in classification studies lending itself to its selection in this study^{30,31}. Gradient Boosted Trees are based on tree algorithms similar to Decision Trees (i.e., the building block) and Random Forest; however, Gradient Boosted Trees are weak learners and use an ensemble approach that build on one tree at a time allowing weaker trees to be strengthened in each new iteration. Alternatively, Random Forest creates iterations independently³². Rationale for why this algorithm demonstrated high performance in this study may be due to its ability to work well with intermediate sample sizes and adapt with nonlinearity³³. However, it is important to note that Gradient Boosted Trees are prone to overfitting and thus, this model should be cross-validated on other datasets prior to being integrated on a larger scale³³.

Following model development, we conducted LDA topic modeling to discover emerging themes from the concerning notes. We found that topics related to patient-clinician communication, HHC services provided, gait challenges, mobility concerns, wounds, and caregivers. Even though it can be difficult to assess without contextual information, in general, communication between clinicians and between patients and clinicians in the HHC setting have shown to be an indicator of clinician concern for patient deterioration^{11,34}. It was found that HHC clinicians would contact providers to request additional HHC services needed to support patient care and then document this communication¹¹. In interviews with HHC clinicians, it was said that documentation is driven by care coordination to communicate concerning information to other care team members via the EHR and to demonstrate the added value of HHC services^{35,36}. This suggests documenting patient statements (e.g., stated, reports, etc.) in narrative notes as a way to relay patient driven concerns and information between clinicians in the HHC setting³⁶.

Other studies have demonstrated that functional limitations, such as gait challenges and mobility concerns^{8,37}, can increase risk for hospitalization during HHC. In the context of concerning notes, functional limitations may be related to increased risk for fall-related injuries which are a frequently cited cause for hospitalizations^{38,39}. Functional limitations may also increase an older adult's likelihood of developing wounds or delay the healing of wounds which has been cited as one of the most common infections leading to hospitalization in HHC⁴⁰. Thus, unsurprisingly, documentation about functional limitation (i.e., gait challenges and mobility concerning) and wounds were themes in concerning narrative notes suggesting these patient cohorts may be at increased risk for deterioration.

Having a broader definition of concern makes it more feasible to identify important social determinants of health such as caregiver support or lack of caregiver support which has not been a previous focus in the HHC setting⁴¹⁻⁴³. In interviews with HHC clinicians, they stated that lack of social support is a risk factor for hospitalization⁴⁴. Therefore, it is interesting that an emerging theme of concerning notes relates to caregivers which previous studies have described as a protective factor^{11,35,45}. However, a possible explanation for this is that the presence of a caregiver could indicate that a patient has health needs unable to be met independently. Alternatively, the documentation of caregiver terms could represent an absence of social support (e.g., no spouse, lives alone). While it is likely that some aspect of social support is documented in concerning notes given its importance in HHC^{35,46}, future exploration is needed to better understand the context of these terms in HHC narrative notes.

While topic modeling provided a snapshot of themes from the coded concerning notes, future studies could consider comparing these to the not concerning notes to identify potential differences or similarities. Alternatively, researchers could consider creating a vocabulary related to each of these domains (e.g., communication, mobility, wounds) and exploring their use in narrative notes. Creating a vocabulary and exploring if these topics are associated with a deterioration outcome such as hospitalization could support stronger model development in capturing patients who may be at risk for deterioration.

The results of this study suggest the feasibility of applying machine learning algorithms to HHC narrative notes to broadly identify clinician concern. Future studies involving longitudinal data over a patient's HHC episode could focus on identifying the proportion of concerning notes for patients during their time in HHC with higher proportions of concerning notes suggesting higher acuity or the need for an intervention. This in turn, could help HHC managers in making clinical assignments balancing patients' care needs. Current work by our team focuses on the added value of this algorithm and how it complements other machine learning algorithms to identify patient deterioration⁴⁷. In our other work, structured and unstructured EHR data were utilized to build risk models¹⁸. It was found that the algorithm developed in this manuscript in combination with another algorithm grounded in a standardized terminology¹⁹ demonstrated the best performance in identifying patients at risk for hospitalization or an ED visit⁴⁷. Future steps include continuing to develop the most comprehensive model to identify HHC patients at risk for deterioration through an early warning system that can be implemented into practice.

Limitations

The study used data from one large, urban home health agency and therefore may not be representative of documentation practices found at other organizations. While this paper adds to the literature by developing algorithms to help identify HHC patients who may be deteriorating there are a few limitations. First, this study uses a small subset of 4,000 narrative notes from a HHC agency with approximately 20% (n = 795) of the notes identified as concerning. Future studies will focus on further characterizing this cohort of patients and applying our findings to the larger organization dataset (approximately 66,000 patients and 2.3 million notes). In addition, while the authors intentionally chose a wide range of algorithms, future studies should explore advanced machine learning algorithms. In addition, alternative topic modeling approaches that consider context could be further explored to better understand each topic. Authors fine-tuned the algorithms using parameter optimization; however, further fine-tuning using alternative software could lead to better performance. Finally, authors intentionally did not include the outcome hospitalization in this study given that one concerning note would not necessarily be a true reflection of a patient's likelihood for deterioration. Future studies applying this algorithm to a larger sample and calculating a proportion of concerning notes will consider this outcome of interest.

Conclusion

This study is the first in HHC to build an algorithm to identify concern broadly in multidisciplinary narrative notes. Our results show that applying machine learning algorithms to HHC narrative notes is feasible and can help to identify patients who may be at risk for deterioration such as hospitalization or an ED visit. Emerging themes from the concerning notes are supported by previous literature and can be further explored to evaluate key risk factors for deterioration contained in HHC narrative notes. The algorithms developed in this study can complement future early warning systems aimed at alerting HHC clinicians to patients at risk for deterioration.

Funding information

This research was supported by grant R01HS027742 from the Agency for Healthcare Research and Quality (PI: Dr. Maxim Topaz). The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality. Members of our team are supported by the National Institute for Nursing Research (T32NR007969-MH and F31NR019919-EK) and the Jonas Scholarship (MH).

References

1. MedPac. [Internet]. [cited 2021 Aug 1]. Available from: http://www.medpac.gov/docs/default-source/reports/mar20_medpac_ch9_sec.pdf
2. Gray E, Currey J, Considine J. Hospital in the home nurses' recognition and response to clinical deterioration. *J Clin Nurs*. 2018;27(9–10):2152–60.
3. Home Health Quality Measures. [Internet]. [cited 2021 Oct 2]. Available from: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/Home-Health-Quality-Measures>
4. Busby J, Purdy S, Hollingworth W. A systematic review of the magnitude and cause of geographic variation in unplanned hospital admission rates and length of stay for ambulatory care sensitive conditions. *BMC Heal Serv Res*. 2015;15(1):1–15.
5. Solberg LI, Ohnsorg KA, Parker ED, Ferguson R, Magnan S, Whitebird RR, et al. Potentially preventable hospital and emergency department events: Lessons from a large innovation project. *Perm J*. 2018;22:17–102.
6. Wood C, Chaboyer W, Carr P. How do nurses use early warning scoring systems to detect and act on patient deterioration to ensure patient safety? A scoping review. 2019;94:166–78.
7. Topaz M, Koleck TA, Onorato N, Smaldone A, Bakken S. Nursing documentation of symptoms is associated with higher risk of emergency department visits and hospitalizations in homecare patients. *Nurs Outlook*. 2021;69(3):435–46.
8. Topaz M, Woo K, Ryvicker M, Zolnoori M, Cato K. Home healthcare clinical notes predict patient hospitalization and emergency department visits. *Nurs Res*. 2020;69(6):448–54.
9. Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Futur Healthc J*. 2019;6(2):94.
10. Topaz M, Adams V, Wilson P, Woo K, Ryvicker M. Free-text documentation of dementia symptoms in home healthcare: A natural language processing study. *Gerontol Geriatr Med*. 2020;6:233372142095986.
11. Hobensack M, Ojo M, Barron Y, Bowles KH, Chae S, Kennedy E, et al. Documentation of hospitalization risk factors in electronic health records (EHRs): a qualitative study with home healthcare clinicians. 2022;ocac023.
12. Melin-Johansson C, Palmqvist R, Rönnerberg L. Clinical intuition in the nursing process and decision-making—A mixed-studies review. *J Clin Nurs*. 2017;26(23–24):3936–49.
13. Boag W, Kovaleva O, McCoy TH, Rumshisky A, Szolovits P, Perlis RH. Hard for humans, hard for machines: predicting readmission after psychiatric hospitalization using narrative notes. *Transl Psychiatry*. 2021;11(1):1–6.
14. Wang L, Miloslavsky E, Stone JH, Choi HK, Zhou L, Wallace ZS. Topic modeling to characterize the natural history of ANCA-Associated vasculitis from clinical notes: A proof of concept study. *Semin Arthritis Rheum*. 2021;51(1):150.
15. Kavvadias S, Drosatos G, Kaldoudi E. Supporting topic modeling and trends analysis in biomedical literature. *J Biomed Inform*. 2020 Oct 1;110:103574.
16. Rumshisky A, Ghassemi M, Naumann T, Szolovits P, Castro VM, McCoy TH, et al. Predicting early psychiatric readmission with natural language processing of narrative discharge summaries. *Transl Psychiatry*. 2016;6(10):e921.
17. Hodge SY, Ali MR, Gordon AL. Recognizing and responding to deterioration in care homes: A scoping review protocol. *JBIM Evid Synth*. 2020;19(2):447–53.
18. Song J, Hobensack M, Bowles KH, McDonald MV, Cato K, Rossetti SC, et al. Clinical notes: An untapped opportunity for improving risk prediction for hospitalization and emergency department visit during home health care. *J Biomed Inform*. 2022;128:104039.
19. Song J, Ojo M, Bowles KH, McDonald MV, Cato KD, Rossetti SC, et al. Detecting Language Associated with Home Healthcare Patient's Risk for Hospitalization and Emergency Department Visit. *Nurs Res*. 2021;
20. Fleiss JL, Levin B, Paik MC. *Statistical Methods for Rates and Proportions*. Statistical Methods for Rates and Proportions John Wiley & Sons, Inc.; Sep 5, 2003.
21. Berthold MR, Cebron N, Dill F, Fatta GD, Gabriel TR, Georg F, et al. Knime: The Konstanz Information Miner [Internet]. [cited 2021 Feb 22]. Available from: <http://www.knime.org>.
22. Jones CD, Falvey J, Hess E, Levy CR, Nuccio E, Barón AE, et al. Predicting hospital readmissions from home healthcare in medicare beneficiaries. *J Am Geriatr Soc*. 2019;67(12):2505–10.
23. Song J, Woo K, Shang J, Ojo M, Topaz M. Predictive risk models for wound infection-related hospitalization or ED visits in home health care using machine-learning algorithms. *Adv Skin Wound Care*.

- 2021;34(8):1–12.
24. Sokolova M, Japkowicz N, Szpakowicz S. Beyond accuracy, f-score, and ROC: A family of discriminant measures for performance evaluation. 2006.
 25. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol*. 2010 Sep 1;5(9):1315–6.
 26. Zhang Y, Jin R, Zhou ZH. Understanding Bag-of-Words Model: A Statistical Framework
 27. Collins S. How extra nursing notes point to deterioration. *Nurs Times*. 2014;110(22):19–20.
 28. Korach ZT, Yang J, Rossetti SC, Cato KD, Kang M-JJ, Knaplund C, et al. Mining clinical phrases from nursing notes to discover risk factors of patient deterioration. *Int J Med Inform*. 2020;135:104053.
 29. Olson RS, La Cava W, Mustahsan Z, Varik A, Moore JH. Data-driven advice for applying machine learning to bioinformatics problems. *Pacific Symp Biocomput*. 2017;0(212669):192–203.
 30. Ehrentraut C, Ekholm M, Tanushi H, Tiedemann J, Dalianis H. Detecting hospital-acquired infections: A document classification approach using support vector machines and gradient tree boosting. *Health Informatics J*. 2018;24(1):24–42.
 31. Li K, Shi Q, Liu S, Xie Y, Liu J. Predicting in-hospital mortality in ICU patients with sepsis using gradient boosting decision tree. *Medicine*. 2021;100(19):e25813.
 32. Ye J, Chow JH, Chen J, Zheng Z. Stochastic gradient boosted distributed decision trees. 2009.
 33. Xie Y, Jiang B, Gong E, Li Y, Zhu G, Michel P, et al. Use of gradient boosting machine learning to predict patient outcome in acute ischemic stroke on the basis of imaging, demographic, and clinical information. 2018;212(1):44–51.
 34. Pesko MF, Gerber LM, Peng TR, Press MJ. Home health care: Nurse–physician communication, patient severity, and hospital readmission. *Health Serv Res*. 2018 Apr 1;53(2):1008–24.
 35. Landers S, Madigan E, Leff B, Rosati RJ, McCann BA, Hornbake R, et al. The future of home health care: A strategic framework for optimizing value. *Home Health Care Manag Pract*. 2016;28(4):262.
 36. Hobensack M, Ojo M, Bowles K, McDonald M, Song J, Topaz M. Home healthcare clinicians’ Perspectives on electronic health records: A qualitative study. *Stud Health Technol Inform*. 2021;284:426–30.
 37. Patel KV, Guralnik JM, Phelan EA, Gell NM, Wallace RB, Sullivan MD, et al. Symptom burden among community-dwelling older adults in the United States. *J Am Geriatr Soc*. 2018;67(2):jgs.15673.
 38. Manis DR, McArthur C, Costa AP. Associations with rates of falls among home care clients in Ontario, Canada: A population-based, cross-sectional study. *BMC Geriatr*. 2020;20(1):1–12.
 39. Adams CM, Tancredi DJ, Bell JF, Catz SL, Romano PS. Associations between home injury falls and prior hospitalizations in community dwelling older adults: A population case-crossover study. *Injury*. 2020;51(2):260–6.
 40. Shang J, Wang J, Adams V, Ma C. Risk factors for infection in home health care: Analysis of national outcome and assessment information set data. *Res Nurs Health*. 2020;43(4):373.
 41. Kino S, Hsu YTT, Shiba K, Chien YSS, Mita C, Kawachi I, et al. A scoping review on the use of machine learning in research on social determinants of health: Trends and research prospects. *SSM - Popul Heal*. 2021;15:100836.
 42. Bompelli A, Wang Y, Wan R, Singh E, Zhou Y, Xu L, et al. Social and behavioral determinants of health in the era of artificial intelligence with electronic health records: A scoping review. *Heal Data Sci*. 2021.
 43. Patra BG, Sharma MM, Vekaria V, Adekkanattu P, Patterson O V, Glicksberg B, et al. Extracting social determinants of health from electronic health records using natural language processing: a systematic review. *J Am Med Informatics Assoc*. 2021;2021(0):1–12.
 44. Navathe AS, Zhong F, Lei VJ, Chang FY, Sordo M, Topaz M, et al. Hospital readmission and social risk factors identified from physician notes. *Health Serv Res*. 2018;53(2):1110.
 45. Mah JC, Stevens SJ, Keefe JM, Rockwood K, Andrew MK. Social factors influencing utilization of home care in community-dwelling older adults: a scoping review. *BMC Geriatr*. 2021;21(1):1–21.
 46. Bick I, Dowding D. Hospitalization risk factors of older cohorts of home health care patients: A systematic review. 2019:111–52
 47. Song J, Hobensack M, Bowles KH, McDonald MV, Cato K, Rossetti SC, et al. Clinical notes: An untapped opportunity for improving risk prediction for hospitalization and emergency department visit during home health care. *J Biomed. Inform*. 2022;128:104039.