

Protocol

Identifying and Reducing Stigmatizing Language in Home Health Care With a Natural Language Processing–Based System (ENGAGE): Protocol for a Mixed Methods Study

Zhihong Zhang^{1,2}, RN, PhD; Pallavi Gupta², MS, PhD; Stephanie Potts-Thompson^{2,3}, MPH; Laura Prescott^{2,3}, MA; Morgan Morrison^{2,3}, MA; Scott Sittig⁴, RHIA, MHI, PhD; Margaret V McDonald⁵, MSW; Chase Raymond^{6,7}, PhD; Jacquelyn Y Taylor³, RN, PNP-BC, PhD; Maxim Topaz^{1,2}, RN, MA, PhD

¹Columbia University Data Science Institute, New York, NY, United States

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

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

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

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

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

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

Corresponding Author:

Maxim Topaz, RN, MA, PhD

Columbia University School of Nursing

560 West 168th Street

New York, NY, 10032

United States

Phone: 1 212 305 3976

Email: mt3315@cumc.columbia.edu

Abstract

Background: Stigmatizing language is common in clinical notes and can adversely affect the quality of patient care. Natural language processing (NLP) is a promising technology for identifying such language across large volumes of clinical notes in electronic health records.

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

Methods: Using a mixed methods study, we will extract electronic health record data for patients admitted to 2 large, diverse home health care (HHC) organizations between January 2019 and December 2021. We propose the following 4 aims: aim 1 is to refine the ontology of stigmatizing language in HHC by (1) interviewing a diverse sample of HHC nurses and patients to identify terms to avoid and (2) analyzing clinical notes from various regions in the United States to categorize stigmatizing language. Aim 2 is to determine the best NLP approach for accurately identifying stigmatizing language by training algorithms and comparing their performance to human annotations. Aim 3 is to 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 minority groups). Aim 4 is to develop the NLP-driven ENGAGE system by (1) testing NLP methods (rule based; "delete, retrieve, and generate"; and transformers) for suggesting alternative wording and (2) designing and refining the user interface for clinical trial preparation.

Results: We received funding from the National Institute on Minority Health and Health Disparities in September 2023. Recruitment began in May 2024, and as of March 2025, interviews have been completed for 9 enrolled participants. We anticipate completing all study aims by April 2027.

Conclusions: This study will leverage extensive data sources to examine stigmatizing language in HHC settings and contribute to the development of systems aimed at effectively reducing the use of such language among HHC nurses.

International Registered Report Identifier (IRRID): DERR1-10.2196/69753

(*JMIR Res Protoc* 2025;14:e69753) doi: [10.2196/69753](https://doi.org/10.2196/69753)

KEYWORDS

natural language processing; racial bias; stigmatizing language; home health care; artificial intelligence; AI

Introduction

Home health care (HHC) is one of the fastest-growing outpatient settings in the United States, 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 (eg, structural resources, levels of education, and patient-per-nurse ratios) [3-5], nurses' biases toward 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 the language used in electronic health record (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 than those whose records were without such language (average visit length 46 vs 70 minutes, respectively) [16,17]. This is concerning because shorter HHC visits are associated with poor outcomes (eg, 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 because, as of April 2021, health care 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 health care [26]. One promising technology is natural language processing (NLP), a branch of artificial intelligence that analyzes free-text data, such as clinical notes in EHRs, to extract meaningful insights [27]. NLP has been used to detect health care provider biases by identifying stigmatizing language in clinical notes [28]. Common NLP techniques include rule-based methods such as NimbleMiner [29], which identify specific stigmatizing terms [30,31], as well as machine learning and deep learning models such as the clinical bidirectional encoder representations from transformers (BERT) model, which can detect more subtle, context-dependent patterns in the language [32]. In our previous research, we applied NimbleMiner to identify stigmatizing language in clinical notes and found that 38% of notes contained such language. In addition, notes about Black patients had up to 50%

higher odds of containing stigmatizing language than those about White patients [30].

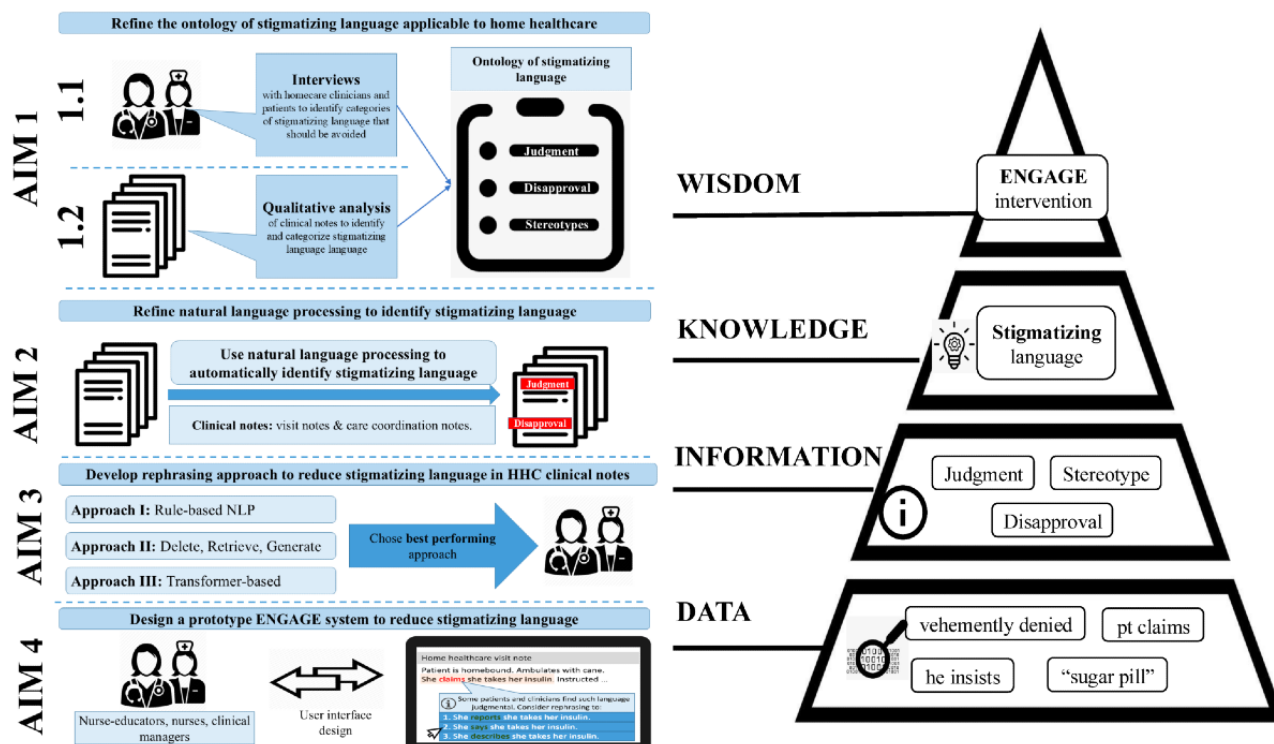
Given the high prevalence of stigmatizing language, what it represents in relation to bias, and its negative impact on patient care, this study proposes an NLP-driven reduce stigmatizing language (ENGAGE) system to automatically identify and address bias and replace stigmatizing language in clinical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

Methods**Ethical Considerations**

This research was approved by Columbia University's institutional review board (AAAU7957). Eligible and willing participants will provide informed consent via Qualtrics (Qualtrics International Inc) or verbal consent (patient only). Participants can choose to opt out at any point throughout the study. All data will be deidentified to protect participant privacy. The participants will be compensated US \$50 in the form of an Amazon electronic gift card for a one-time interview. Nurses serving as content experts will return for a second interview and will receive a US \$100 Amazon electronic gift card. Compensation transparency is ensured through recruitment flyers and consent forms, both of which contain the compensation type and amount.

Study Design

We proposed 4 corresponding study aims that followed a mixed methods study design to achieve study goals (Figure 1). This protocol is reported in accordance with the Standards for Reporting Implementation Studies statement [33], except for the Results section, which is reported according to JMIR's requirements for protocols. 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 the use of stigmatizing language in clinical notes. This 4-aim study design is informed by the data, information, knowledge, and wisdom conceptual framework [34,35]. This framework 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 interviews and clinical notes (data: aims 1 and 2), categorize it (information: aims 1 and 2), analyze its associations (knowledge: aim 3), and apply the 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 health care and interpersonal and individual levels in HHC nursing [36].

Figure 1. Reduce stigmatizing language (ENGAGE) study design. HHC: home health care; NLP: natural language processing.

Study Setting and Data Sources

This study will be conducted within 2 diverse, large HHC organizations: one is a large not-for-profit HHC agency serving patients in New York City 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 United States. EHR data will be extracted for patients admitted to HHC between January 1, 2019, and December 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% other (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.

The following 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 & 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 [37]—(2) administrative data: human resources data will be used to extract HHC nurse characteristics; and (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
Sociodemographics	Age at start of care, race, ethnicity, sex, and geographic location	OASIS ^a
Physiological measures	Height, weight, and BMI	OASIS
Functional status	Activities of daily living and disability	OASIS
Cognitive status	Cognitive impairment	OASIS
Clinical information	Diagnosis and comorbid conditions	OASIS
Presence of stigmatizing language	Categories of stigmatizing language extracted via NLP ^b	Clinical notes

^aOASIS: Outcome and Assessment Information Set.

^bNLP: natural language processing.

Participant Interviews and Recruitment

We will conduct interviews with a sufficient maximum variation sample of 35 HHC nurses and 35 patients in aim 1 [38]. For nurses, we aim to create a diverse sample stratified by race and

ethnicity, years of experience (<5 years vs ≥5 years), and geographic location [39–41]. HHC nurses will be enrolled if they are currently being employed by the participating HHC organizations. About 15 (43%) 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 and ethnicity, sex, and geographic location. Patients will be included if they are aged ≥ 18 years and recently admitted to 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 (ie, phone calls) based on records of recently treated and discharged patients.

Study Procedures

Aim 1: Identify the Ontology of Stigmatizing Language

On the basis of our previous work [29,42-44], 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 to 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 (Zoom Video Communications, Inc), or in person (for patients only). An interview guide with semistructured and open-ended questions will be used. These guides were developed to facilitate discussions with nurses (Multimedia Appendix 1) and patients (Multimedia Appendix 2), aiming to refine the categories of stigmatizing language to avoid. Example questions include the following: “Have you noticed negative, discriminatory, or stigmatizing language used by HHC clinicians? Please explain.” “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 will aim for a maximum variation sample of clinical notes. This sample will be stratified by (1) geographically diverse HHC agencies (Northeastern, Midwestern, and Southern United States and the New York City boroughs); (2) urban, suburban, and rural areas; (3) diverse patient populations; and (4) HHC nurses of varying sex, race, and ethnicity, years of experience, educational levels, and geographic locations. On the basis of our pilot work [29,42-44], we estimate that 10% to 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 2500 notes if necessary to reach knowledge saturation [45]. We will annotate each clinical note for stigmatizing language and its categories (eg, “Stereotyping by race or social class” or “Portraying the patient as difficult”) [13]. Using a hybrid, qualitative approach of inductive and deductive coding [46], we will begin with the 5 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: 2 experienced HHC

nurses (1 White and 1 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 SageMaker Ground Truth [47], with each annotator independently marking instances of stigmatizing language. Results will be merged and reviewed. Interrater reliability will be tracked using κ statistics, aiming for strong agreement (>0.8) [48].

Aim 2: Determine the Optimal NLP Approach for Stigmatizing Language Identification

We will evaluate and compare 3 NLP approaches using SageMaker Ground Truth. The first approach, key-term discovery with NimbleMiner [29,43,49,50], will build on previous work by creating vocabularies of synonyms and excluding irrelevant terms, followed by machine learning classification using models such as XGBoost, random forest, support vector machines, and long short-term memory neural networks, with predictions reviewed until saturation. The second approach involves fine-tuning a publicly available clinical BERT model [51,52] trained on a large set of clinical notes [53] using our HHC notes to improve language representation. The third NLP approach includes 2 aims: feature generation and model training. Feature generation will use techniques such as one-hot encoding; term frequency-inverse document frequency; word embedding techniques (ie, Skip-Gram [54], Glove [55], and FastText [56]); and dynamic (conceptualized) word embedding techniques (ie, ELMO [57] and clinical BERT [52]). During model training, machine learning classifiers will be trained and validated in the model training to identify stigmatizing language in clinical notes. To enable this comparison, the annotated sample of approximately 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

On the basis of our pilot work in hospital and ambulatory settings and HHC, we hypothesize that 2 or more stigmatizing language categories (eg, questioning patient credibility [ie, 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) [36] that we will use in the study, as follows: non-Hispanic Black, Hispanic, Asian or Pacific Islander, American Indian or Alaska Native, and non-Hispanic White. The data available in federally mandated OASIS are 99% complete (ie, no missing data) based on our 2 decades of experience working with these data. Potential covariates will be identified from the data resources of OASIS and administrative data.

Aim 4: Develop an NLP-Driven ENGAGE System

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, and generate” method that modifies stigmatizing attributes [58]; and transformer-based

models (eg, BERT [51], generative pretrained transformers 3 [59], and text-to-text transfer transformer [60]). Two datasets will be created and used for this comparison, including a training set of 2500 clinical notes with stigmatizing language (500 examples per category) and a test set of 4000 sentences. In total, 5 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 nonstigmatizing 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 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 [61,62] that were implemented by our research team in numerous previous studies [63-67]. 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 (R Foundation for Statistical Computing). 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, and message); and timing within clinician workflows.

Figure 2. Iterative development process of the reduce stigmatizing language (ENGAGE) system. HHC: home health care; NLP: natural language processing.

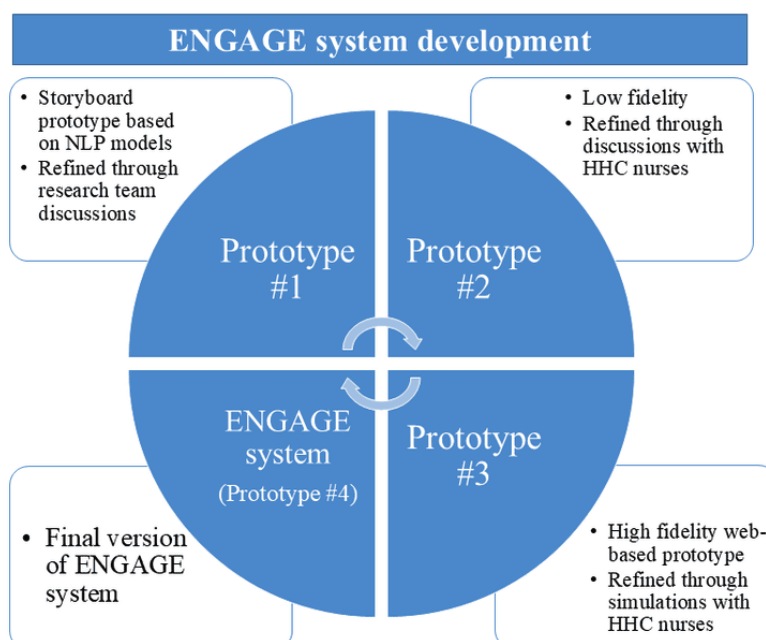
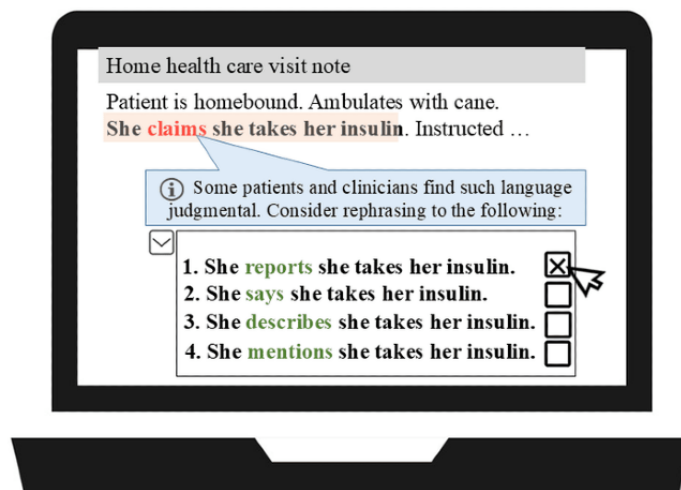


Figure 3. User interface of a potential stigmatizing language reduction system.



Data Analysis Plan

In aim 1, interview audio data will be transcribed by the research assistant, with 20% to 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 the data [68-74]. Qualitative analysis software (NVivo [75]; Lumivero) 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 2 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 the literature.

In aim 2, the performance of 3 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 the receiver operating characteristic curve (AUC-ROC), area under the curve of precision recall (AUC-RP), and the *F* score (a harmonic mean between precision and recall). AUC-ROC is the tradeoff between a true positive rate (sensitivity) and a false positive rate (1-specificity). It has the advantage of being invariant to the class distribution but does not provide sufficient information about the model’s precision [76]. On the other hand, AUC-RP is the tradeoff between recall (true positive rate) and precision [76]. Because 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 or cycles of data annotation (with increments of 2500 clinical notes) and NLP system fine-tuning until the desired performance is achieved.

In aim 3, we 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 and no). Analyses will be conducted at the clinical note level, starting with bivariate assessments of potential confounders (eg, sex, age, disease diagnosis, and comorbidities). Significant variables (*P*<.05) will be included in mixed-effects regression models to examine associations between race and ethnicity and stigmatizing language, accounting for clustering within patients and nurses. If stigmatizing language is infrequent, mixed-effects Poisson or negative binomial regression will be used. We will control the false discovery rate at 0.05 for multiple comparisons.

In aim 4, we will evaluate the best approach for rephrasing stigmatizing language. Each NLP method will generate 4 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 5 reviewers, including diverse HHC nurses, a minority patient, and an 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, and 5=did not change), indicate (1) to what extent stigmatizing language was replaced with a neutral, nonstigmatizing 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 [77].

Results

We received funding from the National Institute on Minority Health and Health Disparities on September 24, 2023, with a project end date of April 30, 2027. Recruitment and enrollment began in May 2024. As of August 2025, we enrolled 13 participants. Table 2 presents the planned timeline and current progress for each study phase.

Table 2. Timeline and study progress.

Phase	Planned timeline	Current progress
Aim 1	May 2024-December 2025	Interviews completed for 9 participants
Aim 2	January 2025-December 2025	Data processed for analysis
Aim 3	January 2026-April 2026	Not yet initiated
Aim 4	May 2026-April 2027	Not yet initiated

Discussion

Anticipated Findings

Reducing racial biases in health care 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 EHRs 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, 5 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. This 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 United States, 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 this study, using a similar population, it is expected that these racial and 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 and BERT-based models, have been evaluated for identifying stigmatizing language in clinical notes [15,28,31,78,79]. 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 such as clinical BERT capture context better but are limited by the training data. Recent advancements in large language models, such as Mistral and Large Language Model Meta AI 3 [80,81], have improved the 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 an original 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 the language used 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. In addition, data annotators will represent diverse racial perspectives. Various “interrater 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.

Conclusions

The ENGAGE study protocol addresses the critical issue of stigmatizing language in HHC through developing an NLP-driven system. This innovative mixed methods study, conducted within 2 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 using 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

The authors thank the ENGAGE study participants, the research team members, and the partnering home health care agencies. The research is funded by the National Institute on Minority Health and Health Disparities (grant 1R01MD018028-01A1).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

SS, JYT, and MT were responsible for conceptualization, acquisition, investigation methodology, project administration resources, supervision, validation, and visualization. ZZ, PG, SS, MVM, JYT, and MT were responsible for data curation. ZZ, PG, and MT were responsible for formal analysis. SS, JYT, and MT were responsible for arriving at the investigation methodology. MT was responsible for software. ZZ, PG, SP-T, LP, MM, SS, MVM, CWR, JYT, and MT were responsible for writing, reviewing, and editing the original draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide for nurses.

[\[DOCX File, 17 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Interview guides for patients.

[\[DOCX File, 16 KB-Multimedia Appendix 2\]](#)

References

- Harris-Kojetin L, Sengupta M, Lendon J, Cefferly C. Long-term care providers and services users in the United States, 2015-2016. Centers for Disease Control and Prevention (CDC). 2019. URL: <https://stacks.cdc.gov/view/cdc/76253> [accessed 2022-01-13]
- March 2019 report to the congress: medicare payment policy. Medicare Payment Advisory Commission. 2019. URL: <https://www.medpac.gov/document/march-2019-report-to-the-congress-medicare-payment-policy/> [accessed 2025-05-29]
- Wei H, Sewell KA, Woody G, Rose MA. The state of the science of nurse work environments in the United States: a systematic review. *Int J Nurs Sci*. Jul 10, 2018;5(3):287-300. [FREE Full text] [doi: [10.1016/j.ijnss.2018.04.010](https://doi.org/10.1016/j.ijnss.2018.04.010)] [Medline: [31406839](https://pubmed.ncbi.nlm.nih.gov/31406839/)]
- Recio-Saucedo A, Dall'Ora C, Maruotti A, Ball J, Briggs J, Meredith P, et al. What impact does nursing care left undone have on patient outcomes? Review of the literature. *J Clin Nurs*. Jun 2018;27(11-12):2248-2259. [FREE Full text] [doi: [10.1111/jocn.14058](https://doi.org/10.1111/jocn.14058)] [Medline: [28859254](https://pubmed.ncbi.nlm.nih.gov/28859254/)]
- Wynendaele H, Willems R, Trybou J. Systematic review: association between the patient–nurse ratio and nurse outcomes in acute care hospitals. *J Nurs Manag*. Apr 15, 2019;27(5):896-917. [doi: [10.1111/JONM.12764](https://doi.org/10.1111/JONM.12764)]
- Blair IV, Steiner J, Havranek E. Unconscious (implicit) bias and health disparities: where do we go from here? *Perm J*. 2011;15(2):71-78. [FREE Full text] [doi: [10.7812/TPP/11.979](https://doi.org/10.7812/TPP/11.979)] [Medline: [21841929](https://pubmed.ncbi.nlm.nih.gov/21841929/)]
- FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. *BMC Med Ethics*. Mar 01, 2017;18(1):19. [FREE Full text] [doi: [10.1186/s12910-017-0179-8](https://doi.org/10.1186/s12910-017-0179-8)] [Medline: [28249596](https://pubmed.ncbi.nlm.nih.gov/28249596/)]
- Zestcott CA, Blair IV, Stone J. Examining the presence, consequences, and reduction of implicit bias in health care: a narrative review. *Group Process Intergroup Relat*. Jul 08, 2016;19(4):528-542. [FREE Full text] [doi: [10.1177/1368430216642029](https://doi.org/10.1177/1368430216642029)] [Medline: [27547105](https://pubmed.ncbi.nlm.nih.gov/27547105/)]
- Groves PS, Bunch JL, Sabin JA. Nurse bias and nursing care disparities related to patient characteristics: a scoping review of the quantitative and qualitative evidence. *J Clin Nurs*. May 22, 2021;30(23-24):3385-3397. [doi: [10.1111/JOCN.15861](https://doi.org/10.1111/JOCN.15861)]
- Hall WJ, Chapman MV, Lee KM, Merino YM, Thomas TW, Payne BK, et al. Implicit racial/ethnic bias among health care professionals and its influence on health care outcomes: a systematic review. *Am J Public Health*. Dec 2015;105(12):e60-e76. [doi: [10.2105/ajph.2015.302903](https://doi.org/10.2105/ajph.2015.302903)]
- Narayan M. CE: addressing implicit bias in nursing: a review. *Am J Nurs*. Jul 2019;119(7):36-43. [doi: [10.1097/01.NAJ.0000569340.27659.5a](https://doi.org/10.1097/01.NAJ.0000569340.27659.5a)] [Medline: [31180913](https://pubmed.ncbi.nlm.nih.gov/31180913/)]
- Topaz M, Song J, Davoudi A, McDonald M, Taylor J, Sittig S, et al. Home health care clinicians' use of judgment language for black and Hispanic patients: natural language processing study. *JMIR Nurs*. Apr 17, 2023;6:e42552. [FREE Full text] [doi: [10.2196/42552](https://doi.org/10.2196/42552)] [Medline: [37067893](https://pubmed.ncbi.nlm.nih.gov/37067893/)]
- Park J, Saha S, Chee B, Taylor J, Beach MC. Physician use of stigmatizing language in patient medical records. *JAMA Netw Open*. Jul 01, 2021;4(7):e2117052. [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.17052](https://doi.org/10.1001/jamanetworkopen.2021.17052)] [Medline: [34259849](https://pubmed.ncbi.nlm.nih.gov/34259849/)]
- P Goddu A, O'Connor KJ, Lanzkron S, Saheed MO, Saha S, Peek ME, et al. Do words matter? Stigmatizing language and the transmission of bias in the medical record. *J Gen Intern Med*. May 26, 2018;33(5):685-691. [FREE Full text] [doi: [10.1007/s11606-017-4289-2](https://doi.org/10.1007/s11606-017-4289-2)] [Medline: [29374357](https://pubmed.ncbi.nlm.nih.gov/29374357/)]
- Barcelona V, Scharp D, Moen H, Davoudi A, Idnay BR, Cato K, et al. Using natural language processing to identify stigmatizing language in labor and birth clinical notes. *Matern Child Health J*. Mar 26, 2024;28(3):578-586. [doi: [10.1007/s10995-023-03857-4](https://doi.org/10.1007/s10995-023-03857-4)] [Medline: [38147277](https://pubmed.ncbi.nlm.nih.gov/38147277/)]
- Andreyeva E, David G, Song H. The effects of home health visit length on hospital readmission. NBER Working Paper Series. 2018. URL: <http://www.nber.org/papers/w24566> [accessed 2025-05-29]
- 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*. Jan 2022;68(1):420-441. [doi: [10.1287/MNSC.2020.3921](https://doi.org/10.1287/MNSC.2020.3921)]
- Smith LM, Anderson WL, Kenyon A, Kinyara E, With SK, Teichman L, et al. Racial and ethnic disparities in patients' experience with skilled home health care services. *Med Care Res Rev*. Dec 03, 2015;72(6):756-774. [doi: [10.1177/1077558715597398](https://doi.org/10.1177/1077558715597398)] [Medline: [26238122](https://pubmed.ncbi.nlm.nih.gov/26238122/)]
- Wang J, Yu F, Cai X, Caprio TV, Li Y. Functional outcome in home health: do racial and ethnic minority patients with dementia fare worse? *PLoS ONE*. May 26, 2020;15(5):e0233650. [doi: [10.1371/JOURNAL.PONE.0233650](https://doi.org/10.1371/JOURNAL.PONE.0233650)]
- Davitt JK, Bourjolly J, Frasso R, Chan S. Understanding racial and ethnic disparities in home health care: practice and policy factors. *Innov Aging*. 2017;1(Suppl 1):956. [doi: [10.1093/geroni/igx004.3442](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*. Nov 26, 2017;28(6):598-607. [doi: [10.1177/1043659617700710](https://doi.org/10.1177/1043659617700710)] [Medline: [28826334](https://pubmed.ncbi.nlm.nih.gov/28826334/)]
22. Squires A, Ma C, Miner S, Feldman P, Jacobs EA, Jones SA. Assessing the influence of patient language preference on 30 day hospital readmission risk from home health care: a retrospective analysis. *Int J Nurs Stud*. Jan 2022;125:104093. [doi: [10.1016/J.IJNURSTU.2021.104093](https://doi.org/10.1016/J.IJNURSTU.2021.104093)]
23. Fernández L, Fossa A, Dong Z, Delbanco T, Elmore J, Fitzgerald P, et al. Words matter: what do patients find judgmental or offensive in outpatient notes? *J Gen Intern Med*. Sep 2021;36(9):2571-2578. [FREE Full text] [doi: [10.1007/s11606-020-06432-7](https://doi.org/10.1007/s11606-020-06432-7)] [Medline: [33528782](https://pubmed.ncbi.nlm.nih.gov/33528782/)]
24. 21st century cures act: interoperability, information blocking, and the ONC health IT certification program. Department of Health and Human Services. URL: https://www.healthit.gov/sites/default/files/cures/2020-03/ONC_Cures_Act_Final_Rule_03092020.pdf [accessed 2021-12-29]
25. Alvarado CS, 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. URL: <https://www.healthit.gov/sites/default/files/electronic-health-record-adoption-and-interoperability-among-u.s.-skilled-nursing-facilities-in-2016.pdf> [accessed 2021-12-29]
26. Matheny ME, Whicher D, Thadaney Israni S. Artificial intelligence in health care: a report from the national academy of medicine. *JAMA*. Feb 11, 2020;323(6):509-510. [doi: [10.1001/jama.2019.21579](https://doi.org/10.1001/jama.2019.21579)] [Medline: [31845963](https://pubmed.ncbi.nlm.nih.gov/31845963/)]
27. Demner-Fushman D, Chapman WW, McDonald CJ. What can natural language processing do for clinical decision support? *J Biomed Inform*. Oct 2009;42(5):760-772. [FREE Full text] [doi: [10.1016/j.jbi.2009.08.007](https://doi.org/10.1016/j.jbi.2009.08.007)] [Medline: [19683066](https://pubmed.ncbi.nlm.nih.gov/19683066/)]
28. Barcelona V, Scharp D, Idnay BR, Moen H, Cato K, Topaz M. Identifying stigmatizing language in clinical documentation: a scoping review of emerging literature. *PLoS One*. 2024;19(6):e0303653. [FREE Full text] [doi: [10.1371/journal.pone.0303653](https://doi.org/10.1371/journal.pone.0303653)] [Medline: [38941299](https://pubmed.ncbi.nlm.nih.gov/38941299/)]
29. Topaz M, Murga L, Bar-Bachar O, McDonald M, Bowles K. NimbleMiner: an open-source nursing-sensitive natural language processing system based on word embedding. *Comput Inform Nurs*. Nov 2019;37(11):583-590. [doi: [10.1097/CIN.0000000000000557](https://doi.org/10.1097/CIN.0000000000000557)] [Medline: [31478922](https://pubmed.ncbi.nlm.nih.gov/31478922/)]
30. Beach MC, Saha S, Park J, Taylor J, Drew P, Plank E, et al. Testimonial injustice: linguistic bias in the medical records of black patients and women. *J Gen Intern Med*. Mar 22, 2021;36(6):1708-1714. [doi: [10.1007/S11606-021-06682-Z](https://doi.org/10.1007/S11606-021-06682-Z)]
31. Himmelstein G, Bates D, Zhou L. Examination of stigmatizing language in the electronic health record. *JAMA Netw Open*. Jan 04, 2022;5(1):e2144967. [FREE Full text] [doi: [10.1001/jamanetworkopen.2021.44967](https://doi.org/10.1001/jamanetworkopen.2021.44967)] [Medline: [35084481](https://pubmed.ncbi.nlm.nih.gov/35084481/)]
32. Scroggins J, Hulchafo II, Harkins S, Scharp D, Moen H, Davoudi A, et al. Identifying stigmatizing and positive/preferred language in obstetric clinical notes using natural language processing. *J Am Med Inform Assoc*. Feb 01, 2025;32(2):308-317. [doi: [10.1093/jamia/ocae290](https://doi.org/10.1093/jamia/ocae290)] [Medline: [39569431](https://pubmed.ncbi.nlm.nih.gov/39569431/)]
33. Pinnock H, Barwick M, Carpenter CR, Eldridge S, Grandes G, Griffiths CJ, et al. StaRI Group. Standards for reporting implementation studies (StaRI) statement. *BMJ*. Mar 06, 2017;356:i6795. [FREE Full text] [doi: [10.1136/bmj.i6795](https://doi.org/10.1136/bmj.i6795)] [Medline: [28264797](https://pubmed.ncbi.nlm.nih.gov/28264797/)]
34. 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.0000000000000107](https://doi.org/10.1097/ANS.0000000000000107)] [Medline: [26836997](https://pubmed.ncbi.nlm.nih.gov/26836997/)]
35. 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](https://doi.org/10.1097/ANS.0b013e3182071813)] [Medline: [21150551](https://pubmed.ncbi.nlm.nih.gov/21150551/)]
36. NIMHD minority health and health disparities research framework. National Institute on Minority Health and Health Disparities. URL: <https://www.nimhd.nih.gov/about/overview/research-framework/> [accessed 2022-01-24]
37. 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. [FREE Full text] [doi: [10.1080/01621424.2012.703908](https://doi.org/10.1080/01621424.2012.703908)] [Medline: [23216513](https://pubmed.ncbi.nlm.nih.gov/23216513/)]
38. Suri H. Purposeful sampling in qualitative research synthesis. *Qual Res J*. 2011;11(2):63-75. [doi: [10.3316/QRJ1102063/FULL/XML](https://doi.org/10.3316/QRJ1102063/FULL/XML)]
39. 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*. Dec 15, 2021;284:426-430. [doi: [10.3233/SHTI210763](https://doi.org/10.3233/SHTI210763)] [Medline: [34920563](https://pubmed.ncbi.nlm.nih.gov/34920563/)]
40. Topaz M, Naylor MD, Holmes JH, Bowles KH. Factors affecting patient prioritization decisions at admission to home healthcare: a predictive study to develop a risk screening tool. *Comput Inform Nurs*. Feb 2020;38(2):88-98. [FREE Full text] [doi: [10.1097/CIN.0000000000000576](https://doi.org/10.1097/CIN.0000000000000576)] [Medline: [31804243](https://pubmed.ncbi.nlm.nih.gov/31804243/)]
41. Bowles KH, Ratcliffe S, Potashnik S, Topaz M, Holmes J, Shih N, et al. Using electronic case summaries to elicit multi-disciplinary expert knowledge about referrals to post-acute care. *Appl Clin Inform*. Dec 16, 2016;7(2):368-379. [FREE Full text] [doi: [10.4338/ACI-2015-11-RA-0161](https://doi.org/10.4338/ACI-2015-11-RA-0161)] [Medline: [27437047](https://pubmed.ncbi.nlm.nih.gov/27437047/)]
42. McDonald MV, Feldman PH, Barrón-Vayá Y, Peng TR, Sridharan S, Pezzin LE. 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*. Feb 2016;22(1):10-19. [FREE Full text] [doi: [10.1111/jep.12383](https://doi.org/10.1111/jep.12383)] [Medline: [26009977](https://pubmed.ncbi.nlm.nih.gov/26009977/)]

43. 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;6:2333721420959861. [FREE Full text] [doi: [10.1177/2333721420959861](https://doi.org/10.1177/2333721420959861)] [Medline: [33029550](https://pubmed.ncbi.nlm.nih.gov/33029550/)]
44. Dreisbach C, Koleck TA, Bourne PE, Bakken S. A systematic review of natural language processing and text mining of symptoms from electronic patient-authored text data. *Int J Med Inform*. May 2019;125:37-46. [FREE Full text] [doi: [10.1016/j.ijmedinf.2019.02.008](https://doi.org/10.1016/j.ijmedinf.2019.02.008)] [Medline: [30914179](https://pubmed.ncbi.nlm.nih.gov/30914179/)]
45. Topaz M, Murga L, Gaddis KM, McDonald MV, Bar-Bachar O, Goldberg Y, et al. Mining fall-related information in clinical notes: comparison of rule-based and novel word embedding-based machine learning approaches. *J Biomed Inform*. Feb 2019;90:103103. [FREE Full text] [doi: [10.1016/j.jbi.2019.103103](https://doi.org/10.1016/j.jbi.2019.103103)] [Medline: [30639392](https://pubmed.ncbi.nlm.nih.gov/30639392/)]
46. 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*. Mar 01, 2006;5(1):80-92. [doi: [10.1177/160940690600500107](https://doi.org/10.1177/160940690600500107)]
47. Training data labeling using humans with Amazon SageMaker ground truth. Amazon Web Services. URL: <https://docs.aws.amazon.com/sagemaker/latest/dg/sms.html> [accessed 2021-12-09]
48. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med (Zagreb)*. 2012;22(3):276-282. [FREE Full text] [Medline: [23092060](https://pubmed.ncbi.nlm.nih.gov/23092060/)]
49. 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*. May 2021;69(3):435-446. [FREE Full text] [doi: [10.1016/j.outlook.2020.12.007](https://doi.org/10.1016/j.outlook.2020.12.007)] [Medline: [33386145](https://pubmed.ncbi.nlm.nih.gov/33386145/)]
50. Blumenthal KG, Topaz M, Zhou L, Harkness T, Sa'adon R, Bar-Bachar O, et al. Mining social media data to assess the risk of skin and soft tissue infections from allergen immunotherapy. *J Allergy Clin Immunol*. Jul 2019;144(1):129-134. [FREE Full text] [doi: [10.1016/j.jaci.2019.01.029](https://doi.org/10.1016/j.jaci.2019.01.029)] [Medline: [30721764](https://pubmed.ncbi.nlm.nih.gov/30721764/)]
51. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. arXiv. Preprint posted online on October 11, 2018. [FREE Full text] [doi: [10.5260/chara.21.2.8](https://doi.org/10.5260/chara.21.2.8)]
52. Alsentzer E, Murphy JR, Boag W, Weng WH, Jin D, Nauman T, et al. Publicly available clinical BERT embeddings. arXiv. preprint posted online on April 6, 2019. [FREE Full text] [doi: [10.18653/v1/w19-1909](https://doi.org/10.18653/v1/w19-1909)]
53. Johnson AE, Pollard TJ, Shen L, Lehman LH, Feng M, Ghassemi M, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. May 24, 2016;3:160035. [FREE Full text] [doi: [10.1038/sdata.2016.35](https://doi.org/10.1038/sdata.2016.35)] [Medline: [27219127](https://pubmed.ncbi.nlm.nih.gov/27219127/)]
54. Guthrie D, Allison B, Liu W, Guthrie L, Wilks Y. A closer look at skip-gram modelling. In: *Proceedings of the 5th International Conference on Language Resources and Evaluation*. 2006. Presented at: ELRA '06; May 22-28, 2006:1222-1225; Genoa, Italy. URL: http://www.lrec-conf.org/proceedings/lrec2006/pdf/357_pdf.pdf
55. 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. Presented at: EMNLP '14; October 25-29, 2014:1532-1543; Doha, Qatar. URL: <https://nlp.stanford.edu/projects/glove/> [doi: [10.3115/v1/d14-1162](https://doi.org/10.3115/v1/d14-1162)]
56. 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*. 2016. Presented at: EACL '17; April 3-7, 2017:427-431; Valencia, Spain. URL: <https://aclanthology.org/E17-2068> [doi: [10.18653/v1/e17-2068](https://doi.org/10.18653/v1/e17-2068)]
57. Sarzynska-Wawer J, Wawer A, Pawlak A, Szymanowska J, Stefaniak I, Jarkiewicz M, et al. Detecting formal thought disorder by deep contextualized word representations. *Psychiatry Res*. Oct 2021;304:114135. [doi: [10.1016/j.psychres.2021.114135](https://doi.org/10.1016/j.psychres.2021.114135)] [Medline: [34343877](https://pubmed.ncbi.nlm.nih.gov/34343877/)]
58. Li J, Jia R, He H, Liang P. Delete, retrieve, generate: a simple approach to sentiment and style transfer. In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. 2018. Presented at: NAACL '18; June 1-6, 2018:1865-1874; New Orleans, LA. URL: <https://aclanthology.org/N18-1169.pdf> [doi: [10.18653/v1/n18-1169](https://doi.org/10.18653/v1/n18-1169)]
59. Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language models are few-shot learners. arXiv. preprint posted online on May 28, 2020. [FREE Full text] [doi: [10.5860/choice.189890](https://doi.org/10.5860/choice.189890)]
60. Raffel C, Shazeer N, Roberts A, Lee K, Narang S. Exploring the limits of transfer learning with a unified text-to-text transformer. *Mach Learn Res*. 2020;79(5):184-212. [FREE Full text] [doi: [10.5260/chara.21.2.8](https://doi.org/10.5260/chara.21.2.8)]
61. Helander MG. *Handbook of Human Computer Interaction: Living Reference Work*. New York, NY: Elsevier; 2014.
62. Ferreira J, Noble J, Biddle R. Agile development iterations and UI design. In: *Proceedings of the 2007 International Conference on Agile Software Development*. 2007. Presented at: AGILE '07; August 13-17, 2007:50-58; Washington, DC. URL: <https://ieeexplore.ieee.org/document/4293575> [doi: [10.1109/AGILE.2007.8](https://doi.org/10.1109/AGILE.2007.8)]
63. Sittig S, Wang J, Iyengar S, Myneni S, Franklin A. Incorporating behavioral trigger messages into a mobile health app for chronic disease management: randomized clinical feasibility trial in diabetes. *JMIR Mhealth Uhealth*. Mar 16, 2020;8(3):e15927. [FREE Full text] [doi: [10.2196/15927](https://doi.org/10.2196/15927)] [Medline: [32175908](https://pubmed.ncbi.nlm.nih.gov/32175908/)]
64. Topaz M, Trifilio M, Maloney D, Bar-Bachar O, Bowles KH. Improving patient prioritization during hospital-homecare transition: a pilot study of a clinical decision support tool. *Res Nurs Health*. Oct 2018;41(5):440-447. [doi: [10.1002/nur.21907](https://doi.org/10.1002/nur.21907)] [Medline: [30203417](https://pubmed.ncbi.nlm.nih.gov/30203417/)]

65. Bowles KH, Holmes JH, Ratcliffe SJ, Liberatore M, Nydick R, Naylor MD. Factors identified by experts to support decision making for post acute referral. *Nurs Res.* 2009;58(2):115-122. [FREE Full text] [doi: [10.1097/NNR.0b013e318199b52a](https://doi.org/10.1097/NNR.0b013e318199b52a)] [Medline: [19289932](https://pubmed.ncbi.nlm.nih.gov/19289932/)]
66. Topaz M, Rao A, Masterson Creber R, Bowles K. Educating clinicians on new elements incorporated into the electronic health record: theories, evidence, and one educational project. *Comput Inform Nurs.* Aug 2013;31(8):375-380. [FREE Full text] [doi: [10.1097/NXN.0b013e318295e5a5](https://doi.org/10.1097/NXN.0b013e318295e5a5)] [Medline: [23774448](https://pubmed.ncbi.nlm.nih.gov/23774448/)]
67. Bowles KH, Chittams J, Heil E, Topaz M, Rickard K, Bhasker M, et al. Successful electronic implementation of discharge referral decision support has a positive impact on 30- and 60-day readmissions. *Res Nurs Health.* Apr 25, 2015;38(2):102-114. [FREE Full text] [doi: [10.1002/nur.21643](https://doi.org/10.1002/nur.21643)] [Medline: [25620675](https://pubmed.ncbi.nlm.nih.gov/25620675/)]
68. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol.* Jan 2006;3(2):77-101. [doi: [10.1191/1478088706qp063oa](https://doi.org/10.1191/1478088706qp063oa)]
69. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: implications for conducting a qualitative descriptive study. *Nurs Health Sci.* Sep 11, 2013;15(3):398-405. [doi: [10.1111/nhs.12048](https://doi.org/10.1111/nhs.12048)] [Medline: [23480423](https://pubmed.ncbi.nlm.nih.gov/23480423/)]
70. 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. [doi: [10.1097/NNR.0b013e318276250c](https://doi.org/10.1097/NNR.0b013e318276250c)] [Medline: [23154659](https://pubmed.ncbi.nlm.nih.gov/23154659/)]
71. Topaz M, Bar-Bachar O, Admi H, Denekamp Y, Zimlichman E. Patient-centered care via health information technology: a qualitative study with experts from Israel and the U.S. *Inform Health Soc Care.* Sep 27, 2020;45(3):217-228. [doi: [10.1080/17538157.2019.1582055](https://doi.org/10.1080/17538157.2019.1582055)] [Medline: [30917717](https://pubmed.ncbi.nlm.nih.gov/30917717/)]
72. Koru G, Alhuwail D, Topaz M, Norcio AF, Mills ME. Investigating the challenges and opportunities in home care to facilitate effective information technology adoption. *J Am Med Dir Assoc.* Jan 2016;17(1):53-58. [doi: [10.1016/j.jamda.2015.10.008](https://doi.org/10.1016/j.jamda.2015.10.008)] [Medline: [26612483](https://pubmed.ncbi.nlm.nih.gov/26612483/)]
73. Topaz M, Seger DL, Goss F, Lai K, Slight SP, Lau JJ, et al. Standard information models for representing adverse sensitivity information in clinical documents. *Methods Inf Med.* Jan 08, 2018;55(02):151-157. [doi: [10.3414/me15-01-0081](https://doi.org/10.3414/me15-01-0081)]
74. 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.* Oct 26, 2018;14(3):217-228. [doi: [10.1558/cam.32815](https://doi.org/10.1558/cam.32815)]
75. NVivo qualitative data analysis software. QSR International. URL: <http://qsrinternational.com/about-us/newsroom/qsr-international-launches-qualitative-data-analysis> [accessed 2025-05-29]
76. Siblini W, Fréry J, He-Guelton L, Oblé F, Wang YQ. Master your metrics with calibration. *arXiv.* preprint posted online April 28, 2020. [FREE Full text] [doi: [10.1007/978-3-030-44584-3_36](https://doi.org/10.1007/978-3-030-44584-3_36)]
77. Kaufmann J, Schering AG. Analysis of Variance ANOVA. Hoboken, NJ. Wiley-Blackwell; 2014.
78. Harrigian K, Zirikly A, Chee B, Ahmad A, Links A, Saha S, et al. Characterization of stigmatizing language in medical records. *PhysioNet.* 2023;(2):312-329. [FREE Full text] [doi: [10.18653/v1/2023.acl-short.28](https://doi.org/10.18653/v1/2023.acl-short.28)]
79. Weiner SG, Lo YC, Carroll AD, Zhou L, Ngo A, Hathaway DB, 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. [FREE Full text] [doi: [10.1097/ADM.0000000000001145](https://doi.org/10.1097/ADM.0000000000001145)] [Medline: [37579100](https://pubmed.ncbi.nlm.nih.gov/37579100/)]
80. Jiang AQ, Sablayrolles A, Mensch A, Bamford C, Chaplot DS, de las Casas D, et al. Mistral 7B. *arXiv.* Preprint posted online on October 10, 2023. [FREE Full text] [doi: [10.5260/chara.21.2.8](https://doi.org/10.5260/chara.21.2.8)]
81. Grattafiori A, Dubey A, Jauhri A, Pandey A, Kadian A, Al-Dahle A, et al. The llama 3 herd of models. *arXiv.* Preprint posted online on July 31, 2024. [FREE Full text] [doi: [10.5860/choice.189890](https://doi.org/10.5860/choice.189890)]

Abbreviations

AUC-ROC: area under the receiver operating characteristic curve
AUC-RP: area under the curve of precision recall
BERT: bidirectional encoder representations from transformers
EHR: electronic health record
ENGAGE: reduce stigmatizing language
HHC: home health care
NLP: natural language processing
OASIS: Outcome and Assessment Information Set
SEB: Stakeholders Engagement Board

Edited by A Schwartz; The proposal for this study was peer-reviewed by: ZEBI OSR-G (O1) - National Institute of Biomedical Imaging and Bioengineering Special Emphasis Panel (National Institutes of Health, USA). See the Multimedia Appendix for the peer-review report; Submitted 06.Dec.2024; accepted 18.Jun.2025; published 25.Sep.2025.

Please cite as:

Zhang Z, Gupta P, Potts-Thompson S, Prescott L, Morrison M, Sittig S, McDonald MV, Raymond C, Taylor JY, Topaz M
Identifying and Reducing Stigmatizing Language in Home Health Care With a Natural Language Processing–Based System (ENGAGE): Protocol for a Mixed Methods Study

JMIR Res Protoc 2025;14:e69753

URL: <https://www.researchprotocols.org/2025/1/e69753>

doi: [10.2196/69753](https://doi.org/10.2196/69753)

PMID:

©Zhihong Zhang, Pallavi Gupta, Stephanie Potts-Thompson, Laura Prescott, Morgan Morrison, Scott Sittig, Margaret V McDonald, Chase Raymond, Jacquelyn Y Taylor, Maxim Topaz. Originally published in JMIR Research Protocols (<https://www.researchprotocols.org>), 25.Sep.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Research Protocols, is properly cited. The complete bibliographic information, a link to the original publication on <https://www.researchprotocols.org>, as well as this copyright and license information must be included.