

Identifying and reducing stigmatizing language in home healthcare: the ENGAGE study protocol

Zhihong Zhang, Pallavi Gupta, Stephanie Potts-Thompson, Laura Prescott, Morgan Morrison, Scott Sittig, Margaret V. McDonald, Chase Raymond, Jacquelyn Y. Taylor, Maxim Topaz

Submitted to: JMIR Research Protocols
on: December 06, 2024

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript..... 5

Supplementary Files..... 28

 Figures 29

 Figure 1..... 29

 Figure 2..... 29

 Figure 3..... 29

 Figure 4..... 29

 Multimedia Appendixes 30

 Multimedia Appendix 1..... 31

Identifying and reducing stigmatizing language in home healthcare: the ENGAGE study protocol

Zhihong Zhang¹ PhD, RN; Pallavi Gupta² PhD, MS; Stephanie Potts-Thompson³ MPH; Laura Prescott³ MA; Morgan Morrison³ MA; Scott Sittig⁴ PhD, MHI, RHIA; Margaret V. McDonald⁵ MSW; Chase Raymond^{6,7} PhD; Jacquelyn Y. Taylor³ PhD, PNP-BC, RN, FARM, FAAN; Maxim Topaz^{1,2} PhD, RN, MA

¹Columbia University Data Science Institute New York US

²Columbia University School of Nursing New York US

³Columbia University School of Nursing Center for Research on People of Color New York US

⁴University of Louisiana at Lafayette Health Sciences Lafayette US

⁵Center for Home Care Policy & Research VNS Health New York US

⁶Department of Linguistics University of Colorado Boulder US

⁷Department of Family Medicine Anschutz Medical Campus University of Colorado Aurora US

Corresponding Author:

Maxim Topaz PhD, RN, MA

Columbia University School of Nursing

560 West 168th Street

New York

US

Abstract

Background: Stigmatizing language is prevalent in clinical notes, adversely affecting patient care quality. Natural language processing (NLP) is a promising technology for analyzing millions of clinical notes in electronic health records.

Objective: This study proposes an NLP-driven ENGAGE system to automatically identify and replace stigmatizing language.

Methods: This mixed-method study will extract electronic health record (EHR) data for patients admitted to two large, diverse home healthcare (HHC) organizations between January 2019, and December 2021. The study will be conducted in four aims: Aim 1 will refine the ontology of stigmatizing language in HHC by (a) interviewing a diverse sample of HHC nurses and patients to identify terms to avoid, and (b) analyzing clinical notes from various U.S. regions to categorize stigmatizing language. Aim 2 will determine the best NLP approach for accurately identifying stigmatizing language by training algorithms and comparing their performance to human annotations. Aim 3 will analyze the prevalence of stigmatizing language based on patients' race and ethnicity using adjusted statistical analyses of a sample of approximately half a million HHC patients (34% racial and ethnic minorities). Aim 4 will develop the NLP-driven ENGAGE system by (a) testing NLP methods (rule-based, "Delete, Retrieve, Generate," and transformers) for suggesting alternative wording, and (b) designing and refining the user interface for clinical trial preparation.

Results: This study was funded by the National Institute on Minority Health and Health Disparities. Recruitment, enrollment, and data curation are ongoing.

Conclusions: This study will leverage extensive data sources to examine stigmatizing language in HHC settings, contributing to developing systems to effectively reduce such language among HHC nurses. Clinical Trial: NA

(JMIR Preprints 06/12/2024:69753)

DOI: <https://doi.org/10.2196/preprints.69753>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ Please make my preprint PDF available to anyone at any time (recommended).

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to the public.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org/>, I will be able to make my accepted manuscript PDF available to anyone at any time.



Original Manuscript

Original Paper

Identifying and reducing stigmatizing language in home healthcare: the ENGAGE study protocol

Zhihong Zhang, PhD, RN¹, Pallavi Gupta, PhD, MS², Stephanie Potts-Thompson, MPH³, Laura Prescott, MA³, Morgan Morrison, MA³, Scott Sittig, PhD, MHI, RHIA⁴, Margaret V. McDonald, MSW⁵, Chase Wesley Raymond, PhD^{6,7}, Jacquelyn Y. Taylor PhD, PNP-BC, RN, FAHA, FAAN³, Maxim Topaz PhD, RN, MA^{1,2,5*}

1 Columbia University Data Science Institute, New York, New York, United States of America

2 Columbia University School of Nursing, New York, New York, United States of America

3 Columbia University School of Nursing, Center for Research on People of Color, New York, New York, United States of America

4 University of Louisiana at Lafayette, Health Sciences, Lafayette, Louisiana, United States of America

5 Center for Home Care Policy & Research, VNS Health, New York, New York, United States of America

6 University of Colorado, Department of Linguistics, Boulder, Colorado, United States of America

7 University of Colorado, Department of Family Medicine, Anschutz Medical Campus, Aurora, Colorado, United States of America

*Corresponding author

Email: mt3315@cumc.columbia.edu

Abstract

Background: Stigmatizing language is prevalent in clinical notes, adversely affecting patient care quality. Natural language processing (NLP) is a promising technology for analyzing millions of clinical notes in electronic health records.

Objective: This study proposes an NLP-driven ENGAGE system to automatically identify and replace stigmatizing language.

Methods: This mixed-method study will extract electronic health record (EHR) data for patients admitted to two large, diverse home healthcare (HHC) organizations between January 2019, and December 2021. The study will be conducted in four aims: Aim 1 will refine the ontology of stigmatizing language in HHC by (a) interviewing a diverse sample of HHC nurses and patients to identify terms to avoid, and (b) analyzing clinical notes from various U.S. regions to categorize stigmatizing language. Aim 2 will determine the best NLP approach for accurately identifying stigmatizing language by training algorithms and comparing their performance to human annotations. Aim 3 will analyze the prevalence of stigmatizing language based on patients' race and ethnicity using adjusted statistical analyses of a sample of approximately half a million HHC patients (34% racial and ethnic minorities). Aim 4 will develop the NLP-driven ENGAGE system by (a) testing NLP methods (rule-based, "Delete, Retrieve, Generate," and transformers) for suggesting alternative wording, and (b) designing and refining the user interface for clinical trial preparation.

Results: This study was funded by the National Institute on Minority Health and Health Disparities.

Recruitment, enrollment, and data curation are ongoing.

Discussion: This study will leverage extensive data sources to examine stigmatizing language in HHC settings, contributing to developing systems to effectively reduce such language among HHC nurses.

Keywords: Natural language processing; Racial bias; Stigmatizing language; Home Healthcare

Introduction

Home healthcare (HHC) is one of the fastest-growing outpatient settings in the United States (US), where 200,000 nurses provide care for more than 5 million patients annually [1,2]. Although the quality of nursing care is affected by numerous factors (e.g., structural resources, levels of education, patient-per-nurse ratios) [3-5], nurses' biases towards their patients influence the delivery of high-quality care [6]. Implicit bias is the negative inclination of one group and its members relative to others unconsciously or unintentionally [6]. Recent literature reviews [7-8] found widespread implicit biases among nurses toward their patients. Specifically, a recent review of 215 studies [9] found that nurses frequently display biases based on patients' race or ethnicity, influencing treatment decisions and impacting patient adherence and outcomes [10,11].

Racial biases can be propagated via language used in electronic health records (EHR) documentation [12-15]. Research has shown that stigmatizing language used in clinical notes can harm patient care [12-15]. Patients for whom stigmatizing language was used had HHC visits that were 24 minutes shorter compared to patients without such language (average visit length = 46 vs 70 minutes, respectively) [16,17]. This is concerning since shorter HHC visits are associated with poor outcomes (e.g., higher risk of hospitalizations) [10-12,16-22]. A recent study found that 10% of 22,959 patients who reviewed their clinical notes felt judged or offended by stigmatizing language [23]. This is crucial since, as of April 2021, healthcare organizations, including HHC, must share EHR data with patients under the 21st Century Cures Act's "Information Blocking" Rule [24]. More than 80% of HHC agencies have EHRs, and reducing stigmatizing language use can decrease racial biases and improve the quality of care and patient outcomes [25].

When applied appropriately, technology can help identify and reduce biases in healthcare [26]. One promising technology is natural language processing (NLP), an algorithm that analyzes and extracts meaning from EHR clinical notes [27]. NLP can detect healthcare provider biases by identifying stigmatizing language in clinical notes [28]. For instance, recent studies found that notes about Black patients had up to 50% higher odds of containing stigmatizing language than those about White patients [29].

Given the high prevalence of stigmatizing language, what it represents in relation to bias, and its negative impact on patient care, this study proposes an NLP-driven ENGAGE (rEduce stigmatiziNG languAGE) system to automatically identify and address bias and replace stigmatizing language in clinical notes.

Methods

Ethical Considerations

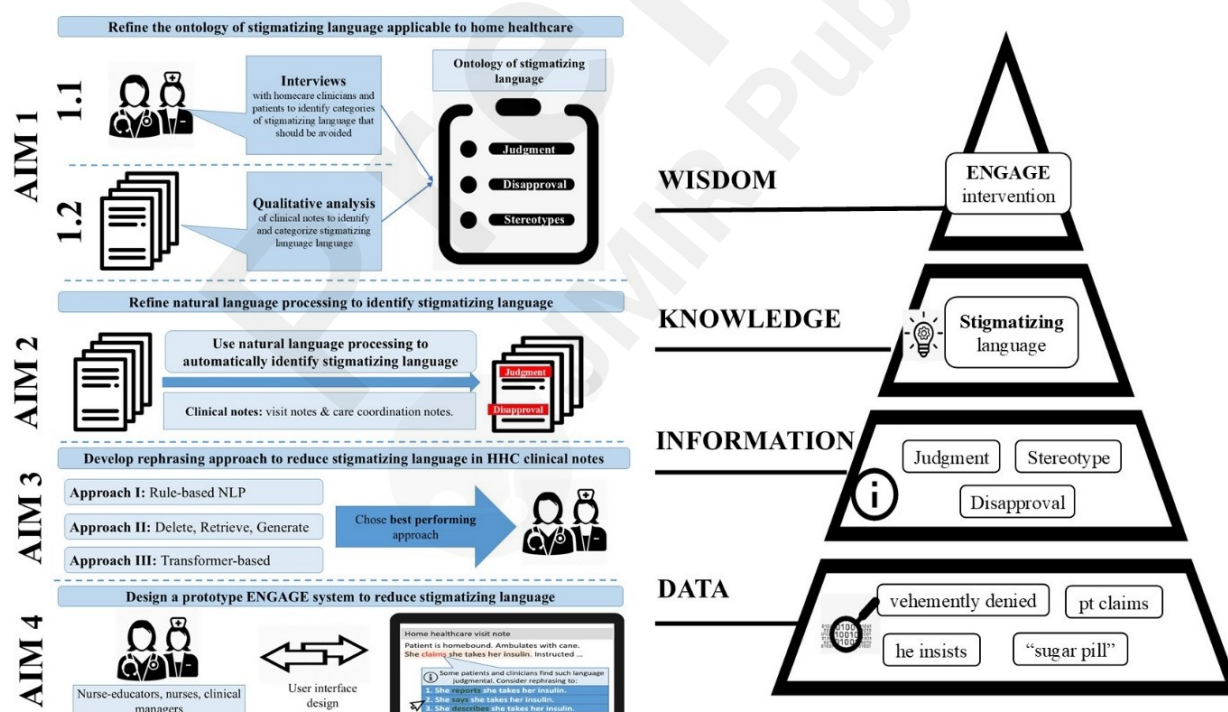
This research was approved by Columbia University's Institutional Review Board (IRB) Protocol number AAAU7957. Recruitment of participants began in May 2024. Eligible and willing

participants will provide informed consent via Qualtrics or verbal consent (patient only). Participants can choose to opt out at any point throughout the study. All data is de-identified. The participants will be compensated a fair amount of \$50 in an Amazon gift card for a one time interview. Those returning for a second interview (nurses only), will receive a \$100 Amazon gift card. Compensation transparency is ensured through recruitment flyers and consent forms, both of which contain the type of compensation and the amount.

Study Design

Four corresponding study aims were proposed that followed a mixed-method study design to achieve study goals (**Figure 1**). In Aim 1, we will adapt the ontology of stigmatizing language for HHC via interviews with patients and nurses and qualitative analysis of clinical notes. In Aim 2, we will develop and compare several NLP approaches to automatically identify stigmatizing language in clinical notes. In Aim 3, we will compare the prevalence of stigmatizing language by patients' race and ethnicity. In Aim 4, we will develop an NLP-driven ENGAGE system to reduce stigmatizing language in clinical notes. The four-aim study design is informed by the Data-Information-Knowledge-Wisdom (DIKW) conceptual framework [30,31]. DIKW suggests that discrete data points generate meaningful information that can be turned into knowledge. Wisdom is the appropriate use of knowledge to manage and solve problems. We will identify stigmatizing language from interview and clinical notes (data, aim 1-2), categorize it (information, aim 1-2), analyze its associations (knowledge, aim 3), and apply findings to develop the intervention (wisdom, aim 4) (**Figure 1**). This study aligns with the National Institute on Minority Health and Health Disparities' framework, focusing on healthcare and interpersonal/individual levels in HHC nursing [32].

Figure 1: ENGAGE Study Design



Study setting and data sources

This study will be conducted within two diverse, large HHC organizations: one is a large not-for-profit HHC agency serving patients in New York City (NYC) and its surrounding suburbs. The other

is a national HHC provider network with more than 300 HHC agencies in more than 30 states in the US. EHR data will be extracted for patients admitted to HHC between 1/1/2019-12/31/2021 (3 years). We expect to include approximately half a million unique patients. The expected patient demographics will be 67% White (non-Hispanic), 15% Black (non-Hispanic), 11% Hispanic, 2.5% Asian, and 4.5% others (including American Indian, Alaska Native, Native Hawaiian, and Pacific Islander). This should yield over 16 million clinical notes, accounting for the involvement of over 10,000 HHC nurses.

Study variables will be extracted from the EHR (**Table 1**): (1) The Outcome and Assessment Information Set (OASIS). OASIS is a comprehensive, Centers for Medicare and Medicaid Services-mandated standardized assessment tool designed to collect nearly 100 items related to a recipient's functional status, clinical status, and service needs during an HHC episode [33]. (2) Administrative data: human resources data will be used to extract HHC nurse characteristics. (3) Clinical notes: The stigmatizing languages and their classifications will be identified at the clinical notes level in aim 2.

Table 1: Study variables and sources.

Variable categories	Variables	Data source
<i>Socio-demographics</i>	Age at start of care, race, ethnicity, sex, geographic location	OASIS
<i>Physiological measures</i>	Height, weight, body mass index (BMI)	OASIS
<i>Functional status</i>	Activities of daily living, disability	OASIS
<i>Cognitive status</i>	Cognitive impairment	OASIS
<i>Clinical information</i>	HIV/AIDS diagnosis, comorbid conditions	OASIS
<i>Presence of stigmatizing language</i>	Categories of stigmatizing language extracted via NLP	Clinical notes

Participant interviews and recruitment

We will conduct interviews with a sufficient maximum variation sample of 35 HHC nurses and 35 patients in aim 1 [34]. For nurses, we aim to create a diverse sample stratified by race/ethnicity, years of experience (less than five years vs. five or more), and geographic location [35-37]. HHC Nurses will be enrolled if they are currently being employed by the participating HHC organizations. About 15 of those nurses will serve as content experts and partake in an additional interview in aim 4. For patients, we will get a sufficient maximum variation sample stratified by race/ethnicity, sex, and geographic location. Patients will be included if they are aged 18 years or older and recently admitted or discharged from HHC within the last 3 years.

Nurses will be recruited through email advertisements and presentations at nursing team meetings. Patients will be recruited through direct outreach (phone calls) based on records of recently treated and discharged patients. Participants will receive a \$50 Amazon e-gift card as a token of appreciation. Nurses serving as content experts will receive an additional \$100 Amazon e-gift card as a token of appreciation.

Study procedures

Aim 1. Identify the ontology of stigmatizing language

Based on our previous work [38-41], the ontology of stigmatizing language will be refined and expanded from the interviews with HHC patients and nurses and a review of clinical notes.

Participant interview

To generate a sufficient maximum variation sample, we plan on conducting about 30-35 interviews

for HHC nurses and patients separately. Interviews will continue until data saturation is reached, lasting up to 2 hours and conducted by phone, Zoom, or in-person (for patients only). An interview guide with semi-structured and open-ended questions will be used. These guides were developed to facilitate discussions with patients and nurses, aiming to refine the categories of stigmatizing language to avoid. Example questions include: “Have you noticed negative, discriminatory, or stigmatizing language used by HHC clinicians? Please explain.” And “We found expressions like 'claims smoking cessation, but ashtray still noted on nightstand' in the HHC clinical notes. Would you consider this judgmental or offensive? Should this language be changed or eliminated?”. All interviews will be audio-recorded for analysis.

Clinical notes

A subset of clinical note samples will be selected to identify stigma language in clinical notes. We aim for a maximum variation sample of clinical notes. This sample will be stratified by a) geographically diverse HHC agencies (Northeast, Midwest, Southern US, and NYC boroughs), b) urban/suburban/rural areas, c) diverse patient populations, and d) HHC nurses of varying sex, race and ethnicity, years of experience, educational levels, and geographical locations. Based on our pilot work [38-41], we estimate that 10-20% of clinical notes will contain stigmatizing language. To capture linguistic patterns fully, we will initially analyze 10,000 clinical notes, with additional batches of 2,500 notes if necessary to reach knowledge saturation [42]. We will annotate each clinical note for stigmatizing language and its categories (e.g., “Stereotyping by race or social class” or “Portraying the patient as difficult”) [13]. Using a hybrid qualitative approach of inductive and deductive coding [43], we will begin with the five categories from our previous study and refine or add categories as needed through discussions with annotators, the study team, and the Stakeholders Engagement Board (SEB, an interdisciplinary team of experts). Four annotators will review each note: two experienced HHC nurses (one White, one from a minority group), a social worker with racial bias detection expertise, and a minority patient who received HHC services. Annotations will be done using Amazon Web Services Ground Truth [44], with each annotator independently marking instances of stigmatizing language. Results will be merged and reviewed. Interrater reliability will be tracked using Kappa statistics, aiming for strong agreement (>0.8) [45].

Aim 2. Determine the optimal NLP approach for stigmatizing language identification

We will evaluate and compare three NLP approaches using AWS SageMaker. The first approach, key-term discovery with NimbleMiner [39,41,46,47], will build on previous work by creating vocabularies of synonyms and excluding irrelevant terms, followed by machine learning classification using models like Support Vector Machines (SVM) and Long Short-Term Memory Neural Networks (LSTM), with predictions reviewed until saturation. The second approach involves fine-tuning a publicly available Clinical Bidirectional Encoder Representations from Transformers (BERT) model [48,49], trained on a large set of clinical notes [50], using our HHC notes to improve language representation. The third NLP approach includes two aims: feature generation and model training. Feature generation will use techniques like one-hot encoding, TF-IDF, word embedding techniques (i.e., Skip-Gram [51], Glove [52], and FastText [53]), and dynamic (conceptualized) word embedding techniques (i.e., ELMO [54] and Clinical BERT [49]). Classifiers like XGBoost, Random Forest, SVM, and LSTM will be trained and validated in the model training to identify stigmatizing language in clinical notes. To enable this comparison, the annotated sample of ~10,000 clinical notes from aim 1 will be split into training (60%), validation (10%), and testing (30%) sets, stratified by stigmatizing language categories.

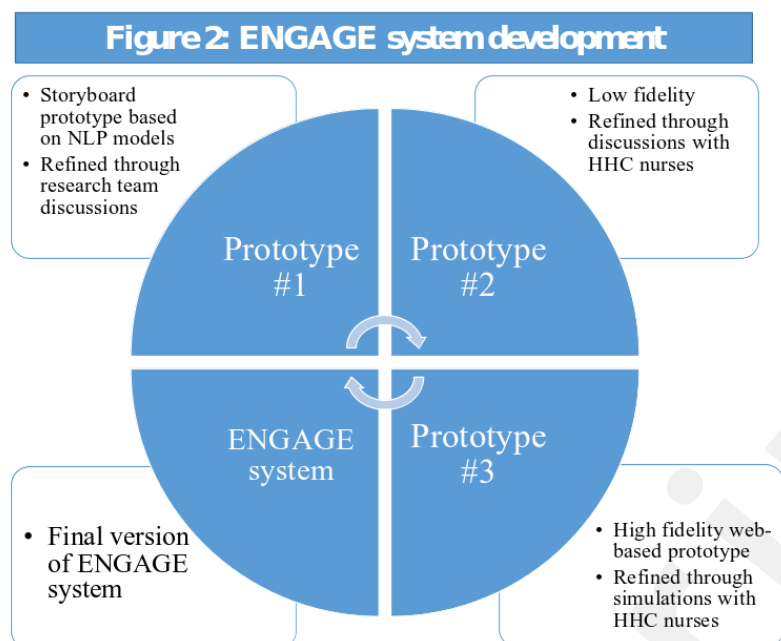
Aim 3. Compare the prevalence of stigmatizing language by patients' race and ethnicity

Based on our pilot work in hospital/ambulatory settings and HHC, we hypothesize that two or more stigmatizing language categories (e.g., questioning patient credibility [i.e., judgment], stereotyping by race or social class) will be associated with the patient's race and ethnicity. We define race and ethnicity based on the categories available in the federally mandated HHC assessment data (OASIS) [32] we will use in the study, as follows: non-Hispanic Black, Hispanic, Asian/Pacific Islander, American Indian/Alaska Native, and non-Hispanic White. The data available in federally-mandated OASIS is 99% complete (i.e., no missing data) based on our two decades of experience working with this data. Potential covariates will be identified from the data resources of OASIS and administrative data.

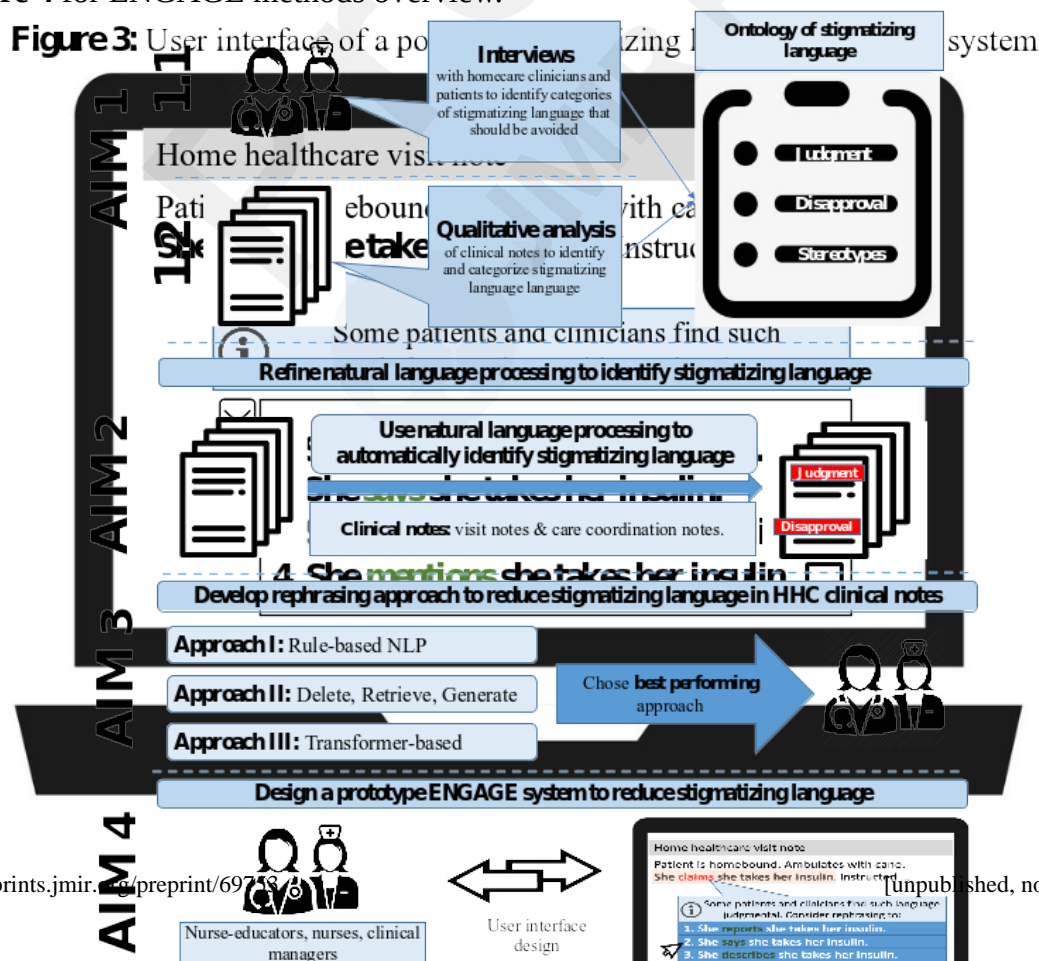
Aim 4. Develop an NLP-driven ENGAGE system to reduce stigmatizing language

To enable the development of an NLP-driven ENGAGE system, we will first identify the best method for rephrasing stigmatizing language without altering meaning. Three approaches will be compared: a rule-based method using synonym lists reviewed by annotators, a "Delete, Retrieve, Generate" method that modifies stigmatizing attributes [55], and transformer-based models (BERT [48], GPT-3 [56], T5 [57]). Two datasets will be created and used for this comparison, including a training set of 2,500 clinical notes with stigmatizing language (500 examples per category) and a test set of 4,000 sentences. Five reviewers will rewrite the sentences and reach a consensus through Delphi rounds, generating sentence pairs for training and testing. For this task, we define NLP performance as: 1) the system's ability to replace stigmatizing language with a non-stigmatizing neutral language; and 2) the system's ability not to alter the meaning of the source sentence significantly. The identified NLP approach with the best performance will help paraphrase stigmatizing language without significantly changing the original text's meaning in the NLP-driven ENGAGE system.

Next, the identified best NLP approach will be incorporated into the NLP-driven ENGAGE system. Iterative user-centered design methodologies will be used (**Figure 2**) to develop the NLP-driven ENGAGE system based on agile software development approaches [58,59] that were implemented



by our research team in numerous previous studies [60,61-64]. We will start with an initial storyboard prototype (Prototype #1) and refine it through team discussions, leading to several low-fidelity prototypes (Prototype #2). These will be reviewed with a subset of HHC nurses from Aim 1 and SEB, resulting in a high-fidelity web-based prototype (Prototype #3) built with the Shiny visualization package in R. This iterative process will continue until a final user interface (**Figure 3**) is developed, addressing key questions about screen layout, visualization of recommendations, delivery methods (pop-up, dashboard icon, message), and timing within clinician workflows. See **Figure 4** for ENGAGE methods overview.



Data analysis plan

In aim 1, interview audio data will be transcribed by the research assistant, with 20-30% validated by another study team member. Data will be analyzed using thematic analysis - a qualitative descriptive approach for identifying, analyzing, and reporting themes within data [65-71]. Qualitative analysis software (NVivo [72]) will be used to implement the analysis. The analysis includes six aims: (1) Familiarization with data by listening to recordings and reading transcripts, (2) Generating initial codes based on interview questions, (3) Data coding by two researchers with dual coding to ensure >90% agreement, (4) Collating codes into themes, (5) Defining and naming themes, and (6) Producing a final report with quotes, linking themes to the research question and literature.

In aim 2, the performance of three NLP approaches will be compared on the testing set to identify the best one for identifying stigmatizing language. For each stigmatizing language category, we will calculate the Area Under Curve of Receiver operating characteristics (AUC-ROC), AUC of precision-recall (AUC-RP), and F-score (a harmonic mean between Precision and Recall). AUC-ROC is the tradeoff between a True Positive Rate (TPR-or Sensitivity) and a False Positive Rate (FPR-or 1-Specificity). It has the advantage of being invariant to the class distribution but does not provide sufficient information about the model's precision [73]. On the other hand, AUC-RP is the tradeoff between TPR and Precision [73]. Since our goal is to maximize the Sensitivity and Precision of the NLP systems in identifying clinical notes with stigmatizing language, we will rank the performance of NLP systems using AUC-RP. We aim to achieve an F-score and AUC-ROC >0.80, which indicates a well-balanced and functioning system. If NLP approaches fail to achieve this performance level, we will conduct another cycle(s) of data annotation (with increments of 2,500 clinical notes) and NLP system fine-tuning until the desired performance is achieved.

Aim 3 will compare the prevalence of stigmatizing language by patients' race and ethnicity. The dependent variable will be the presence of stigmatizing language in the clinical note (Yes/No). Analyses will be conducted at the clinical note level, starting with bivariate assessments of potential confounders (e.g., sex, age, HIV/AIDS, BMI, comorbidities). Significant variables ($p < 0.05$) will be included in mixed-effects regression models to examine associations between race/ethnicity and stigmatizing language, accounting for clustering within patients and nurses. If stigmatizing language is infrequent, mixed-effect Poisson or negative binomial regression will be used. We will control the false discovery rate at 0.05 for multiple comparisons.

Aim 4 will evaluate the best approach for rephrasing stigmatizing language. Each NLP method will generate four paraphrased options from the test set. Five human reviewers will independently select options that replace stigmatizing language without altering the sentence's meaning. The research group will then conduct Delphi rounds to reach a consensus on the best versions. A new group of five

reviewers, including diverse HHC nurses, a minority patient, and a SEB chair, will rate each rephrased sentence on a 5-point Likert scale for effectively replacing stigmatizing language and preserving meaning. This group will first independently review each rephrased sentence and, on a 5-point Likert scale (range: 1 – almost completely, 2- to some extent, 3- unsure, 4- to a small extent, 5- did not change), indicate: 1) To what extent stigmatizing language was replaced with a non-stigmatizing neutral language; and 2) To what extent the meaning of the source sentence was altered? We will generate mean and median scores for each NLP approach and examine whether any of the NLP approaches achieved statistically significantly better performance on questions 1 (replace stigmatizing language) and 2 (not changing the meaning of the sentence) using ANOVA analyses [74].

Results

This study was funded by the National Institute on Minority Health and Health Disparities on September 24th, 2023, with a project end date of April 30th, 2027. Recruitment and enrollment began in May 2023 and is ongoing. Data curation is ongoing, with plans to start analyzing the secondary soon.

Discussion

Reducing racial biases in healthcare is a national priority. This innovative study will leverage extensive data sources to explore stigmatizing language in clinical notes, addressing critical gaps in detecting racial bias in EHR and improving system design to minimize such language used by HHC nurses. The research team includes qualified researchers who will ensure the study's implementation and timely completion.

One expected outcome of this study is the identification of an expanded ontology of stigmatizing language categories. In a previous study, five categories were identified: questioning patient credibility, expressing disapproval of patient reasoning or self-care, stereotyping by race or social class, portraying the patient as "difficult," and emphasizing clinician authority over the patient [13]. However, these categories were derived from 600 encounter notes written by 138 physicians. The current study will expand the data to a larger sample of clinical notes and interviews with patients and nurses. With this larger dataset and diverse perspectives, the expectation is to identify additional categories of stigmatizing language. These expanded and refined categories will allow for a more comprehensive analysis of stigmatizing language in clinical notes.

The prevalence of stigmatizing language is expected to vary across different racial and ethnic groups. Using data from one of the largest not-for-profit HHC agencies in the US, a pilot NLP study was conducted to examine stigmatizing language use among HHC nurses. The study found that stigmatizing language was least prevalent in the Asian group of patients. Compared to this group, the prevalence increased by 22%, 37%, and 39% in the White, Black, and Hispanic groups, respectively [12]. In the current study, using a similar population, it is expected that these racial/ethnic differences will persist. Therefore, it is crucial to develop interventions to reduce racial bias and stigmatizing language in clinical notes.

Several NLP approaches, such as rule-based methods, BERT-based models, have been evaluated for identifying stigmatizing language in clinical notes [15,28,75-77]. These approaches have their limitations. For example, rule-based approaches, which rely on predefined vocabularies, are rigid and often miss context-dependent nuances, while transformer-based models like Clinical BERT capture context better but are limited by the training data. Recent advancements in large language models, such as Mistral and LLaMA 3 [78,79], have improved performance across various NLP benchmarks. Therefore, future research can explore and compare these newer models in addressing stigmatizing language in clinical notes.

Strengths and limitations

This study has several strengths. First, this is the first study that will examine the ontology of stigmatizing language and explore the NLP approach to automatically identify and reduce stigmatizing language use in HHC. Second, the strengths of our study include the rich data resource of approximately 16.7 million clinical notes for about 667,000 unique HHC patients. This rich dataset enables a comprehensive understanding of racial bias in clinical notes. Third, an interdisciplinary team of experts in linguistics, health disparities, HHC nursing, qualitative analysis, and NLP has been assembled to design a nurse-centered NLP-based system. With strong expertise in both content and methodology, the team has carefully considered potential biases and limitations, developing a plan to address them and enhance scientific rigor.

Limitations have also been identified. First, stigmatizing language can be ambiguous and difficult to identify. To mitigate this, the interdisciplinary team of experts includes experts in racial health disparities. Additionally, data annotators will represent diverse racial perspectives. Various 'inter-rater reliability' steps will be included within the protocol to reduce the potential for such ambiguities and others that may be overlooked. Second, developing an effective ENGAGE system will pose challenges. To address this, a comprehensive, iterative development plan will be created with input from diverse HHC clinicians.

Summary

The ENGAGE study protocol addresses the critical issue of stigmatizing language in HHC through developing an NLP-driven system. This innovative mixed-method study, conducted within two large and diverse HHC organizations, aims to refine the ontology of stigmatizing language, develop and test various NLP methods for identifying and replacing such language, and examine the prevalence and impact of stigmatizing language across different racial and ethnic groups. The project will comprehensively analyze the patterns and consequences of stigmatizing language in clinical notes by leveraging extensive EHR data and employing robust statistical and machine learning techniques. With the successful development and iterative refinement of the NLP-driven ENGAGE system, which integrates the most effective NLP approach for rephrasing stigmatizing language without altering the original meaning, the final system will be ready for testing in a clinical trial.

Acknowledgments

We thank the ENGAGE study participants, the research team members, and the partnering HHC agencies.

Funding Statement

The research is funded by the National Institute on Minority Health and Health Disparities, grant number 1R01MD018028-01A1.

Data Availability

This study is in the data collection phase. Data sets will be made available at a later time.

Conflicts of Interest

None declared.

Author Contributions

Conceptualization: Scott Sittig, Jacquelyn Y. Taylor, and Maxim Topaz

Data Curation: Zhihong Zhang, Pallavi Gupta, Scott Sittig, Margaret V. McDonald, Jacquelyn Y. Taylor, and Maxim Topaz

Formal Analysis: Zhihong Zhang, Pallavi Gupta, and Maxim Topaz

Funding Acquisition: Scott Sittig, Jacquelyn Y. Taylor, and Maxim Topaz

Investigation Methodology: Scott Sittig, Jacquelyn Y. Taylor, and Maxim Topaz

Project Administration Resources: Scott Sittig, Jacquelyn Y. Taylor, and Maxim Topaz

Software: Maxim Topaz

Supervision: Scott Sittig, Jacquelyn Y. Taylor, and Maxim Topaz

Validation: Scott Sittig, Jacquelyn Y. Taylor, and Maxim Topaz

Visualization: Scott Sittig, Jacquelyn Y. Taylor, and Maxim Topaz

Writing – original draft: Zhihong Zhang, Pallavi Gupta, Stephanie Potts-Thompson, Laura Prescott, Morgan Morrison, Scott Sittig, Margaret V. McDonald, Chase Wesley Raymond, Jacquelyn Y. Taylor, Maxim Topaz

Writing – review & editing: Zhihong Zhang, Pallavi Gupta, Stephanie Potts-Thompson, Laura Prescott, Morgan Morrison, Scott Sittig, Margaret V. McDonald, Chase Wesley Raymond, Jacquelyn Y. Taylor, Maxim Topaz

Abbreviations

ANOVA: Analysis of Variance

AUC-ROC: Area Under Curve of Receiver operating characteristics

AUC-RP: Area Under Curve of precision-recall

AWS: Amazon Web Services

BERT: Bidirectional Encoder Representations from Transformers

BMI: body mass index

DIKW: Data-Information-Knowledge-Wisdom

EHR: electronic health records
ENGAGE: rEduce stigmatiziNG languAGE
GPT-3: Generative Pre-trained Transformer 3
HHC: home healthcare
HIV/AIDS: human immunodeficiency virus/acquired immunodeficiency syndrome
LLaMA: Large Language Model Meta AI
LSTMs: Long Short-Term Memory
NLP: natural language processing
NYC: New York City
OASIS: Outcome and Assessment Information Set
SEB: Stakeholders Engagement Board
SVM: Support Vector Machine
T5: Text-To-Text Transfer Transformer
TF-IDF: Term frequency-inverse document frequency
TPR: True Positive Rate
US: United States

References

1. Harris-Kojetin L, Sengupta M, Lendon J, Cefferly C. Long-Term Care Providers and Services Users in the United States, 2015-2016. Washington, DC; 2019. CDC website; Accessed January 13, 2022. <https://stacks.cdc.gov/view/cdc/76253>
2. Medicare Payment Advisory Commission. March 2019 Report to the Congress: Medicare Payment Policy. Medicare Payment Advisory Commission; 2019. <https://www.medpac.gov/document/march-2019-report-to-the-congress-medicare-payment-policy/>

3. Wei H, Sewell KA, Woody G, Rose MA. The state of the science of nurse work environments in the United States: A systematic review. *International Journal of Nursing Sciences*; 2018. (5,3): 287-300. doi:10.1016/J.IJNSS.2018.04.010
4. Recio-Saucedo A, Dall'Ora C, Maruotti A, et al. What impact does nursing care left undone have on patient outcomes? Review of the literature. *J Clin Nurs*; 2018. (27,11-12): 2248-2259. doi:10.1111/JOCN.14058
5. Wynendaele H, Willems R, Trybou J. Systematic review: Association between the patient-nurse ratio and nurse outcomes in acute care hospitals. *J Nurs Manag*; 2019. (27,5): 896-917. doi:10.1111/JONM.12764
6. Blair I V, Steiner JF, Havranek EP. Unconscious (Implicit) Bias and Health Disparities: Where Do We Go from Here? *Perm J*; 2011. (15,2): 71-78. doi:10.7812/TPP/11.979
7. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. *BMC Med Ethics*; 2017. (18,1): 1-18. doi:10.1186/S12910-017-0179-8
8. Zestcott CA, Blair I V, Stone J. Examining the presence, consequences, and reduction of implicit bias in health care: A narrative review. *Group processes & intergroup relations*; 2016. (19,4): 528-542. doi:10.1177/1368430216642029
9. Groves P, Bunch J, Sabin J. Nurse bias and nursing care disparities related to patient characteristics: A scoping review of the quantitative and qualitative evidence. *J Clin Nurs*; 2021. (30,23-24): 3385–3397. doi:10.1111/JOCN.15861
10. Hall WJ, Chapman M V, Lee KM, et al. Implicit Racial/Ethnic Bias Among Health Care Professionals and Its Influence on Health Care Outcomes: A Systematic Review. *American Journal of Public Health*; 2015; (105,12): e60-e76. doi:10.2105/AJPH.2015.302903
11. Narayan MC. Addressing Implicit Bias in Nursing: A Review. *Am J Nurs*; 2019. (119,7): 17-

18. doi:10.1097/01.NAJ.0000569340.27659.5A
12. Topaz M, Song J, Davoudi A, et al. Home Healthcare Clinicians Use Judgment Language More Frequently for Black and Hispanic Patients: A Natural Language Processing Study. *JMIR Nursing*; 2022; (6): e42552. doi:10.2196/42552
13. Park J, Saha S, Chee B, et al. Physician Use of Stigmatizing Language in Patient Medical Records. *JAMA Netw Open*; 2021. (4,7): e2117052-e2117052. doi:10.1001/jamanetworkopen.2021.17052
14. P. Goddu A, O'Connor KJ, Lanzkron S, et al. Do Words Matter? Stigmatizing Language and the Transmission of Bias in the Medical Record. *J Gen Intern Med*; 2018. (33,5): 685-691. doi:10.1007/S11606-017-4289-2
15. Barcelona, V, Scharp, D, Moen, H, et al. Using Natural Language Processing to Identify Stigmatizing Language in Labor and Birth Clinical Notes. *Maternal and Child Health Journal*; 2024. (28,3): 578–586. doi:10.1007/s10995-023-03857-4
16. Andreyeva E, David G, Song H. The Effects of Home Health Visit Length on Hospital Readmission. *NBER Working Paper Series*; 2018. <http://www.nber.org/papers/w24566>
17. Song H, Andreyeva E, David G. Time Is the Wisest Counselor of All: The Value of Provider–Patient Engagement Length in Home Healthcare. *Manag Sci*; 2021. (68,1): 420-4441. doi:10.1287/MNSC.2020.3921
18. Smith LM, Anderson WL, Kenyon A, et al. Racial and Ethnic Disparities in Patients' Experience With Skilled Home Health Care Services. *Med Care Res Rev*; 2015. (72,6): 756-774. doi:10.1177/1077558715597398
19. Wang J, Yu F, Cai X, et al. Functional outcome in home health: Do racial and ethnic minority patients with dementia fare worse? *PLoS One*; 2020. (15,5): e0233650. doi:10.1371/journal.pone.0233650

20. Lipphardt A, Leen-Thomele E, Coroian E, et al. Understanding racial and ethnic disparities in home health care: practice and policy factors. *Innov Aging*; 2017. (1,suppl_1): 956. <https://doi.org/10.1093/geroni/igx004.3442>
21. Narayan MC, Scafide KN. Systematic Review of Racial/Ethnic Outcome Disparities in Home Health Care. *J Transcult Nurs*; 2017. (28,6): 598-607. doi:10.1177/1043659617700710
22. Squires A, Ma C, Miner S, et al. Assessing the influence of patient language preference on 30 day hospital readmission risk from home health care: A retrospective analysis. *Int J Nurs Stud*; 2022. (125): 104093. doi:10.1016/j.ijnurstu.2021.104093
23. Fernández L, Fossa A, Dong Z, et al. Words Matter: What Do Patients Find Judgmental or Offensive in Outpatient Notes? *J Gen Intern Med*; 2021. (36,9): 2571-2578. doi:10.1007/S11606-020-06432-7/tables/3
24. Department of Health and Human Services. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program; 2020. Accessed December 29, 2021. https://www.healthit.gov/sites/default/files/cures/2020-03/ONC_Cures_Act_Final_Rule_03092020.pdf
25. Alvarado, C S, Zook, K, Henry, K. Electronic Health Record Adoption and Interoperability among U.S. Skilled Nursing Facilities in 2016. US Office of National Coordinator for Health IT; 2017. Accessed December 29, 2021. <https://www.healthit.gov/sites/default/files/electronic-health-record-adoption-and-interoperability-among-u.s.-skilled-nursing-facilities-in-2016.pdf>
26. Matheny ME, Whicher D, Thadaney Israni S. Artificial Intelligence in Health Care: A Report from the National Academy of Medicine. *JAMA*; 2020. (323,6): 509-510.

doi:10.1001/Jama.2019.21579

27. Demner-Fushman D, Chapman WW, McDonald CJ. What can natural language processing do for clinical decision support? *J Biomed Inform*; 2009. (42,5): 760-772. doi:10.1016/j.jbi.2009.08.007
28. Barcelona, V, Scharp, D, Idnay, B R, et al. Identifying stigmatizing language in clinical documentation: A scoping review of emerging literature. *PloS One*; 2024. (19,6): e0303653. doi:10.1371/journal.pone.0303653
29. Beach MC, Saha S, Park J, et al. Testimonial Injustice: Linguistic Bias in the Medical Records of Black Patients and Women. *J Gen Intern Med*; 2021. (36,6): 1708-1714. doi:10.1007/S11606-021-06682-Z
30. Ronquillo C, Currie LM, Rodney P. The Evolution of Data-Information-Knowledge-Wisdom in Nursing Informatics. *ANS Adv Nurs Sci*; 2016. (39,1): E1-18. doi:10.1097/ANS.000000000000107
31. Matney S, Brewster PJ, Sward KA, Cloyes KG, Staggers N. Philosophical approaches to the nursing informatics data-information-knowledge-wisdom framework. *ANS Adv Nurs Sci*; 2011. (34,1): 6-18. doi:10.1097/ANS.0b013e3182071813
32. National Institute on Minority Health and Health Disparities. NIMHD Minority Health and Health Disparities Research Framework. NIMHD Website; 2020. Accessed January 24, 2022. <https://www.nimhd.nih.gov/about/overview/research-framework/>
33. O'Connor M, Davitt JK. The Outcome and Assessment Information Set (OASIS): a review of validity and reliability. *Home Health Care Serv Q*; 2012. (31,4): 267-301. doi:10.1080/01621424.2012.703908
34. Suri H. Purposeful sampling in qualitative research synthesis. *Qual Res J*; 2011. (11,2): 63-75. doi:10.3316/QRJ1102063/FULL/XML

35. Hobensack M, Ojo M, Bowles K, et al. Home Healthcare Clinicians' Perspectives on Electronic Health Records: A Qualitative Study. *Stud Health Technol Inform*; 2021. (284): 426-430. doi:10.3233/SHTI210763
36. Topaz M, Naylor MD, Holmes JH, et al. Factors Affecting Patient Prioritization Decisions at Admission to Home Healthcare: A Predictive Study to Develop a Risk Screening Tool. *CIN – Comput Informatics Nurs*; 2020. (38,2): 88-98. doi:10.1097/CIN.0000000000000576
37. Bowles KH, Ratcliffe S, Potashnik S, et al. Using Electronic Case Summaries to Elicit Multi-Disciplinary Expert Knowledge about Referrals to Post-Acute Care. *Appl Clin Inform*; 2016. (7,2): 368-379. doi:10.4338/ACI-2015-11-RA-0161
38. McDonald M V, Feldman PH, Barrón-Vayá Y, et al. Outcomes of clinical decision support (CDS) and correlates of CDS use for home care patients with high medication regimen complexity: a randomized trial. *J Eval Clin Pract*; 2016. (22,1): 10-19. doi:10.1111/jep.12383
39. Topaz M, Adams V, Wilson P, et al. Free-Text Documentation of Dementia Symptoms in Home Healthcare: A Natural Language Processing Study. *Gerontol Geriatr Med*; 2020. (6): 1-10. doi:10.1177/2333721420959861
40. Dreisbach C, Koleck TA, Bourne PE, et al. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *Int J Med Inform*; 2019. (125): 37-46. doi:10.1016/j.ijmedinf.2019.02.008
41. Topaz M, Murga L, Bar-Bachar O, et al. NimbleMiner: An Open-Source Nursing- Sensitive Natural Language Processing System Based on Word Embedding. *CIN Comput Informatics, Nurs*; 2019. (37,11): 583-590. doi:10.1097/CIN.0000000000000557
42. Topaz M, Murga L, Gaddis KM, et al. Mining fall-related information in clinical notes:

- comparison of rule-based and novel word embedding-based machine learning approaches. *J Biomed Inform*; 2019. (90,103103): 1-8. doi:10.1016/J.JBI.2019.103103
43. Fereday J, Muir-Cochrane E. Demonstrating Rigor Using Thematic Analysis: A Hybrid Approach of Inductive and Deductive Coding and Theme Development. *Int J Qual Methods*; 2006. (5,1): 80-92. doi:10.1177/160940690600500107
44. Amazon Web Services. Training data labeling using humans with Amazon SageMaker Ground Truth. Amazon SageMaker. Accessed December 9, 2021. <https://docs.aws.amazon.com/sagemaker/latest/dg/sms.html>
45. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*; 2012. (22,3): 276-82. PMID: 23092060
46. Topaz M, Koleck TA, Onorato N, et al. Nursing documentation of symptoms is associated with higher risk of emergency department visits and hospitalizations in homecare patients. *Nurs Outlook*; 2020. (69,3): 435-446. doi:10.1016/j.outlook.2020.12.007
47. Blumenthal KG, Topaz M, Zhou L, et al. Mining social media data to assess the risk of skin and soft tissue infections from allergen immunotherapy. *J Allergy Clin Immunol*; 2019. (144,1): 129-134. doi:10.1016/J.JACI.2019.01.029
48. Devlin J, Chang M-W, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv*; 2018. <https://doi.org/10.48550/arXiv.1810.04805>
49. Alsentzer E, Murphy JR, Boag W, et al. Publicly available clinical BERT embeddings. *arXiv*; 2019. <https://doi.org/10.48550/arXiv.1904.03323>
50. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*; 2016. (3): 160035. doi:10.1038/sdata.2016.35

51. Guthrie D, Allison B, Liu W, et al. A closer look at skip-gram modeling. In: Proceedings of the Fifth International Conference on Language Resources and Evaluation. ELRA; 2006. (6):1222-1225. http://www.lrec-conf.org/proceedings/lrec2006/pdf/357_pdf.pdf
52. Pennington J, Socher R, Manning CD. Glove: Global vectors for word representation. In: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing; 2014.1532-1543. doi:10.3115/v1/D14-1162
53. Joulin A, Grave E, Bojanowski P, Mikolov T. Bag of tricks for efficient text classification. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers. arXiv; 2016. 427–431. <https://aclanthology.org/E17-2068>
54. Sarzynska-Wawer J, Wawer A, Pawlak A, et al. Detecting formal thought disorder by deep contextualized word representations. Psychiatry Res; 2021. (304):114135. doi:10.1016/j.psychres.2021.114135
55. Li J, Jia R, He H, Liang P. Delete, Retrieve, Generate: a Simple Approach to Sentiment and Style Transfer. NAACL HLT 2018 - 2018 Conf North Am Chapter Assoc Comput Linguist Hum Lang Technol -Proc Conf; 2018. (1): 1865-1874. doi:10.18653/V1/N18-1169
56. Brown T, Mann B, Ryder N, et al. Language Models are Few-Shot Learners. Adv Neural Inf Process Syst; 2020. (33): 1877-1901. <https://doi.org/10.48550/arXiv.2005.14165>
57. Raffel C, Shazeer N, Roberts A, Lee K, Narang S. Exploring the limits of transfer learning with a unified text-to-text transformer. J Mach Learn Res; 2020. (79,5): 184-212. <https://doi.org/10.48550/arXiv.1910.10683>
58. Helander M G. Handbook of Human-Computer Interaction. New York, NY: Elsevier; 2014. ISBN:1483295133, 9781483295138

59. Ferreira J, Noble J, Biddle R. Agile Development Iterations and UI Design. In: AGILE. IEEE; 2007. 50-58. doi:10.1109/AGILE.2007.8
60. Sittig S, Wang J, Iyengar S, et al. Incorporating Behavioral Trigger Messages Into a Mobile Health App for Chronic Disease Management: Randomized Clinical Feasibility Trial in Diabetes. JMIR mHealth uHealth; 2020. (8,3): e15927. doi:10.2196/15927
61. Topaz M, Trifilio M, Maloney D, et al. Improving patient prioritization during hospital-homecare transition: A pilot study of a clinical decision support tool. Res Nurs Health; 2018. (41,5): 440-447. doi:10.1002/nur.21907
62. Bowles KH, Holmes JH, Ratcliffe SJ, et al. Factors identified by experts to support decision making for post acute referral. Nurs Res; 2009. (58,2): 115-122. doi:10.1097/NNR.0b013e318199b52a
63. Topaz M, Rao A, Masterson Creber R, Bowles KH. Educating clinicians on new elements incorporated into the electronic health record: theories, evidence, and one educational project. Comput Inform Nurs; 2013. (31,8): 375-379. doi:10.1097/NXN.0b013e318295e5a5
64. Bowles KH, Chittams J, Heil E, et al. Successful electronic implementation of discharge referral decision support has a positive impact on 30- and 60-day readmissions. Res Nurs Health; 2015. (38,2): 102-114. doi:10.1002/nur.21643
65. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol; 2006. (3,2): 77-101. doi:10.1191/1478088706qp063oa
66. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. Nurs Health Sci; 2013. (15,3): 398-405. doi:10.1111/nhs.12048
67. Riegel B, Dickson VV, Topaz M. Qualitative analysis of naturalistic decision making in

- adults with chronic heart failure. *Nurs Res*; 2013. (62,2): 91-98. PMID: 23154659.
68. Topaz M, Bar-Bachar O, Admi H, et al. Patient-centered care via health information technology: a qualitative study with experts from Israel and the U.S. *Informatics Heal Soc Care*; 2019. (45,3): 1-12. doi:10.1080/17538157.2019.1582055
69. Koru G, Alhuwail D, Topaz M, et al. Investigating the Challenges and Opportunities in Home Care to Facilitate Effective Information Technology Adoption. *J Am Med Dir Assoc*; 2016. (17,1): 53-58. doi:10.1016/j.jamda.2015.10.008
70. Topaz M, Seger DL, Goss F, et al. Standard information models for representing adverse sensitivity information in clinical documents. *Methods Inf Med*; 2016. (55,2): 151-157. doi:10.3414/ME15-01-0081
71. Lee J, Callon W, Haywood, Jr. C, Lanzkron SM, Gulbrandsen P, Beach MC. What does shared decision making look like in natural settings? A mixed methods study of patient-provider conversations. *Commun Med*; 2018. (14,3): 217-228. doi:10.1558/cam.32815
72. QSR International Pty Ltd. NVivo qualitative data analysis Software. Published 2018. <http://qsrinternational.com/about-us/newsroom/qsr-international-launches-qualitative-data-analys>
73. Siblini W, Fréry J, He-Guelton L, et al. Master your metrics with calibration. *arXiv*; 2019. (1): 1-14. <https://doi.org/10.48550/arXiv.1909.02827>
74. Kaufmann J, Schering A. Analysis of Variance ANOVA. *Wiley StatsRef Stat Ref: Statistics Reference Online*; 2014. doi:10.1002/9781118445112.STAT06938.
75. Harrigian, K, Ayah, Z, Brant, C, et al. Characterization of stigmatizing language in medical records. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics*; 2023. (2): 312-329. <https://aclanthology.org/2023.acl-short.28.pdf>

76. Weiner SG, Lo YC, Carroll AD, et al. The Incidence and Disparities in Use of Stigmatizing Language in Clinical Notes for Patients With Substance Use Disorder. *J Addict Med*; 2023. (17,4): 424-430. doi:10.1097/ADM.0000000000001145
77. Himmelstein G, Bates D, Zhou L. Examination of Stigmatizing Language in the Electronic Health Record. *JAMA Netw Open*; 2022. (5,1): e2144967. doi:10.1001/jamanetworkopen.2021.44967
78. Jiang, A Q, Sablayrolles, A, Mensch, A, et al. Mistral 7B. arXiv; 2023. (2): 1-9. <https://doi.org/10.48550/arXiv.2310.06825>
79. Dubey, A, Jauhri, A, Pandey, A, et al. The llama 3 herd of models. arXiv; 2023. (1): 1-92. <https://doi.org/10.48550/arXiv.2407.21783>

Supplementary Files

Figures

Multimedia Appendixes

Summary statement.

URL: <http://asset.jmir.pub/assets/ff991793effaf4c213f57e7a77333d40.pdf>

