

Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

DocuSigned by:
Signature: *Andrew L Walker*
4DB87F8FD70D4BA...

Andrew L Walker
Name

4/3/2024 | 3:11 PM EDT
Date


Title Using natural language processing to detect stigmatizing provider language and evaluate associations with opioid analgesic pain management outcomes

Author Andrew L Walker


Degree Doctor of Philosophy

Program Behavioral, Social, and Health Education Sciences


Approved by the Committee

DocuSigned by:

EC7FF9FB1D29460...


Melvin Livingston
Advisor

DocuSigned by:

B98807E2AE0C427...

Hannah LF Cooper
Committee Member

DocuSigned by:

846A48DA378648B...

Jennifer Love
Committee Member

DocuSigned by:

6956B47460DC486...

Abeer Sarker
Committee Member

Committee Member

Committee Member

Accepted by the Laney Graduate School:

Kimberly Jacob Arriola, Ph.D, MPH
Dean, James T. Laney Graduate School

Date

April 5th, 2024

Using natural language processing to detect stigmatizing provider language and evaluate
associations with opioid analgesic pain management outcomes

By

Andrew L Walker

MS, University of Florida, 2019

Advisor: Melvin Livingston, PhD

An abstract of

A dissertation submitted to the Faculty of the James T. Laney School of Graduate Studies of

Emory University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Behavioral, Social, and Health Education Sciences

2024

Abstract

Using natural language processing to detect stigmatizing provider language and evaluate associations with opioid analgesic pain management outcomes

By Andrew L Walker

Objective: In this dissertation, we sought to detect and classify stigmatizing and biased language in intensive care unit (ICU) electronic health records (EHRs) using natural language processing techniques. We evaluated the prevalence of such language across different patient demographics and provider factors, and explored the association of linguistic biases with care outcomes, including opioid analgesic prescription rates, dispensation rates, and rates of self-directed discharge from the ICU. **Methods:** Utilizing the Medical Information Mart for Intensive Care-III (MIMIC-III) dataset, we developed a comprehensive lexicon from literature-driven stem words, expanded with Word2Vec and GPT 3.5, to identify stigmatizing patient labels, doubt markers, and scare quotes. This lexicon was used to search 18 million sentences, 3000 of which were then used to train various classifiers, including bag-of-words and transformer-based models. Supervised learning techniques assessed the distribution of linguistic bias and its clustering within patient records, leveraging sentence-level analysis to connect linguistic features with patient care outcomes. **Results:** We developed lexicons and with high utility in identifying stigmatizing labels and doubt markers, and classifiers showing high accuracy, recall, and precision. Stigmatizing labels and doubt markers was found to be more prevalent among historically marginalized groups, with notable disparities in care outcomes, namely higher likelihoods of self-directed discharge. No significant associations were found between the linguistic features and opioid prescription or dispensation rates. **Discussion:** This dissertation supports the feasibility of using natural language processing to identify stigmatizing and doubt-marking language within medical records. It highlights consistent trends of stigmatizing language, particularly against historically stigmatized patients, and underscores the need for further research and intervention to mitigate these stigmas and downstream health outcomes. **Conclusions:** The high performance of the classifiers, titled CARE-SD: Classifier-based Analysis for Recognizing and Eliminating Stigmatizing and Doubt Marker Labels in Electronic Health Records, underscores their potential for broader application in identifying stigmatizing language within healthcare systems. Study findings also highlight the importance of addressing stigmatizing language as a component of quality care and suggest that methods used in the study can be used to reduce stigmatization in EHR notes and identify areas of intervention.

Using natural language processing to detect stigmatizing provider language and evaluate
associations with opioid analgesic pain management outcomes

By

Andrew L Walker

MS, University of Florida, 2019

Advisor: Melvin Livingston, PhD

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Behavioral, Social, and Health Education Sciences

2024

Acknowledgments

This project is the result of many people who have believed in me and supported me before and throughout my doctoral journey. Thank you to my wife, Isabel Cecilia Walker, for her undying love and support, and for her confidence that a journey like this was possible for me. Thank you to Doug, for taking me on as an advisor when most people had no idea what to do with me. You are a huge reason I stayed in this program, and I cannot possibly have imagined a more empathetic and brilliant advisor to learn from. Thank you to my committee members Abeed, Hannah, and Jenn, for guiding me through this research process, hyping me up, always believing in my work, and lending your expertise and valuable insights at every step. Thank you to Mom, Dad, Mami, and Sal for always believing in me, being proud of me, and having my back with every step of the way. Thank you to the LBBs: Jordan, Coyt, and Shelton, for helping me remember that I'm more than this work and for keeping me grounded. Thank you to Swathi, Snigdha, and Jerik, who I'm so grateful to have met along this process. Your friendship and kindness have kept me sane more times than I can count. Thank you to Annie, for helping me code thousands of chart sentences, and for providing so many important insights along the way that have shaped these valuable tools. Thanks to Em and Anna for all the work and love you give for Streetlight and for our patients-- present and future. Thank you for always believing in my work and my abilities as a researcher.

Finally, thank you to all the Streetlight patients who I continue to carry with me. They have inspired this work and have taught me more than I could ever learn from a PhD.

This road is not always easy, and no one should have to go it alone.

Table of Contents

Introductory Literature Review	1
Defining Stigma and Bias	1
Ecosocial theory + stigma/bias	3
Clinical Outcomes of Associated with Stigma	12
Pain Management	13
Self-directed discharge	13
Origins and targets of stigmatization	14
Stigmatized Chronic Illness Populations frequently prescribed opioids	18
Current approaches to measuring Stigma and Bias	21
Survey methods	21
Administrative Data	22
Vignette studies	22
Implicit Association Tests (IATs)	23
Solution: Mining patient care notes to identify bias and stigma	25
Linguistic manifestations of bias and stigmatization: The Social Categories and Stereotype Communication Framework	26
Stigmatizing labels and negative descriptors	27
Evidentials and markers of doubt	28
Scare Quotes	29
Natural Language Processing to identify stigma and bias in EHR Data	30
Introductory Literature Review References	32
Aim 1: CARE-SD: Classifier-based Analysis for Recognizing and Eliminating Stigmatizing and Doubt Marker Labels in Electronic Health Records: model development and validation	53
Abstract	53
Introduction:	54
Linguistic manifestations of stigmatization	56
Stigmatizing labels	57
Doubt Markers	58
Scare Quotes	59
Natural Language Processing to identify stigma and bias in EHR Data	60
Methods:	60

MIMIC-III Dataset	61
Lexicon development and sample preparation	61
Stigmatizing labels	63
Doubt markers	63
Scare quotes	63
Matching with sentences in MIMIC-III, creating coding samples	64
Annotation process	64
Sentence Classification	65
Results:	66
Lexicon Development	66
Regular Expression Search results	67
Annotation	70
Linguistic Bias Classifier Model Evaluation Results	71
Discussion:	74
Integrative summary of findings	74
Limitations	77
References	79
Appendix 1: Lexicons for Doubt Markers and Stigmatizing Labels	84
Appendix 2: Stigmatizing Labels Ontology	87
Coding Process	87
Link and Phelan Stigma Definition	87
Stigmatizing labels and negative descriptors in charts	88
Coding Rules	90
Appendix 3: Doubt Markers Ontology	92
Coding Process	92
Coding Rules	94
Appendix 4: Scare Quotes Ontology	95
Coding Process	95
Coding Rules	97
Appendix 5: Best Performing Model Hyperparameters	99
Aim 2: Distribution of stigmatizing and doubt-marking language in EHR across patients, providers, and frequently-stigmatized diagnoses	100
Introduction	101
Methods	107
CARE-SD: Classifier-based Analysis for Recognizing and Eliminating Stigmatizing and Doubt Marker Labels in Electronic Health Records	109
Doubt markers	110
Results	113

Patient-level models- demographic predictors	118
Provider-level models- provider type	119
Patient-level models interaction model estimates with Black + African American /Non-Black + African American race variable	120
Discussion	121
Appendix 1: Patient Ethnicity and Provider Recategorizations	138
Appendix 2: Lexicons for Doubt Markers and Stigmatizing Labels	139
Aim 3: Evaluating the relationships between stigmatizing language features in the EHR and patient care outcomes of opioid analgesic prescription rates and self-directed discharge	142
Abstract:	142
Introduction	143
Defining Stigma	143
Stigmatizing labels	144
Doubt Markers	145
Theoretical frameworks for how provider stigmatizing language impacts health outcomes	145
Clinical Outcomes of Associated with Stigma	147
Pain Management	148
Self-directed discharge	148
Methods	149
Stigmatizing labels	150
Doubt markers	151
Results	155
Discussion	157
Conclusions	161
References	161
Appendix 1	170
Dissertation Conclusion	172

List of Tables and Figures

Introductory Literature Review

Figure 1: Conceptual model	10
<u>Aim 1</u>	
Figure 1: Natural language processing analytic pipeline	62
Figure 2: Network Diagram showing doubt marker lexicon expansion	67
Figure 2: Top 20 Matched Terms for each feature	69
Table 2: Annotation sample reliability	70
Table 3: Notable annotation examples	71
Table 4: Linguistic Bias Classifier Model Performance	72

Figure 3: Top 30 Stigmatizing Label Tokens by Importance	73
Appendix 1: Lexicons for Doubt Markers and Stigmatizing Labels	84
<u>Aim 2</u>	
Table 1: Summary of recent findings related to linguistic stigmatization	103
Figure 1: Natural language processing pipeline	108
Table 2: Patient-level descriptive statistics on rates of stigmatizing labels and doubt	115
Table 3: Frequencies of provider types	115
Table 4: Provider-level descriptive statistics on rates of stigmatizing labels and doubt	116
Table 3: Poisson regression results patients	118
Table 4: Poisson regression results providers	119
Table 5: Patient-level interaction model estimates	121
Appendix 1: Patient Ethnicity and Provider Recategorizations	139
Appendix 2: Lexicons for Doubt Markers and Stigmatizing Labels	139
<u>Aim 3</u>	
Figure 1: Conceptual model	147
Table 1: Descriptive results of study variables	156
Table 2: Regression results	157

Introductory Literature Review

Provider biases and stigmatization of patients drive health care inequities. Provider stigmas and biases are widely believed to contribute to discrimination and health disparities among patients.^{2,26} Patients regularly experience stigma and biases by providers as a result of their race, gender, sexual orientation, disease status, drug use, and socioeconomic status, or other labeled characteristics identified by provider teams.¹ Patients with multiple stigmatized identities and characteristics experience greater stigmatization and face unique challenges.²⁷ Leaders across provider specialties and disciplines have long called for investigation of the processes by which provider bias forms and is cultivated among medical teams.²⁸⁻³⁰ Recent efforts to mandate implicit bias trainings for providers have achieved some success in that 5 states currently have enacted implicit bias training mandates, with more expected to follow.³¹ While efforts to curb the effects of stigma and bias are promising, due to the multifaceted, dynamic, and persistent impacts on health, it is of urgent priority for public and clinical health experts to develop novel methods for measuring and intervening on stigma and bias. Drawing from ecosocial theory, relevant sociolinguistic frameworks, and advances in natural language processing, there is great potential for researchers to investigate manifestations of bias and stigma within the data encoded in electronic health record (EHR) notes. Being able to assess instances of bias and stigma in the health record could allow healthcare teams the ability to evaluate and intervene upon provider bias and mitigate its impact on patient care.

Defining Stigma and Bias

Stigma has been defined by social psychologists Link and Phelan as a social process that is characterized by the interplay of labeling, stereotyping and separation, which leads to status loss and discrimination, and importantly, occurs within a context of power, such as that of the

patient-provider relationship.³² Labeling occurs whenever we identify individuals as belonging to a particular group. Stereotyping involves ascribing a proclivity towards a specific behavior or characteristic to members of a labeled group. Separation involves distancing from the group, and drawing lines of “us versus them”. Status loss and discrimination involves a negative evaluation of a groups’ attributes and its members relative to another group. As Link and Phelan posit, all of these processes crucially develop in the context of power and power imbalances, which are extremely relevant in the context of provider-patient relationships.^{33,34}

Link and Phelan’s definition of stigma has been applied towards studying the sociocultural phenomenon across a variety of populations and settings, including healthcare.^{26,35} Within this applied context and others, it is useful to differentiate types of stigmatization experienced across multiple dimensions. Firstly, whether the stigma is enacted or felt. Enacted stigma includes “negative feelings, thoughts, and actions that target members of the stigmatized group”, which may occur in patient-provider relationships in the form of negative attitudes during interactions, sharing of implicit biases, attitudes, and stereotypes among members of the medical team, and in reduced patient autonomy in medical decision-making.³⁵ Conversely, felt stigma involves processes by stigmatized individuals in which they are aware of, process, and may be affected by the biases and discrimination directed towards them because of their stigmatized condition. Felt stigma may or may not be impacted by enacted stigma by the part of providers, and felt, or internalized stigma may cause patients to make decisions to forgo treatment or not disclose important health information to providers. These different aspects of stigma likely also involve different methods of measurement. In the context of the current study, enacted stigmatization can occur in provider notes without patient (or potentially, even provider) knowledge. This can have patient-level impacts even when patients do not perceive themselves

as being stigmatized. Because of this, it is difficult to measure this stigmatization from the perspective of both patients and providers, who may not be aware of the stigmatizing language they use for their patients. Similarly, biases, which may form based on socially-constructed stigmatization, are altered perceptions that cause individuals to give favor to or act against certain things, people, or groups, which often are the result of conscious or unconscious stigmatization.¹ Explicit biases are conscious acts of bias towards others, which the individual enacting the bias is aware of. Examples of explicit biases include overt racism, making racist comments, or deliberate prejudice. Implicit biases are unconscious negative evaluations about others which unknowingly (to the individual enacting the bias) drive attitudes and behaviors against stigmatized groups. Central to the focus of this study, enacted stigma and implicit biases can be identified through the manifestation of biased and stigmatizing language found within health care provider notes.

Ecosocial theory + stigma/bias

The growing study at the intersection of stigma, provider biases and patient health can greatly benefit from the application of Krieger's ecosocial theory, which asserts that embodiment, processes of embodiment, interaction of susceptibility and protective factors, and systems of accountability and agency all contribute to the accumulation of current and historical sociocultural exposures that impact one's biopsychosocial experience.³⁶ By placing power dynamics at the forefront of its constructs, this theory is aptly developed to confront the layered power structures through which patients experience stigma and bias from healthcare providers, teams, and organizations, particularly within clinical situations related to opioid analgesic pain management.¹⁴ Given the complexity of the modes through which provider stigma and bias operate, and the diversity of health outcomes it impacts, ecosocial theory provides a

comprehensive framework for investigating the complex, multifaceted processes through which provider stigma and bias impacts the health of individuals within the various power structures in their lives.³⁶ While this study does not see to incorporate every level described by ecosocial theory, we invoke its framework to contextualize the focus of provider stigmatization and health outcomes.

One of the primary constructs in ecosocial theory, **embodiment** is an individual's bodily manifestation of stressors and exposures from the biopsychosocial environment over the lifetime.³⁶ In the application of stigma, the construct of embodiment has allowed researchers to assess the impact of racism, sexism, and many other forms of stigma from the societal, structural, and policy levels, down to the individual, and even into the intercellular level.¹² Although stigma within the patient-provider relationship can be conceptualized as interpersonal and organizational-level phenomena, the ecosocial theory allows researchers to incorporate the cumulative experiences of stigma which patients may have experienced outside and within their clinical care that may influence health outcomes related to opioid analgesic pain management or self-directed discharge.

Ecosocial theory asserts that our biological and social trajectories, characterized by social arrangements of power and property, production, consumption, and reproduction are defined in ecosocial theory as **pathways of embodiment**.³⁶ This conceptualization allows us to extend the idea of stigma embodiment and investigate stigma as not just a static exposure, but a complex and dynamic process that occurs within socially constructed power structures which act to impact individuals' social and biological developmental trajectories. Bias and stigmas from providers may evolve into stigmatized patient reputations, as often experienced by patients with SCD, who often require multiple hospitalizations per year to treat their vaso-occlusive pain

crises, are frequently misjudged as hyper-utilizers of the system or get labeled as drug-seeking.^{37,38} Throughout multiple hospitalizations across many providers, biases and stigmas that develop against patients can accumulate and result in the formation of harmful reputations with the ability to impact care.^{4,21} Applying a theoretical framework that allows us the complexity to assess the many processes and cumulative effects of these exposures is key in assessing the relationships between stigma, bias, and patient health.

The third construct of ecosocial theory, **cumulative interplay between exposure, susceptibility and resistance**, encourages researchers to assess the interactive effects of multiple types of patient identity in the study of stigma and bias. The construct of cumulative interplay of exposure, susceptibility, and resistance encourages the examination of a variety of both risk and protective factors, leaning away from a solely risk-based model of disease. This is helpful in understanding how individuals can adapt and cope with stigma and bias, and is applied within a contextual framework that avoids blaming patients for the stigma and bias they experience.

Ecosocial's fourth construct of **accountability and agency** helps us design studies that help to identify health disparities and works to apply ecosocial theory as a framework for promoting a moral and ethical commitment to change and resistance. Evaluating stigma and bias in provider clinical notes can help to identify intervention points and understand how teams of providers and/or medical institutions may work to perpetuate stigmatization of marginalized patient populations. Bordieu argues in his work on social classes that "a power mechanism is actually more effective if it is difficult to identify because it is harder to challenge".³⁹ With this in mind, ecosocial theory helps not only further develop novel efforts to measure stigma and bias within the patient chart, but to challenge its presence among medical teams.

As we dive deeper into the study of stigma's formation and transmission among providers and healthcare teams, it is possible that we will need to adopt theoretical lenses from many different disciplines, as is currently being done in the fields of computational linguistics,^{40,41} to network theory,⁴² social psychology,⁴³ and critical race theory and intersectionality,⁴⁴ in order to appropriately capture the multifaceted ways through which providers develop, transmit, and act on stigma and bias. The Ecosocial framework allows us the flexibility to incorporate these theories that may provide more detailed explication of mechanisms through which provider bias and stigma operate, which we can then situate within constructs in the ecosocial theory to assess their impact on the health of individuals who are dynamically embedded in multidimensional levels of space, culture, biology, power, and time. While more individual theories of health behavior are helpful in understanding individual actions, these theories alone are insufficient to completely incorporate the multi-modal, multi-level mechanisms through which provider stigma and bias develops among health care professionals and negatively affects patient care.³² Ecosocial theory has been combined with other theories to still understand mechanisms of specific phenomena while adequately accounting for complex contextual factors.

Ecosocial theory has been used to investigate numerous negative health outcomes linked to experiences of provider bias and stigmatization, including as acting as a barrier to future care,⁴⁵⁻⁴⁷ reduced trust in medical team,^{45,48-50} increased loneliness,⁵⁰ lower medication adherence,^{51,52} reduced likelihood of being prescribed pain medication,^{37,53-57} as well as increased likelihood of a patient leaving against medical advice.⁵⁸ This framework is apt to describe the complex, multi-faceted experience through which patients encounter stigmatization based on race, gender, diagnosis, socioeconomic status, sexual orientation, and many other factors through a variety of experiences that happen before, after, and beyond the patient-provider relationship.

Stigmatization is associated with many negative health outcomes for populations, and is posited to work through a variety of mechanisms and pathways. One proposed pathway is that stigma harms individual health by increasing social isolation, and reducing social support.⁵⁹⁻⁶¹ Evidence of this pathway has been demonstrated by several studies which found an attenuation of effect after controlling for social support in assessing stigma's impact on health outcomes, across a variety of stigmatized identities.^{26,59}

Others have identified the ways in which “felt”, or self-stigma, can work to overburden and deplete an individuals' psychological resources, which may lead to maladaptive coping behaviors like isolation, rumination, suppression, smoking, or drinking.⁶¹ Minority stress theory has been applied widely to self-stigmatization, which helps to understand the excess stress to which individuals from stigmatized groups are exposed as a result of social position.⁶² This framework has identified that the stress of individuals' experiencing discrimination and unfair treatment is associated with adverse physiological outcomes, including higher diastolic blood pressure reactivity⁶³, increased cortisol output⁶⁴, and is particularly harmful with chronic exposure. Ecosocial theory has led the field of public health forward due to its application in developing detailed measures and study designs that study the historical and contemporary impacts of experiences of stigma, bias and discrimination at the structural (policy, built environment, organizations), interpersonal, individual levels.¹²

Levels of Stigma/Bias

Stigma and bias manifests in various interacting levels of societal, structural, environmental, interpersonal, and intraindividual factors which affect patients before, during, and after they leave the clinical setting.⁶⁵ While the focus of this study is on the interactional,

inter-individual, and organizational level as assessed through language in the patient EHR, it is nevertheless important to understand the many and varied forms of stigma and bias experienced by patients, which negatively impacts care, drives health inequities, and erodes the quality of patient-provider relationships.²⁶

Structural stigmatization focuses on higher-level aspects of the built and sociopolitical environment which contribute to exacerbating or mitigating stigmatization as experienced through provider or health organization interactions and systems. Though this level is still understudied, early research has shown that these structural-level factors play a major role in pathways of stigmatization.^{66 8} In the context of healthcare policies, providers participating in the adoption of the Prescription Drug Monitoring Program (PDMP) have reported that these new systems have increased tensions related to stigma of substance use disorders, and have made providers take on the policy enforcer role which places them in a position to view and treat patients as criminals.⁶⁸ Stigma and biases stigma and bias can be amplified by “risk score” algorithms, including those used to determine organ transplantation and clinical severity, further marginalizing patients in stigmatized minority groups.⁶⁹

Organizational-level norms regarding treatment of stigmatized groups likely also impact provider stigma and bias towards patients. Despite new NHLBI recommendations for rapid pain treatment for individuals with sickle cell disease (SCD), a recent study on SCD providers show only 32.4% of physicians, nurses, and PAs aware of new NHLBI recommendations for rapid pain treatment, and that this lack of awareness was associated with higher perceptions of bias towards patients with SCD, who are primarily Black.⁷⁰ Similarly, reduced awareness and enforcement of policies related to trainings on HIV have been found to be associated with increased provider stigma.⁷¹

Direct person-to-person experiences of discrimination, especially explicit bias, is where much of the existing research on racism and sexism has focused.⁷² These experiences can occur at a variety of critical points along the healthcare continuum from triage to treatment to palliative care.⁷³ SCD patients perceived provider discrimination has been found to be associated with higher pain scores, lower medication adherence, and reduced trust in providers.⁵¹ While these direct experiences of discrimination are important to continue to study, they may not capture the full scope of implicit biases that drive clinical decisions.

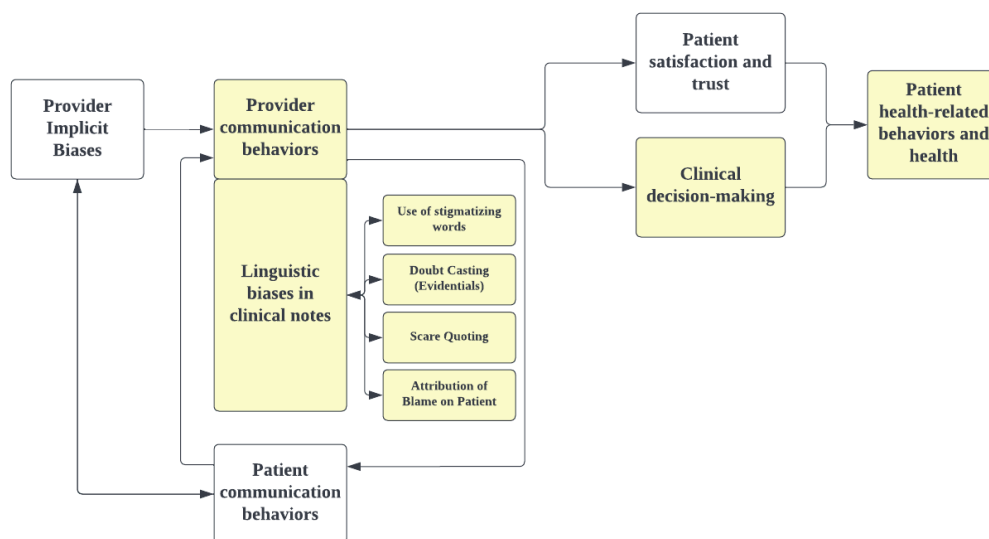
Of critical importance to this work, interactional-level research on discrimination allows for the study of how implicit biases may manifest in patient-provider interactions and/or communication strategies. These may manifest in subtle changes manner or time spent in patient interactions, jokes shared among medical teams which aim to mock patients, or provider mistrust and reduced patient autonomy.^{2,74,75} Often, patient clinical notes are the only record of these interactions, with the narrative controlled entirely by providers.^{21,45} These subtle forms of biases and stigmatizations at the focus of the present study are likely more common and widespread among medical professionals due to their dynamic and concealable nature.⁷⁶

Theoretical framework for provider stigma and bias impacts on health outcomes

Our understanding of how stigma and bias impact health outcomes in clinical settings is guided by ecosocial theory, along with an existing conceptual model posited by Hagiwara and colleagues⁷⁴. Described by Figure 1 below, our model adaptation has operationalized the construct of *provider communication behaviors* as linguistic features in clinical notes, which work to impact patient behaviors and health by affecting patient satisfaction, trust, and clinical

decision-making. Ecosocial theory helps frame the study as one component of multifaceted experiences of bias. Hagiwara's model helps us to focus on the role of provider biased language in charts within the interactional level of the patient-provider relationship. This model also posits that provider communication behaviors are part of a reciprocal feedback loop within patient-provider relationships. In this loop, negative implicit biases held by providers are hypothesized to produce negative provider communication behaviors in patient interactions, which are then met with reduced quality of patient communication. These processes work in tandem to impact patient trust and satisfaction, as well as clinical decision making in the context of the patient's healthcare team. While this conceptual model does not seek to completely eliminate the impacts of higher-level constructs of systemic racism, it provides a useful framework for investigating the role of provider bias and stigmatizing language within the patient-provider relationship, and for planning interventions to improve health outcomes for frequently stigmatized patient populations.

Figure 1: Conceptual model for how stigmatizing language in notes affects health outcomes



How does stigma/bias impact clinical decision-making?

Given the large scope of impact of provider stigma on the population health outcomes, it is imperative to understand the mechanisms of how stigma and biases work to affect clinical interactions, provider communication, and decision-making along the proposed pathway.

One proposed theoretical model, the Dual Process Theory, helps to understand how humans make decisions dependent on contexts and trained behaviors.¹⁴ This theory posits that there are two types of pathways which govern our decision-making process. Type 1 has been defined as the fast, unconscious and intuitive types of decisions, which over time require little cognitive resources. Type 2 processes require more cognitive resources, are slower, and involve a greater degree of analytic and critical thinking. The mental shortcuts and heuristics which define Type 1 make up the majority of our decision-making, despite being most vulnerable to error. When investigating the impact of stigma and biases on provider decision-making, recent studies have identified situations with high “clinical ambiguity”, such as pain assessments, as key windows where provider bias may be utilized as a way to fill in missing information with stigma and biases.¹⁴ In these and other ambiguous situations with no clear clinical answer, Type 1 pathways may irrationally override Type 2, defined as “dysrationalia override”, whenever providers insist on a healthcare plan of their own, ignore evidence-based clinical decision rules, or act on implicit biases and stigmas towards patients.⁷⁷

Providers may also share biases and stigma regarding marginalized patients as a way to promote ingroup favoritism and gain favor among other providers by engaging in outgroup derogation.⁷⁸ Within healthcare provider teams, stigma and provider explicit and implicit biases hold strong influence over patient care decisions,¹⁴ and erode patient perceived trust and quality of provider-patient interactions.^{51,79} Stigma and biases can be actively transmitted through shared language across medical teams through a variety of explicit (patient charts, medical orders) and

implicit (subtle cues, remarks, or insinuations made off record)^{18,80} mediums. These interactional-level instances of bias and stigmatization have the ability to transmit biases across providers, demonstrated in a study that found student providers witnessing racial discrimination during training experiences increased the likelihood of providers expressing stigmatizing attitudes themselves.⁸¹ Similarly, as provider notes encode interactions with patients and are read among provider teams, these stigmas and biases have the ability to become encoded into the medical record and transmitted across providers.²¹ Finally, all of these levels may interact with and exacerbate the previously discussed intra-individual experiences of felt, or self-stigmatization.⁸²

This study will aim to assess the enacted stigma at the interactional and organizational level among providers towards patients as assessed via free-text notes in the electronic health record, and assess potential relationships with health outcomes of patients leaving against medical advice. Provider notes traverse the levels between organizational and interactional discrimination by supporting a shared dialog between providers that may facilitate sharing of stigma and biases towards patients, which influence interactions between providers and patients and drive clinical decision-making.

Clinical Outcomes of Associated with Stigma

Previous research has identified several key outcomes theorized to be linked to provider stigma and implicit biases, including insufficient pain management strategies and increased instances of leaving against medical advice.

Pain Management

Many studies have pointed to disparities in pain management strategies among frequently stigmatized group. Previous research has amassed substantial evidence of widespread reduced opioid prescription rates among Black and Hispanic patients^{37,52,83} and women,^{13,84–86} occurring in nearly all age groups, from pediatrics⁸⁷ to end-of-life care.^{88,89} Due to the inherent subjectivity of patient-reported pain scores, pain management situations are hypothesized to embody the “high clinical ambiguity” that has been posited to be more susceptible to the impact of implicit biases.⁵⁵ The outcomes related to pain management are often related to the type of stigma experienced by patients, and is further described in sections on each of the types of stigma experienced.

Self-directed discharge

Experiences of discrimination and bias by race and socioeconomic status have also been shown to drive higher rates of leaving against medical advice^{58,90}, and increased reluctance to seek care^{47,51,91,92}, which can lead to increased risk of mortality and disease complications across a variety of chronic conditions.^{93,94} While leaving against medical advice happens because of a variety of socioeconomic and individual patient factors, providers factors such as hospital setting and structure and providers’ clinical communication style and experience, along with patient perceptions of trust likely play a significant role.^{95–97} While there has been limited research on provider-level predictors of leaving against medical advice, studies cite “failure to orient the patient to treatment on intake, punitive or threatening atmosphere on the inpatient unit, difficulties in doctor-patient relationship, [and] failure to establish a supportive provider-patient relationship” as common issues that may be directly related to provider biases and stigmas towards marginalized patient populations.⁹⁸

Origins and targets of stigmatization

While stigma and bias in medicine has only recently emerged as a priority of public and clinical health organizations, it has been deeply embedded within clinical science and practice for centuries. Alongside centuries of systemic abuse of non-white, non-male patients, the medical field has struggled with the limits of generalizability of foundational clinical research findings which were and still are largely derived from cohorts of almost entirely White, cis-gendered men, which have largely excluded individuals who have been societally marginalized on the intersectional dimensions of race, gender, socioeconomic, or disease status.^{99,100}

Racial bias and stigma

While a comprehensive history of racial violence and injustices of medicine is beyond the scope of this review, it is important to consider to provide some level of historical context to understand modern provider racial biases in medicine. Medical experimentation and exploitation of Black people, especially Black women, prevailed throughout the 17th, 18th, and 19th centuries.¹⁰¹ Enslaved Black people's bodies were often used as "anatomical material", stolen from graveyards, or coerced into experimental trials, for many American medical colleges well into the 1920s. Medical research in the 19th century largely rationalized slavery,^{102,103} supported by the eugenics movement which espoused the myths of intrinsic biological racial differences and genetic inferiority of Black and other non-White people.¹⁰⁴ Many of these claims were used to assert that Black patients could withstand greater amounts of pain, that their bodies were "tougher" than White bodies, and thus did not need as much pain medication or anesthesia.^{105,106} Despite current awareness of the racist myths perpetuated by the eugenics movement, many of these biases persist subconsciously today in providers and are still embedded into medical

education systems. A 2016 study showed that large proportions of medical students (nearly 50%) believed Black bodies to be biologically different than White bodies, particularly that they are “stronger” or “tougher” than White bodies, and feel less pain.^{17,107} Furthermore, these beliefs are associated with inaccurate pain management strategies and underreporting of patient pain. These biases and stigmas towards patients based on race are often perpetuated by “hidden curriculum”¹⁸ of medical education, and are emboldened for student providers when students witness acts of discrimination or stigmatizing provider behavior.² As recently as 2017, a nursing textbook was removed from publication after receiving public criticism for publishing a section on “cultural differences in responses to pain”, which espoused many racial biases and stigmas related to the experience of pain.¹⁰⁸

The history of racism in the US medical system shows that stigma and biases towards patients due to race have remained persistent, despite changing in forms throughout the years. In spite of the persistence of racism in medicine, there have been many forms of anti-racist resistance cultivated by Black, Latino, Indigenous, and other marginalized racial groups over the years to provide care and protect communities against systemic injustices. The National Medical Association was formed in 1895 as an alliance of African American providers aimed at improving healthcare and medical educational opportunities for Black Americans. The Black Panther movement helped form free healthcare clinics and expand sickle cell disease screenings in the 1970s.¹⁰⁹

Anti-Black racial biases continue to contribute to inequalities in pain management decisions and provider perceptions of patient pain.¹⁷ This effect can be found even in children, where Black children are 60%, and Hispanic children 50% as likely as White children to be prescribed opioids.⁵⁷ Another recent study showed significant disparities in analgesic and opioid

prescription rates in among a cohort of nearly one million children being treated for appendicitis, wherein Black children had 20% the odds of being prescribed opioids to treat severe pain as White children, and only 10% the odds of being prescribed any analgesic medications when presenting with appendicitis.⁷⁶ In order to continue combat the myriad of ways in which systemic violence and racism affects patients today, we must question the structures which perpetuate them and find ways to hold the medical system accountable.

Gender Bias

Of course, provider biases and stigmas are not limited to race. Modern medicine is also struggling to come to terms with the sexist history of abuse, misdiagnosis, and mistreatment of women, who are treated as the “other” to the “normal” of being a cis-gendered man.^{11,110} Women have traditionally been viewed in medicine always in comparison to the male form, with the “complications” of reproduction and hormones. This has led to inequities in treatment and research, where men’s health issues are often treated and investigated more intensively than women.¹¹¹ Women have historically been excluded from major research studies and clinical trials, an issue that persists to this day.¹¹¹ The resulting lack of research, as well as deficits in teaching about gender-specific approaches to clinical evaluation, diseases specific to women, and psychological aspects of health, combined with prevalent societal subordination of women likely contributes to the present-day problem of providers not believing women when they report symptoms.¹¹² Across a variety of conditions, providers distrust women especially when reporting pain and providing pain management.^{11,86} This disbelief has been found to begin as early as adolescence in women,¹³ who report that their pain-related complaints are more frequently dismissed than men. Women also frequently report taking significant efforts in making sure they present themselves and their symptoms in a way that would be most “appropriate” and socially

visible when presenting to providers.⁸⁵ These significant inequities of providers' belief in patients who identify as women have been hypothesized to lead to high amounts of under or mis-diagnosed treatment plans, reduction in provider trust and patient quality of life, and poorer quality of care.¹¹

Socioeconomic/Insurance status

While race and gender play a significant role in driving provider biases, it is important to also consider the role that socioeconomic and insurance status plays in the formation of provider biases which impact patient care. In 1971, Julian Tudor Hart, a renowned British general practitioner, posited the “inverse care law,” which stated that people with the greatest healthcare needs receive the least healthcare.¹¹³ Beyond the significant structural barriers present for low-income patients to access healthcare, it is hypothesized that stigma plays a large role in reducing healthcare access.^{114,115} A recent study examining patients with Affordable Care Act coverage or no insurance coverage reported feelings of perceived stigma derived from poor patient-provider communication, where patients often felt belittled or not respected by providers.¹¹⁴ Similar to the way race and gender impact providers' beliefs around patient pain, recent studies have shown that providers also are less inclined to believe pain from patients of low socioeconomic status.¹⁵

Across biases and stigmas surrounding race, gender, and insurance status, it is clear that there are epistemic injustices taking place where marginalized patient groups are disproportionately disbelieved while reporting pain and other symptoms. This inequity must be addressed in order to hold our medical system accountable to believe and validate the experiences of all patients. Researchers have posited that this level of epistemic injustice, or degree to which marginalized social groups are portrayed as untrustworthy or unbelievable, are

critical drivers of misdiagnoses, reduced quality of life, and trust among patients from a variety of marginalized backgrounds.¹¹⁶

Stigmatized Chronic Illness Populations frequently prescribed opioids

Patients with chronic illnesses like sickle cell disease, HIV/AIDS, OUD (Opioid Use Disorder), and obesity, often experience unique forms of stigmatization and bias from healthcare team members unique to their condition. These stigmas and biases may be especially difficult to bear over the increased volume of healthcare interactions which individuals with chronic illnesses often experience.

Sickle Cell Disease

Patients with SCD, who are primarily Black, require pain medication during vaso-occlusive crises (VOCs), and often experience discrimination from providers of all types centered around disbelief of pain.⁵ During VOCs, SCD patients often require high doses of opioids to treat intense pain crises, which fuels the false assumption that patients with SCD are “Drug Seeking”, or are addicted to opioids.⁷⁰ SCD patients perceived discrimination and provider-based experiences of injustice have been shown to be associated with higher pain scores, lower med adherence, reduced trust.^{51,79,117} Experiences of distrust have also been associated with higher rates of hospital self-discharge, associated with a variety of negative clinical outcomes.⁹⁰ In 2014, the National Heart Lung and Blood Institute issued evidence-based recommendations for treatment of SCD, which emphasized the need to believe patient pain, and induce rapid and appropriate pain treatment during VOCs.¹¹⁸ Despite progress in these recommendations, a 2019 study by Masese and colleagues on barriers and facilitators to proper SCD care showed that only a third of surveyed providers are aware of recommendations, nearly two thirds blame patient behavior as the primary barrier to care. Many surveyed providers

endorsed stigmatization of SCD patients, particularly long-term patients who are dubbed “frequent flyers” or “sicklers”, who develop “drug-seeking” reputations among patient teams and may be less readily prescribed opioids to treat their pain crises.⁷⁰

HIV/AIDS

The ongoing stigmatization associated with HIV/AIDS diagnosis is exemplary of the intersectional nature of biases, which is often wrapped in racial, sexual behaviors and/or minority status, and drug use stigmatization.^{9,16} Patients have reported that during encounters with clinical providers, that stigmatization commonly manifests as providers taking unnecessary or excessive efforts to prevent self-exposure, to using stigmatizing language with them, to denying access to proper treatment or other services.¹¹⁹ Patients with HIV/AIDS who are admitted to the hospital frequently require opioids for chronic pain, with many patients being concurrently treated for opioid use disorder, and thus may be potentially perceived with the stigmas that accompany opioid use.⁸ HIV-related stigma has been found to undermine patient access to care, reduce rates of viral load testing, and reduce treatment adherence of antiretroviral and Pre-Exposure Prophylaxis (PrEP) therapies.^{35,92} While much of the research on HIV/AIDS stigma has focused on patients’ perspectives of HIV-related stigma, provider perspectives have remained understudied.⁹

ODD

Opioid use disorder remains one of the most stigmatizing conditions worldwide.¹²⁰ Recent conservative estimates have indicated there are 2.1 million to have OUD in US, though may not account for many individuals who go underdiagnosed due to stigmas related to seeking

care.¹²¹ People with OUD encounter stigmatization at multiple ecosocial levels, from the political climate surrounding opioid prescriptions and syringe exchange programs, to hospital policies which place the role of drug policy enforcer onto the daily responsibilities of providers. Many people, providers included, believe that people with OUD are culpable for their illness, viewed as “addicts”, or people who are “weak” and exhibit “bad character”.¹²² A national survey of primary care providers 2021 found high levels of stigmatizing attitudes related to OUD. Furthermore, this study found that OUD stigma was negatively associated with the likelihood of prescribing OUD medications, and negatively associated with supporting increased access to Medications for Opioid Use Disorder (MOUD) such as buprenorphine and methadone.¹²³ Much of the way provider stigma is experienced by people with OUD happens through language, which often dehumanizes people with OUD and encourages discrimination, which negatively affect treatment outcomes among people with OUD.¹²⁴

Obesity

As rates of obesity have risen in the US, patients with obesity report worse communication with physicians via perceiving less respect and attention from providers, and spending less time with physicians.¹²⁵ It has been hypothesized that this reduction in time spent with patients has significant impacts on rapport or relationship-building, which may further drive weight-based stigmatization.^{125,126} The common stereotype among medical providers is that patients with obesity have little willpower and will fail to adhere to provider recommendations on behavior change to lose weight, despite the numerous genetic, environmental, and socio ecological factors that may have led to patients having obesity.¹²⁷

Research has found that patients with obesity often experience a greater prevalence of explicit stigmatization, often experienced as derogatory statements and “cynical humor” directed at patients, who are often described by providers as deserving of the ridicule due to their culpability.¹²⁸ These jokes create environments that perpetuate stereotypes and negative attitudes and harm patients.¹²⁷ Consequences of weight-based stigma and biases include deterioration of provider-patient relationships and trust, reduced adherence to medical plans, and delay of care among patients experiencing weight-based bias by their providers.^{127,129,130}

Current approaches to measuring Stigma and Bias

One of the primary research challenges in the field of measuring stigma and provider bias and its effect on patient health outcomes is the difficulty in measuring stigma and bias as it plays out in real world interactions. Current methodologies have demonstrated that providers of many types, surveyed internationally, uphold a variety of implicit biases towards stigmatized groups. While methods have been used to show how these biases may impact hypothetical patient care decision-making, there is a lack of research into the mechanisms through which biases are transmitted and enacted which lead to patient experiences of stigmatization and disparities in care.

Survey methods

In the study of provider biases and stigmas, self-reported survey methods have been successfully employed to assess explicit biases, attitudes, and beliefs related to race, drug use, and obesity.^{131–133} However, survey methods aimed to assess provider-held biases and stigmas suffer from the potential biases of self-report and halo effects, where providers are often hesitant to endorse stigmatized or biased beliefs, resulting in underreporting. These surveys may be able

to use deception to mask the intent of the survey, but as a whole they typically also only are able to evaluate conscious explicit mental processes, and may not be able to capture the subtleties of implicit or automatic biases.¹ Additionally, similar to the challenges posed by asking participants to report experiences of lifetime discrimination in a consistent way which was assessed by Krieger and Van Dyke, it is a difficult task to accurately and truthfully recall the role one has played and may actively play in perpetuating stigma and biases.^{134,135} This is particularly difficult given the proposed inverse relationship between knowledge of bias and stigma and willingness to believe in stigmatizing and biased beliefs.⁷⁶ Because of this, studies focused on provider bias and stigma are at risk of misclassification of the exposure, and underestimation of the effect of discrimination on health outcomes. It is imperative for the study of provider bias and stigma to use alternative sources of data to measure provider stigma and bias.

Administrative Data

In light of the challenges posed by survey data, many researchers of bias in healthcare have relied upon simplified constructions of race, gender, or insurance status which are readily available in most EHR datasets.¹³⁶ While these studies have offered key insights by identifying healthcare inequities and guiding research efforts, this approach lacks critical information on the mechanisms of bias and stigma, and may lead to indirect inferences of stigma and bias without specifically seeking to measure either.^{72,137}

Vignette studies

Vignette studies, which utilize hypothetical narrative patient descriptions and assess provider attitudes towards patients and projected clinical decisions, have helped to illuminate the potential impact of stigma and biases on patient healthcare outcomes. Goddu and colleagues tested this method experimentally by randomly assigning medical students into groups receiving two nearly-identical narratives of a 28 year-old with sickle cell disease, with one narrative including stigmatizing linguistic features, and the control condition assigned to read a narrative with neutral descriptions of the patient.¹³⁸ Providers in the stigmatized language condition were more likely endorse negative attitudes toward the patient and significantly less aggressive pain management strategies.¹³⁸ Despite the potential vignette studies to provide insights into provider attitudes and behavior intentions, there are still high possibilities that respondents are on “best behavior”, so methods should be taken in order to still attempt to deceive the exact intent of the study, such as implementing participant and researcher blinding strategies to avoid artificial adaptations to provider participant behaviors.¹ Additionally, many studies utilizing vignette approaches to study stigma and bias have relied on samples of lay or pre-clinical medical students, which may not reflect real-world clinical decision-making.^{107,139} Finally, these methods still focus on behavioral intention and have yet been unable to capture actual provider behavior in a way that could be deployed to evaluate and intervene on real-time manifestations of provider bias and stigmatization.

Implicit Association Tests (IATs)

Another potential solution to the study of stigma and bias has been pursued through the application of the Implicit Association Test (IAT).^{2,140} The IAT was created in 1998, and involves

the process of participants being asked to complete timed computerized exams which have participants categorize certain words based on their known stereotypes. Participant response rates are then measured, as well as their % of word attributions in line with pre-existing stereotype conditions.¹⁴⁰ This process has been thoroughly validated and implemented extensively over the last near 15-years of use, and has been shown to effectively predict stigma and biases in the forms of racism, sexism, ageism, and homophobia.² IAT scores have often been compared with provider patient care attitudes or perceptions that are hypothesized to be relevant to care delivery outcomes.¹ A recent review of studies using the IAT found that increased provider implicit biases against Black patients were associated with reduced feelings of respect, and warmth, and collaboration reported by Black patients, as well as reduced adherence to treatment plans, reduced likelihood of being recommended necessary interventions, and worse psychosocial health outcomes such as depression and life satisfaction.³ The majority of studies utilizing the IAT to investigate provider bias and its impact on health outcomes found providers to harbor low to moderate anti-Black bias, though studies assessing anti-Hispanic/Latino/Latina bias have suggested similar sized moderate negative biases held among providers.³ While the IAT has been one of the most widely known assessment tools for implicit bias, its validity has been widely debated, particularly around what other factors or constructs besides implicit biases may be influencing respondent decisions.¹⁴¹ Further, IAT studies typically involve recruitment and active participation of providers in contexts outside of their typical care environment, and may still not accurately reflect real-time provider stigma and bias as they are activated to impact care decisions for patients.

Solution: Mining patient care notes to identify bias and stigma

With limitations of current approaches to studying stigma and bias in mind, the richness of EHR free text clinical data offers new opportunities to study provider stigma and bias. EHR clinical notes can allow for the analysis of longitudinal care histories of patients, accumulation of stigmatized reputations, and examine evidence of stigmatization as it is transmitted within the medical record shared across provider teams.

Clinical notes play a key role in the transmission of information and attitudes towards patients that impact decision-making from one provider to another.¹⁸ Through this transmission of information, patients can quickly develop “reputations” among the provider team, and different perceptions of patients can take on a self-fulfilling momentum of their own.⁴ Within the context of clinical charts, specific features of language commonly used by providers have been identified that may perpetuate biases across providers and impact clinical decision-making through the induction of doubt, mistrust, and stigmatizing language about patients and the validity of their reported pain or symptoms.^{19–21,142} Charts of black patients were found by Beach and colleagues to have 48% higher odds of containing inappropriate “scare quotes”, and 32% higher odds of including evidential words, both of which have been studied as linguistic features used to specifically connote distrust or invalidation of point-of-view.²¹

Research on the transmission of provider bias through clinical notes has identified many common manifestations of both negative and positive bias within patient charts²⁰, and early research into this text and language-based transmission of bias towards patients has found associations with disparities in care across race, gender, drug use, and other marginalized conditions.^{1,20,21,143}

Linguistic manifestations of bias and stigmatization: The Social Categories and Stereotype Communication Framework

Prior researchers have begun to apply linguistic approaches to assessing numerous manifestations of stigma and bias as they are expressed in language, which provides a rich foundation of approaches to assess provider stigma in clinical notes.

Linguistic bias has been defined as “a systematic asymmetry in word choice as a function of the social category to which the target belongs”.⁴¹ This definition borrows directly from Link and Phelan’s definition of bias. Although there are a variety of ways in which linguistic bias can manifest,¹⁴⁴ they all facilitate the transmission of “essentialist beliefs about social categories”, or work to suggest and reinforce that members of particular groups contain inherent shared characteristics making them fundamentally distinct from others across situations.

Beukeboom and Burgers recently developed a framework to understand how stigmatization can manifest and be formed through language, via the Social Categories and Stereotype Communication Framework.¹⁴⁵ This framework posits that stereotypes are communicated through systematic differences in how other groups are 1) labeled, and 2) how their behaviors and characteristics are described. Recent research investigating linguistic bias in the medical record has identified salient features of bias within the language of provider notes, namely focused on the expression of degree of trust or doubt in patient testimony through quotations and evidential use, as well as the usage of stigmatizing and negative descriptor language.^{21,138,146} In applying Beukeboom and Burgers’ framework towards the study of stigma in the electronic health record, we position stigmatizing labels, which also includes negative descriptors as ways in which patients may be differentially labeled, and doubt markers and scare

quotes as ways in which the behaviors and characteristics of patients may be differentially described.

Stigmatizing labels and negative descriptors

Stigmatizing labels to describe groups are often used to perpetuate stereotypes, and when used by providers, can lead to feelings of stigmatization and reduced trust among their patients. Much of the recent work on identifying and reducing stigmatizing labels has come from providers seeking to improve care for patients with substance use disorders. A recent NIDA study published a list of words to avoid using around patients with substance use disorders, including "addict", "abuser", "user", or "junkie", which have been found to be associated with perceived stigmatization by patients.¹⁴⁷ Similar studies have been applied to other chronic illness populations, identifying terms like "sickler" or "frequent flier" which may be used to further stigmatize patients with chronic illnesses who are often admitted into the hospital.^{138,148,149} While some providers may argue that these terms may be useful in flagging unwanted patient behaviors or mental states, a recent study has shown that patients exposed to language written about them by providers which included stigmatizing labels resulted in patients feeling unfairly judged, labeled, and disrespected.¹⁵⁰

Recent research led by Michael Sun and colleagues on over 40,000 clinical notes has found disparities in presence of "Negative Descriptor" words, evaluated by the Health Equity Commission of the Society of General Internal Medicine, which included commonly used terms in the medical record such as "(non-)adherent, aggressive, agitated, angry, challenging, combative, (non-)compliant, confront, (non-)cooperative, defensive, exaggerate, hysterical, (un-)pleasant, refuse, and resist". This study found that compared to White patients, Black patients had 2.54 times the odds of having at least one negative descriptor written in their history

and physical notes.¹⁵¹ Research into stereotype expression in language has found that even seemingly innocuous category labels may prompt others to perceive target individual actions and characteristics as “static” aspects of their identity, and exaggerate differences across groups and similarities within them.⁴¹ These labels can be used to justify clinical decision-making, withholding of resources, or to confer doubt upon patient testimonies.¹⁴⁴ While current recommendations encourage use of person-centered, neutral language in medical charts, it is important to evaluate the presence of known stigmatizing labels within provider notes to mitigate the transmission of bias in the medical record.

Evidentials and markers of doubt

Linguistic features such as evidentials, defined as “the linguistic coding of epistemology”,¹⁵² are words that are frequently used in chart language to question the veracity of patients. Among the many words used as evidentials, words and expressions used to confer doubt or uncertainty such as: *allegedly*, *apparently*, or verbs like *claimed*, are often used to when describing patient testimonies, for example: “patient *claimed* their pain was 10/10”.²¹ Providers may use words when describing patient testimony in combination with stigmatizing labels or negative descriptors of patients to transmit their stance, or expression of attitudes, feelings, and judgment about patients to other providers which may impact future treatment and care decisions.¹⁵³ Disparities have been found among usage of these terms across race and gender, where patients who were women and patients who were Black were found to have significantly higher evidentials in their provider notes than patients who were men or White.²¹ The extent to which providers use evidentials is posited to be reflective of the amount of doubt and uncertainty a provider has on patient testimony, and is thus hypothesized to impact trust within the

patient-provider relationship, and related outcomes like leaving against medical advice or in pain management strategies.^{20,21}

Scare Quotes

Another linguistic marker of uncertainty that has been previously identified in patient charts are “scare quotes”, which involve the utilization of quotation marks to mock, cast doubt, challenge patient credibility, or insinuate low health literacy when describing the testimony of another individual.²¹ While use of scare quotes has been documented since the 1950s, some argue that the linguistic phenomenon has been increasing increasing in recent years, both from the rise of “air quoting” gesture in the 80s and 90s, and was commonly employed by Trump prior to and throughout his presidency.^{154,155}

While quotations in charts can be useful to describe patient symptoms using their exact language and document patient wishes or concerns, recent linguistic research has identified a troubling prevalence of providers utilizing quotations in ways to mock, manipulate, and regulate the voices of patients. For example, consider the ambiguity added to the sentence: “Patient reports 10/10 pain related to sickle cell crisis.”, when you add “Patient reports ‘10/10’ pain related to ‘sickle cell crisis’”. Because of the quotation marks, both 10/10 and sickle cell crisis could be inferred as being untrue or uncertain. Similar to evidentials and negative patient descriptors, scare quotes have been found to be more prevalent among patients who were Black and among patients who were women.²¹ Early research on scare quoting in patient charts has recommended that providers utilize quotes only when absolutely necessary to reflect a patient’s exact experience, wishes, or concerns, and that even when made in earnest, unnecessary quotation opens patient testimony up to a level of uncertainty or inference to other providers who may question the veracity of patient conditions or experiences.¹⁵⁶

Natural Language Processing to identify stigma and bias in EHR Data

While linguistic studies have acted to guide researchers into identifying manifestations of stigma and bias in text, due to the nature of qualitative, in-depth assessments, these methods can have significant limitations in being able to be deployed to rapidly identify stigma and bias in medical notes in a way in which we can intervene on it and hold providers accountable. Recent advances in computational linguistics are allowing researchers to harness the human-annotated insights on linguistic bias and stigma and scale up categorization to larger amounts of unlabeled data.

Advances in transformer-based language learning models and contextual word embeddings have demonstrated high model performance in categorization tasks using large amounts of medical data, like that of MIMIC-III datasets.¹⁵⁷ Transformer-based NLP prediction models have been successfully applied to unstructured clinical notes within EHR systems towards tasks involving symptom extraction¹⁵⁸, and clinical relationship extraction,¹⁵⁹ which can be used to extract key behavioral, social determinant of health, and adverse drug event information only encoded in unstructured notes. Transformer-based methods have also been applied in detecting biased text within user reviews or in detecting biased text in Wikipedia articles.^{24,146}

The power of many of these transformer-based models lie in their ability to “pay attention” to specific words and word contexts of any given input text.¹⁶⁰ This allows models to more accurately disambiguate word sense, or differentiate meanings of the same word which change given context, for example: “I deposited money at the bank” versus “I climbed up the river bank”. This increased ability to be able to identify important differences in these sentences

allows for more sophisticated internal models of language and higher accuracy across many NLP-related tasks, from translation, to recognizing word dependencies, to categorization of text.

Much of the existing research on stigma and bias in language has been rooted in linguistics and pursued through qualitative studies. Current approaches are extremely time and labor intensive, drawing on expertise of trained human annotators to sample and identify bias in medical notes. The ability to accurately detect bias in patient charts in real-time could offer healthcare teams opportunities to evaluate and intervene in biased chart language. This task could be made possible through the development of scalable, computationally intensive systems to keep up with the high-volume data from live clinical provider teams.¹⁶¹

In combination with utilizing state of the art modeling methods, computational linguistics researchers in the social sciences have recently called for the utilization of “Grounded Computational Theory” in order to maintain the role of human expert annotation in the loop to evaluate model output and processes in major NLP tasks, particularly related to those aimed at impacting human well-being and health.¹⁶² These methods have been advocated for in order to mitigate the potential harmful effects which ML/AI algorithms may inadvertently have on the populations they intend to support.

This project aims to apply advanced methods in natural language processing to detect and assess the presence of negatively biased language in provider clinical notes for patients with chronic illnesses who receive opioid pain medications during inpatient hospitalizations. This project will extend qualitative and NLP methods using Grounded Computational Theory to detect provider stigma bias in clinical notes in an effort to increase the accountability, evaluation, and intervention of negative bias that threatens to deteriorate quality of care for patients among a variety of uniquely stigmatized groups.

Introductory Literature Review References

1. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. *BMC Medical Ethics*. 2017;18(1):19. doi:10.1186/s12910-017-0179-8
2. Maina IW, Belton TD, Ginzberg S, Singh A, Johnson TJ. A decade of studying implicit racial/ethnic bias in healthcare providers using the implicit association test. *Social Science & Medicine*. 2018;199:219-229. doi:10.1016/j.socscimed.2017.05.009
3. Hall WJ, Chapman MV, Lee KM, et al. Implicit Racial/Ethnic Bias Among Health Care Professionals and Its Influence on Health Care Outcomes: A Systematic Review. *Am J Public Health*. 2015;105(12):e60-e76. doi:10.2105/AJPH.2015.302903
4. Fontana AF. Patient Reputations: Manipulator, Helper, and Model. *Archives of General Psychiatry*. 1971;25(1):88-93. doi:10.1001/archpsyc.1971.01750130090011
5. Jenerette CM, Brewer C. Health-Related Stigma in Young Adults With Sickle Cell Disease. *Journal of the National Medical Association*. 2010;102(11):1050-1055. doi:10.1016/S0027-9684(15)30732-X
6. Jenerette CM, Brewer CA, Ataga KI. Care Seeking for Pain in Young Adults with Sickle Cell Disease. *Pain Management Nursing*. 2014;15(1):324-330. doi:10.1016/j.pmn.2012.10.007
7. Haywood C, Tanabe P, Naik R, Beach MC, Lanzkron S. The impact of race and disease on sickle cell patient wait times in the emergency department. *Am J Emerg Med*. 2013;31(4):651-656. doi:10.1016/j.ajem.2012.11.005
8. Cunningham CO. Opioids and HIV Infection: From Pain Management to Addiction

- Treatment. *Top Antivir Med.* 2018;25(4):143-146.
9. Geter A, Herron AR, Sutton MY. HIV-Related Stigma by Healthcare Providers in the United States: A Systematic Review. *AIDS Patient Care and STDs.* 2018;32(10):418-424. doi:10.1089/apc.2018.0114
 10. van Boekel LC, Brouwers EPM, van Weeghel J, Garretsen HFL. Stigma among health professionals towards patients with substance use disorders and its consequences for healthcare delivery: Systematic review. *Drug and Alcohol Dependence.* 2013;131(1):23-35. doi:10.1016/j.drugalcdep.2013.02.018
 11. Govender V, Penn-Kekana L. Gender biases and discrimination: a review of health care interpersonal interactions. *Global Public Health.* 2008;3(sup1):90-103. doi:10.1080/17441690801892208
 12. Krieger N. Measures of Racism, Sexism, Heterosexism, and Gender Binarism for Health Equity Research: From Structural Injustice to Embodied Harm—An Ecosocial Analysis. <https://doi.org/10.1146/annurev-publhealth-040119-094017>. doi:10.1146/annurev-publhealth-040119-094017
 13. Iglar EC, Defenderfer EK, Lang AC, Bauer K, Uihlein J, Davies WH. Gender differences in the experience of pain dismissal in adolescence. *J Child Health Care.* 2017;21(4):381-391. doi:10.1177/1367493517727132
 14. Hirsh AT, Hollingshead NA, Ashburn-Nardo L, Kroenke K. The Interaction of Patient Race, Provider Bias, and Clinical Ambiguity on Pain Management Decisions. *The Journal of Pain.* 2015;16(6):558-568. doi:10.1016/j.jpain.2015.03.003
 15. Summers KM, Deska JC, Almaraz SM, Hugenberg K, Lloyd EP. Poverty and pain: Low-SES people are believed to be insensitive to pain. *Journal of Experimental Social*

- Psychology*. 2021;95:104116. doi:10.1016/j.jesp.2021.104116
16. Whittle HJ, Palar K, Ranadive NA, Turan JM, Kushel M, Weiser SD. “The land of the sick and the land of the healthy”: Disability, bureaucracy, and stigma among people living with poverty and chronic illness in the United States. *Social Science & Medicine*. 2017;190:181-189. doi:10.1016/j.socscimed.2017.08.031
 17. Hoffman KM, Trawalter S, Axt JR, Oliver MN. Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *PNAS*. 2016;113(16):4296-4301. doi:10.1073/pnas.1516047113
 18. Hafferty FW, Franks R. The hidden curriculum, ethics teaching, and the structure of medical education. *Academic Medicine*. 1994;69(11):861-871.
 19. P. Goddu A, O’Conor 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
 20. Park J, Saha S, Chee B, Taylor J, Beach MC. Physician Use of Stigmatizing Language in Patient Medical Records. *JAMA Network Open*. 2021;4(7):e2117052-e2117052. doi:10.1001/jamanetworkopen.2021.17052
 21. 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*. Published online March 22, 2021. doi:10.1007/s11606-021-06682-z
 22. Fodeh SJ, Finch D, Bouayad L, et al. Classifying clinical notes with pain assessment using machine learning. *Med Biol Eng Comput*. 2018;56(7):1285-1292. doi:10.1007/s11517-017-1772-1
 23. Sheikhalishahi S, Miotto R, Dudley JT, Lavelli A, Rinaldi F, Osmani V. Natural Language

- Processing of Clinical Notes on Chronic Diseases: Systematic Review. *JMIR Medical Informatics*. 2019;7(2):e12239. doi:10.2196/12239
24. Recasens M, Danescu-Niculescu-Mizil C, Jurafsky D. Linguistic models for analyzing and detecting biased language. In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. ; 2013:1650-1659.
 25. Johnson, Alistair, Pollard, Tom, Mark, Roger. MIMIC-III Clinical Database. Published online 2015. doi:10.13026/C2XW26
 26. Hatzenbuehler ML, Phelan JC, Link BG. Stigma as a Fundamental Cause of Population Health Inequalities. *Am J Public Health*. 2013;103(5):813-821. doi:10.2105/AJPH.2012.301069
 27. Biello KB, Hughto JMW. Measuring Intersectional Stigma Among Racially and Ethnically Diverse Transgender Women: Challenges and Opportunities. *Am J Public Health*. 2021;111(3):344-346. doi:10.2105/AJPH.2020.306141
 28. Narayan MC. CE: Addressing Implicit Bias in Nursing: A Review. *AJN The American Journal of Nursing*. 2019;119(7):36-43. doi:10.1097/01.NAJ.0000569340.27659.5a
 29. Fujisawa D, Hagiwara N. Cancer Stigma and its Health Consequences. *Curr Breast Cancer Rep*. 2015;7(3):143-150. doi:10.1007/s12609-015-0185-0
 30. Institute of Medicine. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. (Smedley BD, Stith AY, Nelson AR, eds.). The National Academies Press; 2003. doi:10.17226/10260
 31. Cooper LA, Saha S, van Ryn M. Mandated Implicit Bias Training for Health Professionals—A Step Toward Equity in Health Care. *JAMA Health Forum*. 2022;3(8):e223250. doi:10.1001/jamahealthforum.2022.3250

32. Link BG, Phelan JC. Conceptualizing Stigma. *Annu Rev Sociol.* 2001;27(1):363-385.
doi:10.1146/annurev.soc.27.1.363
33. Goodyear-Smith F, Buetow S. Power issues in the doctor-patient relationship. *Health Care Anal.* 2001;9(4):449-462. doi:10.1023/A:1013812802937
34. Beisecker AE. Patient Power in Doctor-Patient Communication: What Do We Know? *Health Communication.* 1990;2(2):105-122. doi:10.1207/s15327027hc0202_4
35. Penner LA, Phelan SM, Earnshaw V, Albrecht TL, Dovidio JF. Patient stigma, medical interactions, and health care disparities: A selective review. In: *The Oxford Handbook of Stigma, Discrimination, and Health.* Oxford library of psychology. Oxford University Press; 2018:183-201.
36. Krieger N. Theories for social epidemiology in the 21st century: an ecosocial perspective. *International Journal of Epidemiology.* 2001;30(4):668-677. doi:10.1093/ije/30.4.668
37. Pentin PL. Drug Seeking or Pain Crisis? Responsible Prescribing of Opioids in the Emergency Department. *AMA Journal of Ethics.* 2013;15(5):410-415.
doi:10.1001/virtualmentor.2013.15.5.ecas2-1305
38. McCaffery M, Grimm MA, Pasero C, Ferrell B, Uman GC. On the Meaning of “Drug Seeking.” *Pain Management Nursing.* 2005;6(4):122-136. doi:10.1016/j.pmn.2005.08.002
39. Bourdieu P. What Makes a Social Class? On The Theoretical and Practical Existence Of Groups. *Berkeley Journal of Sociology.* 1987;32:1-17.
40. Caliskan A, Lewis M. Social biases in word embeddings and their relation to human cognition. Published online 2020. doi:10.31234/osf.io/d84kg
41. Beukeboom CJ, Burgers C. Linguistic Bias. In: *Oxford Research Encyclopedia of Communication.* Oxford University Press; 2017.

doi:10.1093/acrefore/9780190228613.013.439

42. Perry BL, Felix E, Bolton M, Pullen EL, Pescosolido BA. Public Stigma and Personal Networks: Confronting the Limitations of Unidimensional Measures of Social Contact. *J Health Soc Behav*. Published online February 28, 2022:00221465211072311.
doi:10.1177/00221465211072311
43. Bargh JA. The cognitive monster: The case against the controllability of automatic stereotype effects. In: *Dual-Process Theories in Social Psychology*. The Guilford Press; 1999:361-382.
44. Jackson-Best F, Edwards N. Stigma and intersectionality: a systematic review of systematic reviews across HIV/AIDS, mental illness, and physical disability. *BMC Public Health*. 2018;18(1):919. doi:10.1186/s12889-018-5861-3
45. Buchman DZ, Ho A, Goldberg DS. Investigating Trust, Expertise, and Epistemic Injustice in Chronic Pain. *Bioethical Inquiry*. 2017;14(1):31-42. doi:10.1007/s11673-016-9761-x
46. Hall NY, Le L, Majmudar I, Mihalopoulos C. Barriers to accessing opioid substitution treatment for opioid use disorder: A systematic review from the client perspective. *Drug and Alcohol Dependence*. 2021;221:108651. doi:10.1016/j.drugalcdep.2021.108651
47. Brennan-Cook J, Bonnabeau E, Aponte R, Augustin C, Tanabe P. Barriers to Care for Persons with Sickle Cell Disease: The Case Manager's Opportunity to Improve Patient Outcomes. *Prof Case Manag*. 2018;23(4):213-219. doi:10.1097/NCM.0000000000000260
48. Cockroft JD, Adams SM, Bonnet K, Matlock D, McMillan J, Schlundt D. "A scarlet letter": Stigma and other factors affecting trust in the health care system for women seeking substance abuse treatment in a community setting. *null*. 2019;40(2):170-177.
doi:10.1080/08897077.2018.1544184

49. Boag W, Suresh H, Celi LA, Szolovits P, Ghassemi M. Racial Disparities and Mistrust in End-of-Life Care. *arXiv:180803827 [stat]*. Published online August 15, 2018. Accessed September 6, 2021. <http://arxiv.org/abs/1808.03827>
50. Martin SR, Cohen LL, Mougianis I, Griffin A, Sil S, Dampier C. Stigma and Pain in Adolescents Hospitalized for Sickle Cell Vasooocclusive Pain Episodes. *The Clinical Journal of Pain*. 2018;34(5):438-444. doi:10.1097/AJP.0000000000000553
51. Haywood C, Lanzkron S, Bediako S, et al. Perceived Discrimination, Patient Trust, and Adherence to Medical Recommendations Among Persons with Sickle Cell Disease. *J Gen Intern Med*. 2014;29(12):1657-1662. doi:10.1007/s11606-014-2986-7
52. Keister LA, Stecher C, Aronson B, McConnell W, Hustedt J, Moody JW. Provider Bias in prescribing opioid analgesics: a study of electronic medical Records at a Hospital Emergency Department. *BMC Public Health*. 2021;21(1):1-9. doi:10.1186/s12889-021-11551-9
53. Burgess DJ, Van Ryn M, Crowley-Matoka M, Malat J. Understanding the Provider Contribution to Race/Ethnicity Disparities in Pain Treatment: Insights from Dual Process Models of Stereotyping. *Pain Medicine*. 2006;7(2):119-134. doi:10.1111/j.1526-4637.2006.00105.x
54. Singhal A, Tien YY, Hsia RY. Racial-Ethnic Disparities in Opioid Prescriptions at Emergency Department Visits for Conditions Commonly Associated with Prescription Drug Abuse. *PLOS ONE*. 2016;11(8):e0159224. doi:10.1371/journal.pone.0159224
55. Ahluwalia SC, Giannitrapani KF, Dobscha SK, Cromer R, Lorenz KA. “Sometimes you wonder, is this really true?”: Clinician assessment of patients’ subjective experience of pain. *Journal of Evaluation in Clinical Practice*. 2020;26(3):1048-1053.

doi:10.1111/jep.13298

56. Carr DB. Patients with Pain Need Less Stigma, Not More. *Pain Medicine*. 2016;17(8):1391-1393. doi:10.1093/pm/pnw158
57. Groenewald CB, Rabbitts JA, Hansen E, Palermo TM. Racial differences in opioid prescribing for children in the United States. *Pain*. 2018;159(10):2050-2057. doi:10.1097/j.pain.0000000000001290
58. Simon R, Snow R, Wakeman S. Understanding why patients with substance use disorders leave the hospital against medical advice: A qualitative study. *Substance Abuse*. 2020;41(4):519-525. doi:10.1080/08897077.2019.1671942
59. Nachegea JB, Morroni C, Zuniga JM, et al. HIV-Related Stigma, Isolation, Discrimination, and Serostatus Disclosure: A Global Survey of 2035 HIV-Infected Adults. *Journal of the International Association of Physicians in AIDS Care*. 2012;11(3):172-178. doi:10.1177/1545109712436723
60. O'Donnell AT, O'Carroll T, Toole N. Internalized Stigma and Stigma-Related Isolation Predict Women's Psychological Distress and Physical Health Symptoms Post-Abortion. *Psychology of Women Quarterly*. 2018;42(2):220-234. doi:10.1177/0361684317748937
61. Hatzenbuehler ML, Nolen-Hoeksema S, Dovidio J. How Does Stigma "Get Under the Skin"? The Mediating Role of Emotion Regulation. *Psychol Sci*. 2009;20(10):1282-1289. doi:10.1111/j.1467-9280.2009.02441.x
62. Meyer IH. Resilience in the study of minority stress and health of sexual and gender minorities. *Psychology of Sexual Orientation and Gender Diversity*. 2015;2:209-213. doi:10.1037/sgd0000132
63. Gyll M, Matthews KA, Bromberger JT. Discrimination and unfair treatment: relationship

- to cardiovascular reactivity among African American and European American women. *Health Psychol.* 2001;20(5):315-325. doi:10.1037//0278-6133.20.5.315
64. Townsend SSM, Major B, Gangi CE, Mendes WB. From “In the Air” to “Under the Skin”: Cortisol Responses to Social Identity Threat. *Pers Soc Psychol Bull.* 2011;37(2):151-164. doi:10.1177/0146167210392384
65. Link B, Hatzenbuehler ML. Stigma as an Unrecognized Determinant of Population Health: Research and Policy Implications. *Journal of Health Politics, Policy and Law.* 2016;41(4):653-673. doi:10.1215/03616878-3620869
66. Krieger N. Got Theory? On the 21st c. CE Rise of Explicit use of Epidemiologic Theories of Disease Distribution: A Review and Ecosocial Analysis. *Curr Epidemiol Rep.* 2014;1(1):45-56. doi:10.1007/s40471-013-0001-1
67. Hughes LD, Gamarel KE, King WM, Goldenberg T, Jaccard J, Geronimus AT. State-Level Policy Stigma and Non-Prescribed Hormones Use among Trans Populations in the United States: A Mediation Analysis of Insurance and Anticipated Stigma. *Annals of Behavioral Medicine.* 2022;56(6):592-604. doi:10.1093/abm/kaab063
68. Allen B, Harocopos A, Chernick R. Substance Use Stigma, Primary Care, and the New York State Prescription Drug Monitoring Program. *Behavioral Medicine.* 2020;46(1):52-62. doi:10.1080/08964289.2018.1555129
69. Oliva JD. Dosing Discrimination: Regulating PDMP Risk Scores. Published online January 18, 2021. doi:10.2139/ssrn.3768774
70. Masese R, Bulgin D, Douglas C, Shah N, Tanabe P. Barriers and facilitators to care for individuals with sickle cell disease in central North Carolina: The emergency department providers’ perspective. *PLOS ONE.* 2019;14:e0216414. doi:10.1371/journal.pone.0216414

71. Stringer KL, Turan B, McCormick L, et al. HIV-Related Stigma Among Healthcare Providers in the Deep South. *AIDS Behav.* 2016;20(1):115-125.
doi:10.1007/s10461-015-1256-y
72. Krieger N. Discrimination and health inequities. *Int J Health Serv.* 2014;44(4):643-710.
doi:10.2190/HS.44.4.b
73. Hunt J, Bristowe K, Chidyamatara S, Harding R. ‘So isolation comes in, discrimination and you find many people dying quietly without any family support’: Accessing palliative care for key populations – an in-depth qualitative study. *Palliat Med.* 2019;33(6):685-692.
doi:10.1177/0269216319835398
74. Hagiwara N, Lafata JE, Mezuk B, Vrana SR, Fetters MD. Detecting implicit racial bias in provider communication behaviors to reduce disparities in healthcare: Challenges, solutions, and future directions for provider communication training. *Patient Educ Couns.* 2019;102(9):1738-1743. doi:10.1016/j.pec.2019.04.023
75. Fruh SM, Nadglowski J, Hall HR, Davis SL, Crook ED, Zlomke K. Obesity Stigma and Bias. *The Journal for Nurse Practitioners.* 2016;12(7):425-432.
doi:10.1016/j.nurpra.2016.05.013
76. Gopal DP, Chetty U, O’Donnell P, Gajria C, Blackadder-Weinstein J. Implicit bias in healthcare: clinical practice, research and decision making. *Future Healthc J.* 2021;8(1):40-48. doi:10.7861/fhj.2020-0233
77. Ely JW, Graber ML, Croskerry P. Checklists to Reduce Diagnostic Errors. *Academic Medicine.* 2011;86(3):307-313. doi:10.1097/ACM.0b013e31820824cd
78. Paradies Y, Truong M, Priest N. A Systematic Review of the Extent and Measurement of Healthcare Provider Racism. *J GEN INTERN MED.* 2014;29(2):364-387.

doi:10.1007/s11606-013-2583-1

79. Elander J, Beach MC, Jr CH. Respect, trust, and the management of sickle cell disease pain in hospital: comparative analysis of concern-raising behaviors, preliminary model, and agenda for international collaborative research to inform practice. *Ethnicity & Health*. 2011;16(4-5):405-421. doi:10.1080/13557858.2011.555520
80. Fallin-Bennett K. Implicit Bias Against Sexual Minorities in Medicine: Cycles of Professional Influence and the Role of the Hidden Curriculum. *Academic Medicine*. 2015;90(5):549-552. doi:10.1097/ACM.0000000000000662
81. Burke SE, Dovidio JF, Perry SP, et al. Informal Training Experiences and Explicit Bias against African Americans among Medical Students. *Soc Psychol Q*. 2017;80(1):65-84. doi:10.1177/0190272516668166
82. Rivera AV, DeCuir J, Crawford ND, Amesty S, Lewis CF. Internalized stigma and sterile syringe use among people who inject drugs in New York City, 2010-2012. *Drug Alcohol Depend*. 2014;144:259-264. doi:10.1016/j.drugalcdep.2014.09.778
83. Morden NE, Chyn D, Wood A, Meara E. Racial Inequality in Prescription Opioid Receipt — Role of Individual Health Systems. *New England Journal of Medicine*. 2021;385(4):342-351. doi:10.1056/NEJMsa2034159
84. Werner A, Isaksen LW, Malterud K. 'I am not the kind of woman who complains of everything': Illness stories on self and shame in women with chronic pain. *Social Science & Medicine*. 2004;59(5):1035-1045. doi:10.1016/j.socscimed.2003.12.001
85. Werner A, Malterud K. It is hard work behaving as a credible patient: encounters between women with chronic pain and their doctors. *Social Science & Medicine*. 2003;57(8):1409-1419. doi:10.1016/S0277-9536(02)00520-8

86. Hoffmann DE, Tarzian AJ. The Girl Who Cried Pain: A Bias against Women in the Treatment of Pain. *Journal of Law, Medicine & Ethics*. 2001;29(1):13-27.
doi:10.1111/j.1748-720X.2001.tb00037.x
87. Goyal MK, Kuppermann N, Cleary SD, Teach SJ, Chamberlain JM. Racial Disparities in Pain Management of Children With Appendicitis in Emergency Departments. *JAMA Pediatr*. 2015;169(11):996-1002. doi:10.1001/jamapediatrics.2015.1915
88. Boag W, Suresh H, Celi LA, Szolovits P, Ghassemi M. Racial Disparities and Mistrust in End-of-Life Care. In: *Proceedings of the 3rd Machine Learning for Healthcare Conference*. PMLR; 2018:587-602. Accessed February 1, 2023.
<https://proceedings.mlr.press/v85/boag18a.html>
89. Vaughn JR. *Understanding Clinical Pain Management and Patient Experiences of Pain from Electronic Health Records*. Massachusetts Institute of Technology; 2021. Accessed February 1, 2023. <https://dspace.mit.edu/handle/1721.1/140086>
90. Haywood C, Lanzkron S, Ratanawongsa N, Bediako SM, Lattimer-Nelson L, Beach MC. Hospital Self-Discharge Among Adults with Sickle Cell Disease: Associations with Trust and Interpersonal Experiences with Care. *J Hosp Med*. 2010;5(5):289-294.
doi:10.1002/jhm.643
91. Evensen CT, Treadwell MJ, Keller S, et al. Quality of care in sickle cell disease. *Medicine (Baltimore)*. 2016;95(35). doi:10.1097/MD.0000000000004528
92. Katz IT, Ryu AE, Onuegbu AG, et al. Impact of HIV-related stigma on treatment adherence: systematic review and meta-synthesis. *Journal of the International AIDS Society*. 2013;16(3S2):18640. doi:10.7448/IAS.16.3.18640
93. Platt OS, Brambilla DJ, Rosse WF, et al. Mortality in sickle cell disease. Life expectancy

- and risk factors for early death. *N Engl J Med*. 1994;330(23):1639-1644.
doi:10.1056/NEJM199406093302303
94. Shankar SM, Arbogast PG, Mitchel E, Cooper WO, Wang WC, Griffin MR. Medical care utilization and mortality in sickle cell disease: A population-based study. *American Journal of Hematology*. 2005;80(4):262-270. doi:https://doi.org/10.1002/ajh.20485
95. Jeremiah J, O'Sullivan P, Stein MD. Who leaves against medical advice? *J Gen Intern Med*. 1995;10(7):403-405. doi:10.1007/BF02599843
96. Paul G, Gautam P, Mahajan RK, Gautam N, Ragavaiah S. Patients Leaving Against Medical Advice-A National Survey. *Indian J Crit Care Med*. 2019;23(3):143-148.
doi:10.5005/jp-journals-10071-23138
97. Ayed IA. What Makes Patients Leave Against Medical Advice? *Journal of Taibah University Medical Sciences*. 2009;4(1):16-22. doi:10.1016/S1658-3612(09)70077-0
98. Brook M, Hilty DM, Liu W, Hu R, Frye MA. Discharge Against Medical Advice From Inpatient Psychiatric Treatment: A Literature Review. *PS*. 2006;57(8):1192-1198.
doi:10.1176/ps.2006.57.8.1192
99. Clark LT, Watkins L, Piña IL, et al. Increasing Diversity in Clinical Trials: Overcoming Critical Barriers. *Current Problems in Cardiology*. 2019;44(5):148-172.
doi:10.1016/j.cpcardiol.2018.11.002
100. McCarthy CR. Historical background of clinical trials involving women and minorities. *Academic Medicine*. 1994;69(9):695-698.
101. Nuriddin A, Mooney G, White AIR. Reckoning with histories of medical racism and violence in the USA. *The Lancet*. 2020;396(10256):949-951.
doi:10.1016/S0140-6736(20)32032-8

102. Guillory JD. The Pro-Slavery Arguments of Dr. Samuel A. Cartwright. *Louisiana History: The Journal of the Louisiana Historical Association*. 1968;9(3):209-227.
103. TIDYMAN P. ART. VI. A Sketch of the most remarkable Diseases of the Negroes of the Southern States, with an account of the method of treating them, accompanied by physiological observations. *The Philadelphia Journal of the Medical and Physical Sciences (1820-1827)*. 1826;3(6):306.
104. Cartwright SA. *Report on the Diseases and Physical Peculiarities of the Negro Race*. New Orleans Medical and Surgical Journal; 1851.
105. Pernick MS. *A Calculus of Suffering: Pain, Professionalism, and Anesthesia in Nineteenth-Century America*. NY: Columbia University Press; 1985. Accessed November 30, 2022. <http://deepblue.lib.umich.edu/handle/2027.42/116015>
106. Trawalter S, Hoffman KM. Got Pain? Racial Bias in Perceptions of Pain. *Social and Personality Psychology Compass*. 2015;9(3):146-157. doi:10.1111/spc3.12161
107. Green AR, Carney DR, Pallin DJ, et al. Implicit Bias among Physicians and its Prediction of Thrombolysis Decisions for Black and White Patients. *J Gen Intern Med*. 2007;22(9):1231-1238. doi:10.1007/s11606-007-0258-5
108. Publisher apologises for “racist” text in medical book. *BBC News*. <https://www.bbc.com/news/blogs-trending-41692593>. Published October 20, 2017. Accessed November 3, 2022.
109. Bassett MT. Beyond Berets: The Black Panthers as Health Activists. *Am J Public Health*. 2016;106(10):1741-1743. doi:10.2105/AJPH.2016.303412
110. Hamberg K. Gender Bias in Medicine. *Womens Health (Lond Engl)*. 2008;4(3):237-243. doi:10.2217/17455057.4.3.237

111. Holdcroft A. Gender bias in research: how does it affect evidence based medicine? *J R Soc Med.* 2007;100(1):2-3. doi:10.1177/014107680710000102
112. Autry AM, Meurer LN, Barnabei VM, et al. A longitudinal women's health curriculum: a multi-method, multiperspective needs assessment. *Am J Obstet Gynecol.* 2002;187(3 Suppl):S12-14. doi:10.1067/mob.2002.127369
113. Fiscella K, Shin P. The Inverse Care Law: Implications for Healthcare of Vulnerable Populations. *The Journal of Ambulatory Care Management.* 2005;28(4):304-312.
114. Allen H, Wright BJ, Harding K, Broffman L. The Role of Stigma in Access to Health Care for the Poor. *Milbank Q.* 2014;92(2):289-318. doi:10.1111/1468-0009.12059
115. Reutter LI, Stewart MJ, Veenstra G, Love R, Raphael D, Makwarimba E. "Who Do They Think We Are, Anyway?": Perceptions of and Responses to Poverty Stigma. *Qual Health Res.* 2009;19(3):297-311. doi:10.1177/1049732308330246
116. Saulnier KM. Telling, Hearing, and Believing: A Critical Analysis of Narrative Bioethics. *Bioethical Inquiry.* 2020;17(2):297-308. doi:10.1007/s11673-020-09973-y
117. Ezenwa MO, Molokie RE, Wilkie DJ, Suarez ML, Yao Y. Perceived Injustice Predicts Stress and Pain in Adults with Sickle Cell Disease. *Pain Management Nursing.* 2015;16(3):294-306. doi:10.1016/j.pmn.2014.08.004
118. NHLBI, NIH. Evidence-Based Management of Sickle Cell Disease: Expert Panel Report, 2014 | NHLBI, NIH. Published September 2014. Accessed February 13, 2021. <https://www.nhlbi.nih.gov/health-topics/evidence-based-management-sickle-cell-disease>
119. Fredericksen R, Edwards T, Crane HM, et al. Patient and provider priorities for self-reported domains of HIV clinical care. *AIDS Care.* 2015;27(10):1255-1264. doi:10.1080/09540121.2015.1050983

120. Yang L, Wong LY, Grivel MM, Hasin DS. Stigma and substance use disorders: an international phenomenon. *Curr Opin Psychiatry*. 2017;30(5):378-388.
doi:10.1097/YCO.0000000000000351
121. Dydyk AM, Jain NK, Gupta M. Opioid Use Disorder. In: *StatPearls*. StatPearls Publishing; 2022. Accessed November 21, 2022. <http://www.ncbi.nlm.nih.gov/books/NBK553166/>
122. Cheetham A, Picco L, Barnett A, Lubman DI, Nielsen S. The Impact of Stigma on People with Opioid Use Disorder, Opioid Treatment, and Policy. *Subst Abuse Rehabil*. 2022;13:1-12. doi:10.2147/SAR.S304566
123. Stone EM, Kennedy-Hendricks A, Barry CL, Bachhuber MA, McGinty EE. The role of stigma in U.S. primary care physicians' treatment of opioid use disorder. *Drug and Alcohol Dependence*. 2021;221:108627. doi:10.1016/j.drugalcdep.2021.108627
124. Werder K, Curtis A, Reynolds S, Satterfield J. Addressing Bias and Stigma in the Language We Use With Persons With Opioid Use Disorder: A Narrative Review. *J Am Psychiatr Nurses Assoc*. 2022;28(1):9-22. doi:10.1177/10783903211050121
125. Richard P, Ferguson C, Lara AS, Leonard J, Younis M. Disparities in physician-patient communication by obesity status. *Inquiry*. 2014;51:0046958014557012.
doi:10.1177/0046958014557012
126. Lindhardt CL, Rubak S, Mogensen O, Lamont RF, Joergensen JS. The experience of pregnant women with a body mass index >30 kg/m² of their encounters with healthcare professionals. *Acta Obstet Gynecol Scand*. 2013;92(9):1101-1107. doi:10.1111/aogs.12186
127. Phelan SM, Burgess DJ, Yeazel MW, Hellerstedt WL, Griffin JM, van Ryn M. Impact of weight bias and stigma on quality of care and outcomes for patients with obesity. *Obes Rev*. 2015;16(4):319-326. doi:10.1111/obr.12266

128. Wear D, Aultman JM, Varley JD, Zarconi J. Making Fun of Patients: Medical Students' Perceptions and Use of Derogatory and Cynical Humor in Clinical Settings. *Academic Medicine*. 2006;81:454-462. doi:10.1097/01.ACM.0000222277.21200.a1
129. Drury CAA, Louis M. Exploring the association between body weight, stigma of obesity, and health care avoidance. *J Am Acad Nurse Pract*. 2002;14(12):554-561. doi:10.1111/j.1745-7599.2002.tb00089.x
130. Aldrich T, Hackley B. The impact of obesity on gynecologic cancer screening: an integrative literature review. *J Midwifery Womens Health*. 2010;55(4):344-356. doi:10.1016/j.jmwh.2009.10.001
131. Phelan SM, Dovidio JF, Puhl RM, et al. Implicit and explicit weight bias in a national sample of 4,732 medical students: The medical student CHANGES study. *Obesity*. 2014;22(4):1201-1208. doi:10.1002/oby.20687
132. Dahl RA, Vakkalanka JP, Harland KK, Radke J. Investigating Healthcare Provider Bias Toward Patients Who Use Drugs Using a Survey-based Implicit Association Test: Pilot Study. *J Addict Med*. 2022;16(5):557-562. doi:10.1097/ADM.0000000000000970
133. Sabin JA, Nosek BrianA, Greenwald AG, Rivara FP. Physicians' Implicit and Explicit Attitudes About Race by MD Race, Ethnicity, and Gender. *J Health Care Poor Underserved*. 2009;20(3):896-913. doi:10.1353/hpu.0.0185
134. Krieger N, Smith K, Naishadham D, Hartman C, Barbeau EM. Experiences of discrimination: Validity and reliability of a self-report measure for population health research on racism and health. *Social Science & Medicine*. 2005;61(7):1576-1596. doi:10.1016/j.socscimed.2005.03.006
135. Van Dyke ME, Kramer MR, Kershaw KN, Vaccarino V, Crawford ND, Lewis TT.

- Inconsistent Reporting of Discrimination Over Time Using the Experiences of
Discrimination Scale: Potential Underestimation of Lifetime Burden. *American Journal of
Epidemiology*. 2022;191(3):370-378. doi:10.1093/aje/kwab151
136. Sheikhalishahi S, Miotto R, Dudley JT, Lavelli A, Rinaldi F, Osmani V. Natural Language
Processing of Clinical Notes on Chronic Diseases: Systematic Review. *JMIR Medical
Informatics*. 2019;7(2):e12239. doi:10.2196/12239
137. National Research Council. *Chapter 4 of Measuring Racial Discrimination (2004)*,
National Research Council, Washington, DC, The National Academies Press. National
Academies Press (US); 2007. Accessed August 25, 2022.
<http://www.ncbi.nlm.nih.gov/books/NBK9804/>
138. Goddu AP, O’Conor KJ, Katie J. O’Conor, et al. Do Words Matter? Stigmatizing Language
and the Transmission of Bias in the Medical Record. *Journal of General Internal Medicine*.
2018;33(5):685-691. doi:10.1007/s11606-017-4289-2
139. Haider AH, Sexton J, Sriram N, et al. Association of unconscious race and social class bias
with vignette-based clinical assessments by medical students. *JAMA*. 2011;306(9):942-951.
doi:10.1001/jama.2011.1248
140. Greenwald AG, McGhee DE, Schwartz JLK. Measuring individual differences in implicit
cognition: The implicit association test. *Journal of Personality and Social Psychology*.
1998;74(6):1464-1480. doi:10.1037/0022-3514.74.6.1464
141. Schimmack U. The Implicit Association Test: A Method in Search of a Construct. *Perspect
Psychol Sci*. 2021;16(2):396-414. doi:10.1177/1745691619863798
142. Martin K, Ricciardelli R, Dror I. How forensic mental health nurses’ perspectives of their
patients can bias healthcare: A qualitative review of nursing documentation. *Journal of*

- Clinical Nursing*. 2020;29(13-14):2482-2494. doi:10.1111/jocn.15264
143. Ashford RD, Brown AM, Curtis B. The Language of Substance Use and Recovery: Novel Use of the Go/No–Go Association Task to Measure Implicit Bias. *Health Communication*. 2019;34(11):1296-1302. doi:10.1080/10410236.2018.1481709
144. Beukeboom CJ. Mechanisms of linguistic bias: How words reflect and maintain stereotypic expectancies. In: *Social Cognition and Communication*. Sydney symposium of social psychology. Psychology Press; 2014:313-330.
145. Beukeboom CJ, Burgers C, Reviews R of CROAHQL. How Stereotypes Are Shared Through Language - A Review and Introduction of the Social Categories and Stereotypes Communication (SCSC) Framework. *Review of Communication Research*. Published online January 1, 2019. Accessed February 26, 2024.
https://www.academia.edu/38218166/Beukeboom_and_Burgers_2019_How_Stereotypes_Are_Shared_Through_Language_A_Review_and_Introduction_of_the_Social_Categories_and_Stereotypes_Communication_SCSC_Framework
146. Zhang H, Lu AX, Abdalla M, McDermott M, Ghassemi M. Hurtful words: quantifying biases in clinical contextual word embeddings. In: *Proceedings of the ACM Conference on Health, Inference, and Learning*. CHIL '20. Association for Computing Machinery; 2020:110-120. doi:10.1145/3368555.3384448
147. Abuse NI on D. Words Matter - Terms to Use and Avoid When Talking About Addiction. National Institute on Drug Abuse. Published November 29, 2021. Accessed January 22, 2022.
<https://www.drugabuse.gov/nidamed-medical-health-professionals/health-professions-education/words-matter-terms-to-use-avoid-when-talking-about-addiction>

148. Abuse NI on D. Words Matter - Terms to Use and Avoid When Talking About Addiction. National Institute on Drug Abuse. Published November 29, 2021. Accessed June 14, 2022. <https://nida.nih.gov/nidamed-medical-health-professionals/health-professions-education/words-matter-terms-to-use-avoid-when-talking-about-addiction>
149. Glassberg J, Tanabe P, Richardson L, DeBaun M. Among emergency physicians, use of the term “Sickler” is associated with negative attitudes toward people with sickle cell disease. *Am J Hematol*. 2013;88(6):532-533. doi:10.1002/ajh.23441
150. 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
151. Sun M, Oliwa T, Peek ME, Tung EL. Negative Patient Descriptors: Documenting Racial Bias In The Electronic Health Record. *Health Affairs*. 2022;41(2):203-211. doi:10.1377/hlthaff.2021.01423
152. Chafe WL, Nichols J. *Evidentiality: The Linguistic Coding of Epistemology*. Ablex; 1986.
153. Benamara F, Taboada M, Mathieu Y. Evaluative Language Beyond Bags of Words: Linguistic Insights and Computational Applications. *Computational Linguistics*. 2017;43(1):201-264. doi:10.1162/COLI_a_00278
154. Garber M. The Scare Quote: 2016 in a Punctuation Mark. The Atlantic. Published December 23, 2016. Accessed September 23, 2022. <https://www.theatlantic.com/entertainment/archive/2016/12/the-scare-quote-2016-in-a-punctuation-mark/511319/>
155. Saner E. Donald Trump and the ‘rise’ of scare ‘quotes.’ *The Guardian*. <https://www.theguardian.com/science/shortcuts/2017/mar/14/donald-trump-and-the-rise-of->

- scare-quotes. Published March 14, 2017. Accessed February 9, 2023.
156. Beach MC, Saha S. Quoting Patients in Clinical Notes: First, Do No Harm. *Ann Intern Med.* 2021;174(10):1454-1455. doi:10.7326/M21-2449
157. [1904.03323] Publicly Available Clinical BERT Embeddings. Accessed September 22, 2021. <https://arxiv.org/abs/1904.03323>
158. Steinkamp JM, Bala W, Sharma A, Kantrowitz JJ. Task definition, annotated dataset, and supervised natural language processing models for symptom extraction from unstructured clinical notes. *Journal of Biomedical Informatics.* 2020;102:103354. doi:10.1016/j.jbi.2019.103354
159. Yang X, Yu Z, Guo Y, Bian J, Wu Y. Clinical Relation Extraction Using Transformer-based Models. Published online August 16, 2021. doi:10.48550/arXiv.2107.08957
160. Vaswani A, Shazeer N, Parmar N, et al. Attention is All you Need. In: *Advances in Neural Information Processing Systems.* Vol 30. Curran Associates, Inc.; 2017. Accessed November 21, 2022. <https://proceedings.neurips.cc/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html>
161. Dash S, Shakyawar SK, Sharma M, Kaushik S. Big data in healthcare: management, analysis and future prospects. *Journal of Big Data.* 2019;6(1):54. doi:10.1186/s40537-019-0217-0
162. Nelson LK. Computational Grounded Theory: A Methodological Framework. *Sociological Methods & Research.* 2020;49(1):3-42. doi:10.1177/0049124117729703

Aim 1: CARE-SD: Classifier-based Analysis for Recognizing and Eliminating Stigmatizing and Doubt Marker Labels in Electronic Health Records: model development and validation

Abstract

Objective: To detect and classify features of stigmatizing and biased language in intensive care EHR using natural language processing techniques. **Materials and Methods:** We first created a lexicon and regular expression lists from literature-driven stem words for linguistic features of stigmatizing patient labels, doubt markers, and scare quotes within electronic health records. The lexicon was further extended using Word2Vec and GPT 3.5, and refined through human evaluation. These lexicons were used to search for matches across 18 million sentences from the de-identified Medical Information Mart for Intensive Care-III (MIMIC-III) dataset. For each linguistic bias feature, 1000 sentence matches were sampled, labeled by expert clinical and public health annotators, and used to train random forest, naïve bayes, logistic regression, and transformer-based classifiers. **Results:** Lexicons development from expanded literature stem-word lists resulted in a doubt marker lexicon containing 58 words, and a stigmatizing labels lexicon containing 127 words. Regular expression matching with stigmatizing label lexicon yielded 8,950 notes, with doubt marker lexicon yielding 3,682 notes, and quote regular expressions yielded 4,806 unique notes. Following annotation and model training to refine whether these notes actually contained the stigmatizing linguistic features, the classifiers for doubt markers and stigmatizing labels had the highest performance, with positive-label recall and precision values ranging from .71 to .86, and accuracies aligning closely with human annotator agreement (.87). Scare quote detection had much lower performance, likely due to a smaller sample of positive instances. **Discussion:** The study showed the feasibility of supervised classifiers in automatically

identifying stigmatizing labels and doubt markers in medical text. Stigmatizing terms like “needy”, “refused”, “uncooperative”, and “noncompliant”, and doubt terms like “claimed” and “insisted”, were more likely to be flagged as stigmatizing or invalidating. Further research should include different datasets, and draw from patient and provider perspectives when deriving meaning from medical chart text. **Conclusions:** Classifiers developed in this study showed high model performance and can be applied to identifying patterns of stigmatizing labels and doubt markers in other healthcare systems. Models and lexicons, titled CARE-SD: Classifier-based Analysis for Recognizing and Eliminating Stigmatizing and Doubt Marker Labels in Electronic Health Records, are available on GitHub and Physionet.

Introduction:

Provider biases and stigmatization of patients drive healthcare inequities

Provider stigmas and biases are widely believed to contribute to discrimination and health inequities among patients.^{1,2} Patients regularly experience stigma and biases by providers as a result of their race, gender, sexual orientation, disease status, drug use, and socioeconomic status, or other labeled characteristics identified by provider teams.³ Our understanding of how stigma and bias impact health outcomes in clinical settings is guided by ecosocial theory, along with an existing conceptual model posited by Hagiwara and colleagues⁴. Our model adaptation has operationalized the construct of *provider communication behaviors* as linguistic features in the EHR, which work to impact patient behaviors and health by affecting patient-provider relationships, satisfaction, trust, and clinical decision-making. This model also posits that provider communication behaviors are part of a reciprocal feedback loop within patient-provider relationships. In this loop, negative implicit biases held by providers are hypothesized to produce

negative provider communication behaviors in patient interactions, which are then met with reduced quality of patient communication. These processes work in tandem to impact patient trust and satisfaction, as well as clinical decision making in the context of the patient's healthcare team. While current interventional approaches to reducing provider stigmatization have typically involved online or workshop-based education on implicit biases, these strategies may have limited effect over time and are often ill-equipped to provide ongoing accountability to ensure long-term behavioral change³. The ability to identify potential instances of patient stigmatization within the electronic health record could help to inform and target future interventions and facilitate real-time audits of healthcare team communication.

This paper aims to apply advanced methods in natural language processing to detect and assess the presence of stigmatizing and biased language in the EHR for patients within the ICU. This study extends qualitative and NLP methods using Grounded Computational Theory to detect provider stigma bias in EHR in an effort to increase the accountability, evaluation, and intervention of negative bias that threatens to deteriorate quality of care for patients among a variety of uniquely stigmatized groups. Building from other research in linguistic stigma and bias in medical charts, this study represents the development of the most comprehensive, automated classification system for doubt markers and stigmatizing labels, applied on the largest corpus of de-identified electronic health records to date.⁵ Doubt marker and stigmatizing label lexicons, as well as classification models, are available for others to utilize this pipeline to identify stigmatizing language in EHR.

Stigma defined

Stigma has been defined by social psychologists Link and Phelan as a social process that is characterized by the interplay of *labeling*, *stereotyping* and *separation*, which leads to status loss and discrimination, and importantly, occurs within a context of power, such as that of the patient-provider relationship.⁶ *Labeling* occurs whenever we identify individuals as belonging to a particular group. *Stereotyping* involves ascribing a proclivity towards a specific behavior or characteristic to members of a labeled group. *Separation* involves distancing from the group, and drawing lines of “us versus them”. Status loss and discrimination involve a negative evaluation of a groups’ attributes and its members relative to another group. As Link and Phelan posit, all of these processes crucially develop in the context of power and power imbalances, which are extremely relevant in the context of provider-patient relationships.^{7,8} Borrowing from Link and Phelan, linguistic bias has been defined as “a systematic asymmetry in word choice as a function of the social category to which the target belongs”.⁹ Although there are a variety of ways in which linguistic bias can manifest,¹⁰ they all facilitate the transmission of “essentialist beliefs about social categories”, resulting in the separation and stigmatization of others.

Linguistic manifestations of stigmatization

Beukeboom and Burgers recently developed a framework to understand how stigmatization can manifest and be formed through language, via the Social Categories and Stereotype Communication Framework.¹¹ This framework posits that stereotypes are communicated through systematic differences in how other groups are 1) labeled , and 2) how their behaviors and characteristics are described. Recent research investigating linguistic bias in the EHR has identified salient features of bias within the language of provider notes, namely

focused on the usage of stigmatizing labels, negative descriptors, as well as the expression of doubt in patient testimony through linguistic doubt markers and “scare quotes”.¹²⁻¹⁴ In applying Beukeboom and Burgers’ framework towards the study of stigma in the electronic health record, we position stigmatizing labels, which also includes negative descriptors as ways in which patients may be differentially labeled, and doubt markers and scare quotes as ways in which the behaviors and characteristics of patients may be differentially described.

Stigmatizing labels

Stigmatizing labels to describe groups are often used to perpetuate stereotypes, and when used by providers, can lead to feelings of stigmatization and reduced trust among their patients. Much of the recent work on identifying and reducing stigmatizing labels has come from providers seeking to improve care for patients with substance use disorders. A recent NIDA study published a list of words to avoid using around patients with substance use disorders, including "addict", "abuser", "user", or "junkie", which have been found to be associated with perceived stigmatization by patients.¹⁵ Similar studies have been applied to other chronic illness populations, identifying terms like “sickler” or “frequent flier” which may be used to further stigmatize patients with chronic illnesses who are often admitted into the hospital.^{13,16} While some providers may argue that these terms may be useful in flagging unwanted patient behaviors or mental states, a recent study has shown that patients exposed to language written about them by providers which included stigmatizing labels resulted in patients feeling unfairly judged, labeled, and disrespected.¹⁷ Given the recent changes in increased patient medical chart access and transparency championed by the 21st Century Cures Act, it is imperative to evaluate the quality of provider medical chart language to reduce patient stigmatization.^{18,19}

Recent research led by Michael Sun and colleagues on over 40,000 clinical notes has found disparities in presence of “Negative Descriptor” words, evaluated by the Health Equity Commission of the Society of General Internal Medicine, which included commonly used terms in the EHR such as “(non-)adherent, aggressive, agitated, angry, challenging, combative, (non-)compliant, confront, (non-)cooperative, defensive, exaggerate, hysterical, (un-)pleasant, refuse, and resist”. This study found that compared to White patients, Black patients had 2.54 times the odds of having at least one negative descriptor written in their history and physical notes²⁰. Research into stereotype expression in language has found that even seemingly innocuous category labels may prompt others to perceive target individual actions and characteristics as “static” aspects of their identity, and exaggerate differences across groups and similarities within them.⁹ These labels can be used to justify clinical decision-making, withholding of resources, or to confer doubt upon patient testimonies.¹⁰

Doubt Markers

Linguistic features such as evidentials, defined as “the linguistic coding of epistemology”,²¹ are words that are frequently used in chart language to question the veracity of patients. Among the many words used as evidentials, words and expressions used to confer doubt or uncertainty such as: *allegedly*, *apparently*, or verbs like *claimed*, are often used to when describing patient testimonies, for example: “patient *claimed* their pain was 10/10”.²² Providers may use words when describing patient testimony in combination with stigmatizing labels or negative descriptors of patients to transmit their stance, or expression of attitudes, feelings, and judgment about patients to other providers which may impact future treatment and care decisions.²³ Inequities have been found among usage of these terms across race and gender, where patients who were women and patients who were Black were found to have significantly

higher frequencies of evidentials in their provider notes than patients who were men or White.²²

The extent to which providers use evidentials is posited to be reflective of the amount of doubt and uncertainty a provider has on patient testimony, and is thus hypothesized to impact trust within the patient-provider relationship, and related outcomes like leaving against medical advice or in pain management strategies.^{22,24}

Scare Quotes

Another linguistic marker of uncertainty that has been previously identified in patient charts are “scare quotes”, which involve the utilization of quotation marks to mock, cast doubt, challenge patient credibility, or insinuate low health literacy when describing the testimony of another individual.²² While use of scare quotes has been documented since the 1950s, some argue that the linguistic phenomenon has been increasing in recent years, both from the rise of “air quoting” gesture in the 80s and 90s, and was commonly employed by former President Donald Trump prior to and throughout his presidency.^{25,26}

While quotations in charts can be useful to describe patient symptoms using their exact language and document patient wishes or concerns, recent linguistic research has identified a troubling prevalence of providers utilizing quotations in ways to mock, manipulate, and regulate the voices of patients. For example, consider the ambiguity added to the sentence: “Patient reports 10/10 pain related to sickle cell crisis.”, when you add “Patient reports ‘10/10’ pain related to ‘sickle cell crisis’”. Because of the quotation marks, both 10/10 and sickle cell crisis could be inferred as being untrue or uncertain. Similar to evidentials and negative patient descriptors, scare quotes have been found to be more prevalent among patients who were Black and among patients who were women.^{12,22}

Natural Language Processing to identify stigma and bias in EHR Data

While linguistic studies have acted to guide researchers into identifying manifestations of stigma and bias in text, due to the nature of qualitative, in-depth assessments, these methods can have significant limitations in being able to be deployed to rapidly identify stigma and bias in medical notes in a way in which we can intervene on it. Recent advances in computational linguistics are allowing researchers to harness human-annotated insights on linguistic bias and stigma and scale up categorization to larger amounts of unlabeled data.

In combination with utilizing state of the art modeling methods, computational linguistics researchers in the social sciences have recently called for the utilization of “Grounded Computational Theory” in order to maintain the role of human expert annotation in the loop to evaluate model output and processes in major NLP tasks, particularly related to those aimed at impacting human well-being and health.²⁷ Within the context of healthcare, human-in-the-loop systems are vital in the development of clinical applications.²⁸ This process of iterative guidance is key to identify and prevent incorrect, biased, and potentially harmful machine learning-driven recommendations before they are deployed in patient care. This expert-driven framework has been advocated for in order to mitigate the potential harmful effects which clinical machine learning applications may inadvertently have on the populations they intend to support.

Methods:

Our methods for identifying sentences within the MIMIC-III EHR dataset containing doubt markers, stigmatizing labels, and scare quotes within provider notes consisted of three steps:

- 1) Lexicon development and sample preparation
- 2) Sentence-level annotation, and

3) Supervised classification using bag-of-words and transformer-based models

MIMIC-III Dataset

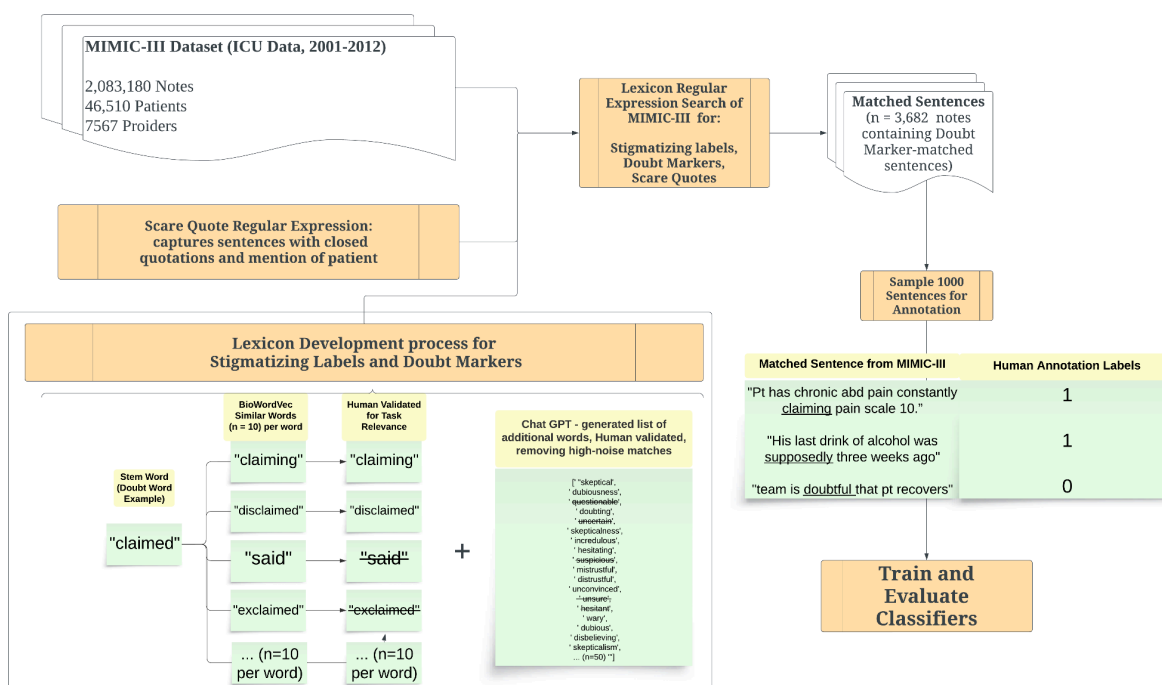
The Medical Information Mart for Intensive Care, or “MIMIC-III”, is a freely-available database of comprehensive, de-identified EHR, free-text notes, and event documentation for over 40,000 patients admitted to the ICU at Beth Israel Deaconess Medical Center in Boston, MA from 2001 to 2012.²⁹ This dataset contains over 1.2 million clinical provider notes, across nearly 50,000 admissions. Because this dataset contains freely-available, EHR from ICU providers from a diverse range of conditions and age ranges, it is a valuable resource for developing bias and stigma detection algorithms in provider language, particularly for patients living with chronic illnesses who may be more likely to be admitted to critical care units.³⁰

Lexicon development and sample preparation

The lexicon development process for doubt markers and stigmatizing patient labels began with a stem-word list describing words previously identified as demarcating doubt or perpetuating stigmatizing patient labels within medical charts. We expanded these word lists to include misspellings or words with high semantic similarity and relevance in the domain by using two subsequent techniques: 1) BioWordVec, a word embeddings model trained on medical text, which generated the top 10 most semantically similar words for each stem word,^{31,32} and 2) GPT 3.5, which suggested an additional 25 words and spelling deviations for each lexicon, following chain-of-thought prompting related to each linguistic bias feature.³³ Following the first round of expansion, we manually validated the list of generated words for task relevance, and assessed human annotation interrater reliability on whether each word was relevant to each specific bias feature. After the second round of GPT 3.5 expansions, we assessed

10-20 sample matches from the highest top-frequency terms' to remove any extremely high-frequency word matches from the lexicon which were not related to transmission of stigmatizing labels or doubt markers and could have significant impact on the annotation sample. This iterative process, reliant on expert-driven inquiry, and complemented by unsupervised, supervised, and transfer learning methods, reflects the strategies championed by computational grounded theory framework, and ensures our results are informed and validated by human domain experts.²⁷ Our analytic pipeline is outlined in Figure 1, with intermediary results described in Appendix 1.

Figure 1: Natural language processing analytic pipeline for lexicon development, regular expression matching, annotation, and classifier model training for stigmatizing linguistic features in MIMIC-III.



Stigmatizing labels

Stigmatizing label lexicon development was guided by literature on stigmatizing language in medical care, specifically from the NIDA “Words Matter” publication, Sun’s “Negative Patient Descriptors: Documenting Racial Bias in the Electronic Health Record”, as well as Zestcott’s “Health Care Providers’ Negative Implicit Attitudes and Stereotypes of American Indians”.^{15,20,34} The initial stem word list consisted of 18 words:

"abuser", "junkie", "alcoholic", "drunk", "drug-seeking", "nonadherent", "agitated", "angry", "combative", "noncompliant", "confront", "noncooperative", "defensive", "hysterical", "unpleasant", "refuse", "frequent-flyer", "reluctant".

Doubt markers

Doubt marker lexicon development was guided by literature on use of “doubt markers” in medical care, specifically led by Beach and colleagues, which identified words such as “claims”, “insists”, and “adamant” or “apparently”, which have been found to be used to discredit or invalidate patient testimony. The 6 words included on the initial stem list were: "adamant", "claimed", "insists", "allegedly", "disbelieves", "dubious".

Scare quotes

Scare quote sample preparation was created by searching the MIMIC-III notes using regular expression `((?=\.\.*)"(?=\b(pt|patient|pateint|he|she|they)\b))` which caught matching closed quotes, and references to patients by “patient” derivations and pronouns, in order to more accurately capture quotes with patient attributions. Finally, several words were added to filter rows, where matches with quoted words were commonly referring to answers for “alert and oriented” examinations -- i.e. “Patient Name”, “Hospital”, “Year”, etc.

Matching with sentences in MIMIC-III, creating coding samples

When completed, the final lexicons or regular expression for each linguistic bias feature was used to filter through patient free-text clinical notes, which had been tokenized at the sentence level to allow for easier readability and classification feasibility. All duplicate sentences were removed from the dataset, and charts labeled as EEG or Radiology were removed in order to restrict to charts more likely to have subjective narrative and patient history text data.

After combining all terms in each linguistic bias lexicon, we used regular expressions to combine them using the pipe “OR” operator, and created a matched data frame of sentences with their matched terms. Each of the three resulting data frames were randomly sampled in non-replacement groups of 100 (for double-coder reliability scores), 400 (coded by AT, a Physician’s Assistant), and 500 (coded by AW, a behavioral data scientist).

Annotation process

Coding ontologies for each of the 3 linguistic bias features were developed originally by AW, then iterated on during the first round of reliability coding. The original ontologies were inspired by research led by Beach, Park, and Goddu on the role of stigmatizing language in patient charts^{13,22,24}. Qualitative annotators met once to discuss each of the three coding ontologies and guiding theories, as well as co-code 5 sentence examples from each linguistic bias dataset. Following the first meeting, each coder completed the same set of 100 sentences for each of the linguistic bias feature datasets. After inter-rater reliability was assessed, the coders met to discuss disagreements and sentences marked as “close calls”, or difficult labeling decisions, and

“exemplary” sentences, which were particularly obvious examples to review. After all disagreements were adjudicated by the coders, they then solo-coded 400 (AT) and 500 (AW) sentences to complete the 1000 samples for each linguistic bias feature.

Sentence Classification

The annotation data was used to train supervised models for the binary classification task of identifying sentences which do, or do not contain each of the linguistic bias features. This supervised learning task was carried out using four models: Naive Bayes, Logistic Regression, Random Forest, as well as the state-of-the-art transformer-based RoBERTA model. Sentences from clinical notes were tokenized into 1-2 word unigrams and bigrams

A grid search approach was used for hyperparameter optimization for these models, using the training data set, which was split at 80/20%. For Naive Bayes, Logistic Regression, and Random Forest, we utilized a stratified k-fold, with 5 splits, in order to create training and test sets which preserved the percentage of samples for each class.

Following model training, each model was evaluated on a held-out 20% of the dataframe, in which we prioritized the performance metrics of 1) positive-class precision, 2) positive-class recall, and 3) macro F_1 -score to select the highest-performing model of each model type. Hyperparameter values for each of the best-performing models for each linguistic feature are available in Appendix 2-4. Positive-class precision, or the proportion of true positives divided by the total number of positive predictions, was prioritized to reduce false positives and develop models highly likely to identify actual stigmatizing language. Recall, or the proportion of all true positives labeled correctly, was prioritized to ensure our classifier can identify as many true cases

of stigmatization as possible. F_1 scores, or the harmonic means of recall and precision, provide a measure of balance between the two performance metrics for each model across both positive and negative class labels. We applied bootstrapping to model evaluation by assessing prediction and ground-truth labels of 1000 samples, generated without replacement. The performance metrics of f_1 , precision, accuracy, recall were aggregated to calculate the confidence intervals of all model metrics.

Finally, we assessed best-performing model text feature importance and feature logistic regression coefficients in order to evaluate the degree to which certain matched terms and phrases contribute to the linguistic bias label predictions of the models. Feature importance was assessed using Gini importance mean impurity reduction method for decision trees in the random forest classifiers, and regression coefficients were calculated from the logistic regression classifiers.³⁵⁻³⁷ Code for all analyses for this study are available on GitHub:

https://github.com/drew-walkerr/Diss_Detecting_Provider_Bias.

Results:

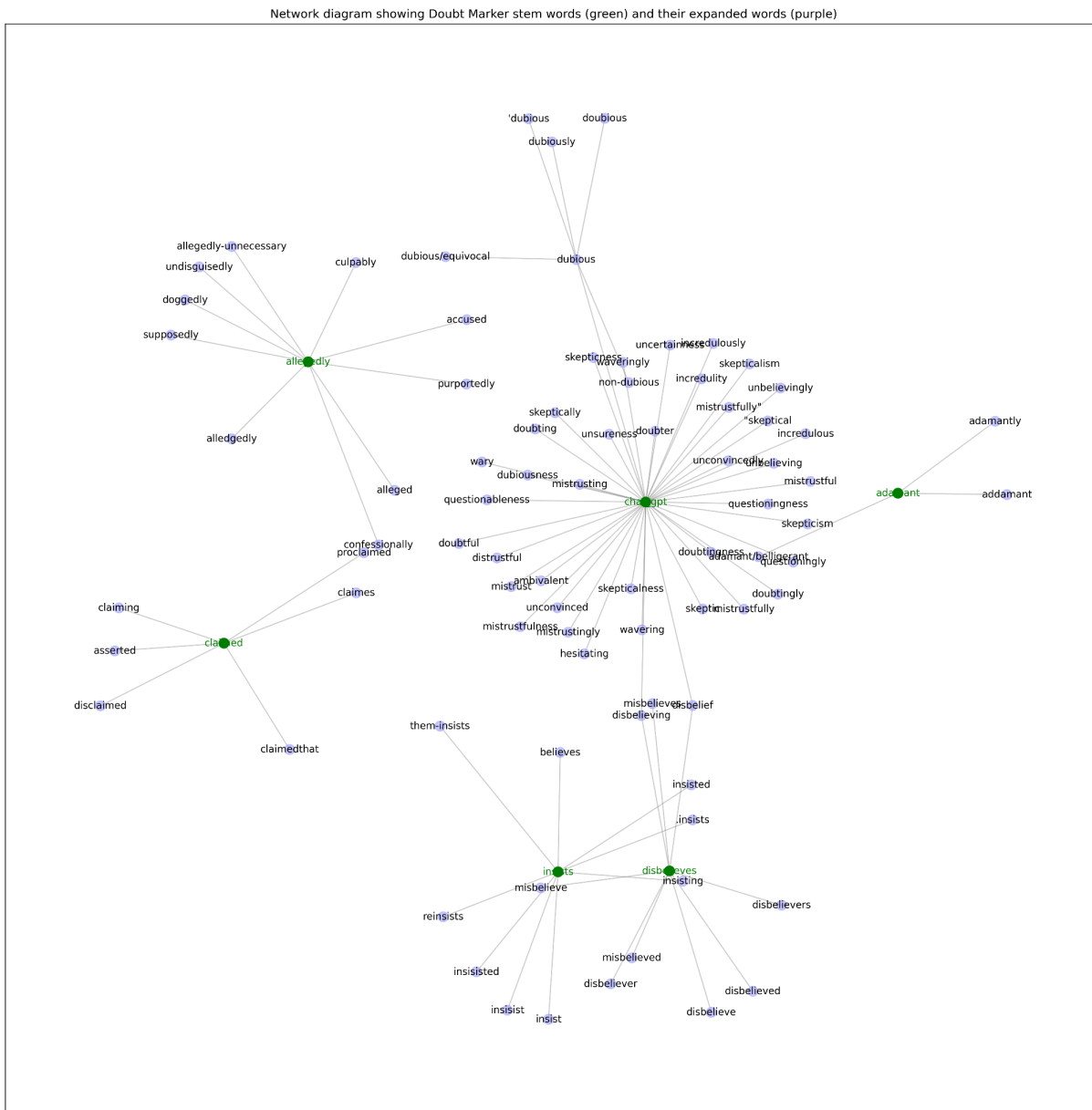
Lexicon Development

For the stigmatizing labels lexicon expansion and annotator pruning, the initial list of 18 was expanded to 180, which was then assessed by annotators DW and SD, removing 83 terms (Annotator agreement = 75%). Final decisions were adjudicated by DW. The final expanded and pruned list of stigmatizing labels used to search the MIMIC-III dataset totaled 127 words, and is provided in Appendix 1. Following assessments of most frequent term matches, we removed the following terms due to high proportion of noise referring to illness characteristics or clinical situations, rather than patients or patient testimonies: 'difficult', 'suspicious', 'aggressive', 'unstable', 'dramatic', 'unreliable', 'entitled', 'invalid', 'violent', 'dangerous'.

For the doubt markers lexicon expansion and annotator pruning, the initial list of 6 terms was expanded to 60, which was then pruned by annotators to remove 2 terms (Annotator agreement = 80%). The final expanded list of doubt markers used to search the MIMIC-III dataset totaled 58 words, and is provided in Appendix 1. Following regular expression searching and assessments of most frequent term matches, we removed the following terms due to high proportion of noise referring to uncertainty in illness or clinical presentations, rather than patient testimonies:

'suspicion', 'suspicious', 'questionable', 'questioning', 'uncertain', 'hesitancy', 'hesitant', 'unsure'.

Figure 2: Network Diagram showing doubt marker stem words (green, also includes GPT 3.5) to show expansions of initial stem words to expanded lexicon words (purple)



Regular Expression Search results

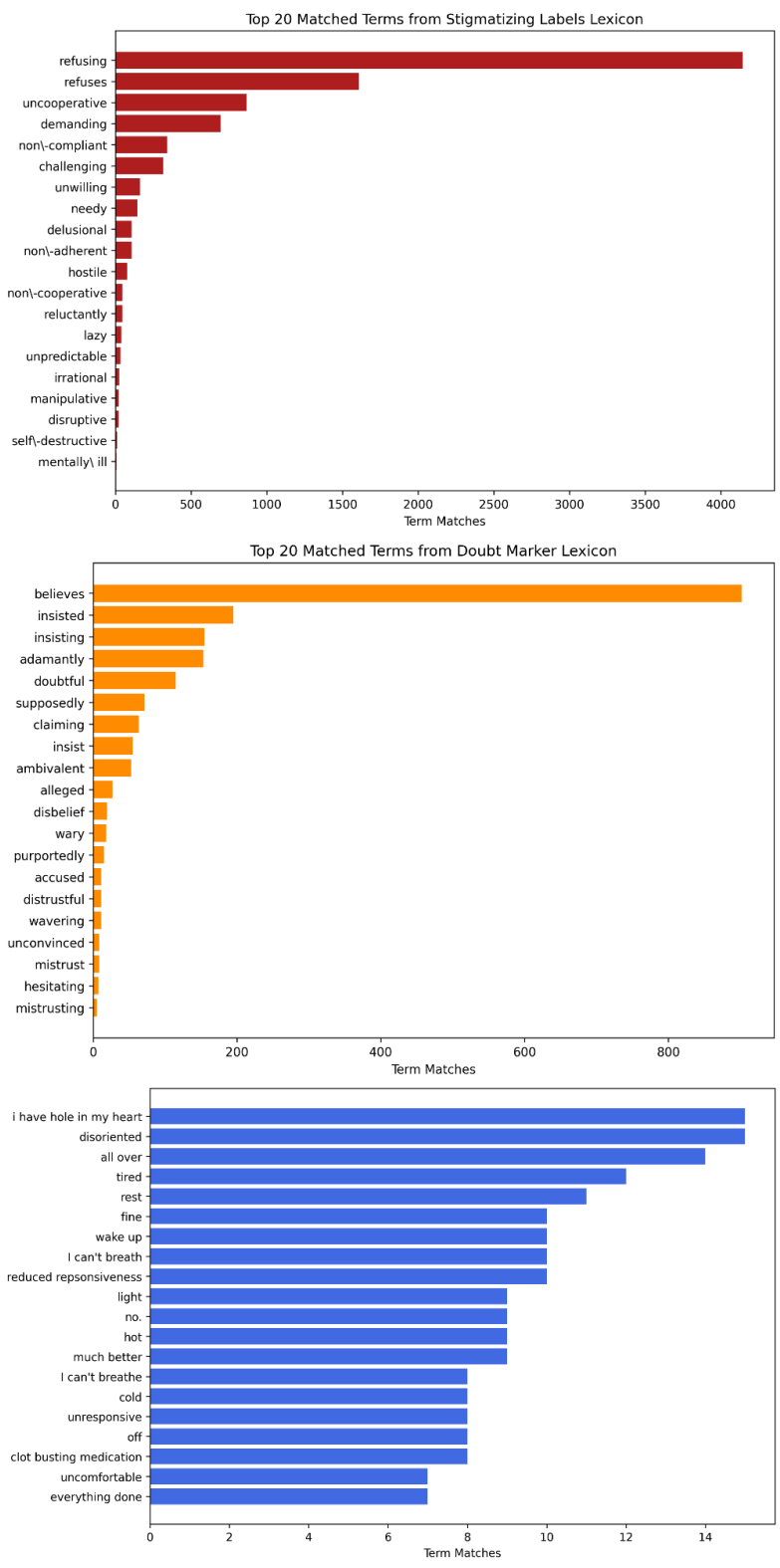
Results describing the text data of the preprocessed MIMIC-III full sample, as well as of search results for each of the stigmatizing labels, doubt markers, and scare quotes regular expression (regex) matched data frames are summarized in Table 1.

Table 1: Summary statistics of MIMIC-III Dataset, compared with linguistic bias corpa

	MIMIC-III Sample	Stigmatizing Label Corpus	Doubt Marker Corpus	Scare Quotes Corpus
Number of unique notes	814,548 notes	8,950 notes	3,682 notes	4,806 notes
Avg note length	654 words	623 words	937 words	763 words
Number of total sentences	18,288,213 sentences	10,278 sentences	3,856 sentences	5,156 sentences
Average sentence length	12 words	48 words	35 words	55 words
Number of patients	11,633 patients	3,483 patients	2,368 patients	2,830 patients
Number of providers	1,879 providers	1,056 providers	800 providers	677 providers

The most frequent matching terms from our lexicon, along with the most commonly occurring trigrams within quoted text, are provided for each of the 3 bias features in Figure 2. For stigmatizing labels, versions of ‘refusing’ and ‘refuses’ were by far the most frequently matched terms. In the doubt marker label lexicon, ‘believes’ was the most frequently matched term, followed by insisted and insisting. Scare quote quoted text frequent words, bigrams, and trigrams were less led by any particular phrases or words, but were mostly used describing patient chief complaints, descriptions of symptoms or condition.

Figure 2: Top 20 Matched Terms Stigmatizing Labels, Doubt Markers, and Scare Quotes



Annotation

Annotation coding ontologies, detailing the labeling instructions for each of the three linguistic bias features, were informed largely through the literature-based stem word operationalizations of each set of linguistic bias features. Appendices 1-3 detail the coding ontologies for each corpus. Table 2 provides the interrater agreement and kappa score for the first 100 samples of each linguistic feature, and positive class frequency in final 1000 sentence sample.

Table 2: Annotation sample reliability, linguistic bias features positive class frequencies, and notable examples

Bias Feature	Agreement	Interrater reliability (Kappa)	Frequency in final sample (N=1000)
Stigmatizing Labels	87%	.74	43.9%
Doubt Markers	87%	.73	31.0%
Scare Quotes	87%	.73	20.7%

Table 3 provides notable positive class examples for each of the 3 linguistic features.

Table 3: Notable annotation examples for stigmatizing labels, doubt markers, and scare quotes

Bias Feature	Notable Sentence Examples (Flagged as containing biased feature, matching or quoted text <u>underlined</u>)
Stigmatizing Labels	<p>Pt very <u>uncooperative</u>, will barely allow any nursing care.</p> <p>Neuro: Is very <u>needy</u>, needs to be encouraged to do more for herself.</p> <p><u>Refuses</u> blood draws, calling out, asking for dilaudid despite med not being due, requiring much emotional support.</p>
Doubt Markers	<p>“Pain control (acute pain, chronic pain) Assessment: Pt has chronic abd pain constantly <u>claiming</u> pain scale 10.”</p> <p>“His last drink of alcohol was <u>supposedly</u> three weeks ago.”</p> <p><u>Insisting</u> on making phone calls becomes very irritated if don't jump to his requests.</p>
Scare Quotes	<p>Easily frustrated, especially when asked questions to assess orientation...pt states "<u>you have already asked me this 100 times</u>".</p> <p>NO nausea/vomiting although pt states he does not want to eat because he "<u>fears</u>" being nauseated.</p> <p>At 10:06, pt put on call light to request "<u>pain pill</u>" and then put on the call light 5 more times over the next 8 minutes to make same demand, at one point saying to the nurse responding to the call light, "What the hell is going on."</p>

Linguistic Bias Classifier Model Evaluation Results

Table 3 displays the results of the best performing models across types and linguistic bias features. A complete list of the best performing model hyperparameters can be found in Appendix 5. We were able to achieve the highest performance on doubt markers and stigmatizing

labels, with scare quote models underperforming other linguistic bias models across nearly every evaluation metric.

The best performing model type for doubt markers and scare quotes was RoBERTA, with Logistic Regression achieving the best performance in classifying stigmatizing label sentences. Run duration was much higher for RoBERTA models, as compared with Random Forest, Logistic Regression, and Naive Bayes classifiers.

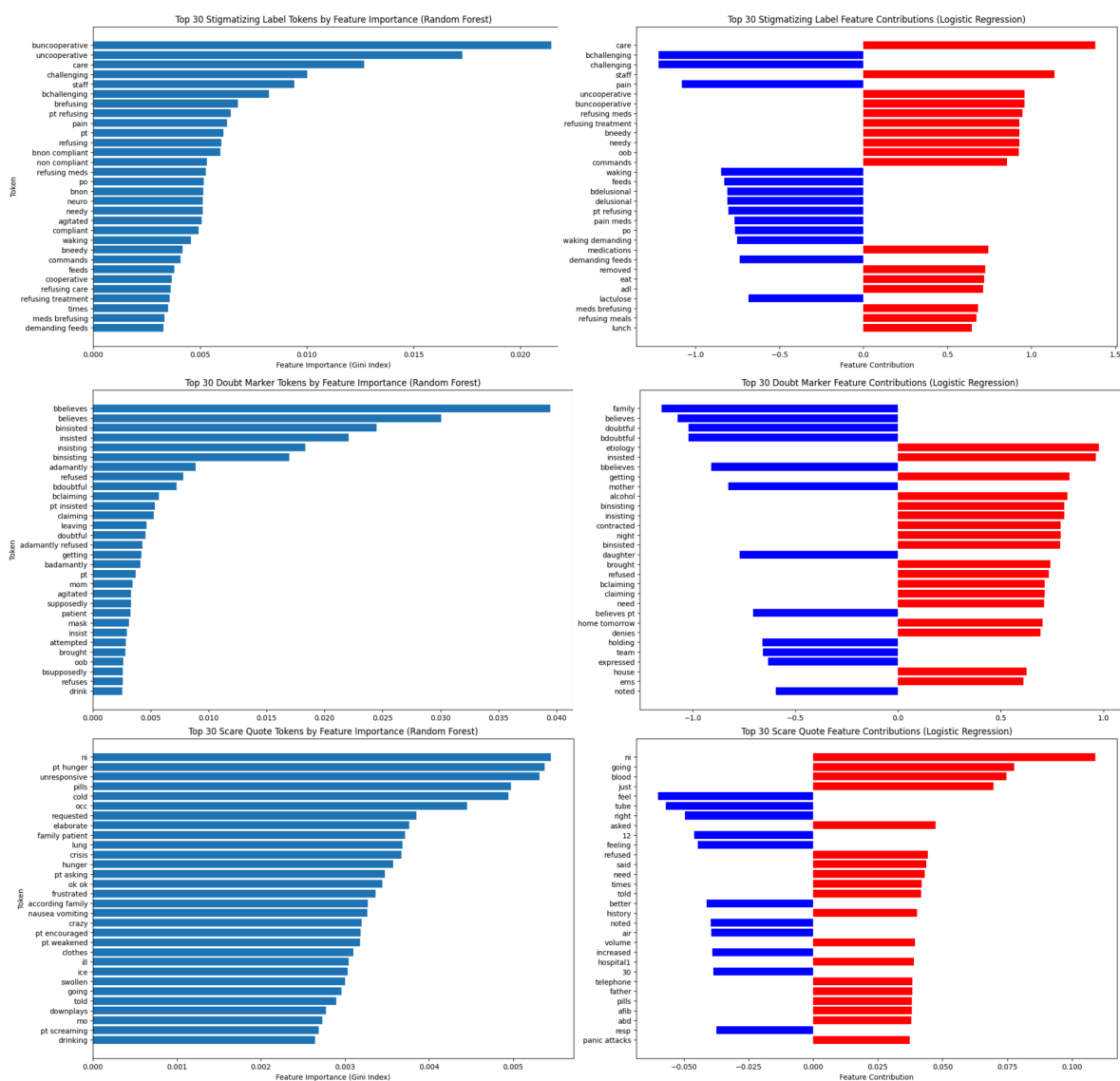
Table 4: Linguistic Bias Classifier Model Performance (best model of each feature in bold)

Bias Feature	Model	Accuracy	Precision (Positive)	Recall (Positive)	F1 (Positive)	Macro Precision	Macro Recall	Macro F1
Stigmatizing Labels	RoBERTA	.69 (.63, .76)	.63 (.54, .72)	.75 (.67, .84)	.69 (.62, .76)	.70 (.64, .76)	.70 (.64, .76)	.69 (.63, .75)
	Random Forest	.79 (.73, .84)	.77 (.68, .86)	.73 (.63, .82)	.75 (.67, .82)	.77 (.68, .86)	.73 (.63, .82)	.75 (.67, .82)
	Logistic Regression	.81 (.75, .86)	.75 (.66, .85)	.84 (.75, .91)	.79 (.72, .86)	.75 (.66, .84)	.84 (.75, .91)	.79 (.72, .86)
	Naive Bayes	.71 (.64, .77)	.62 (.53, .70)	.86 (.79, .93)	.72 (.64, .78)	.62 (.53, .70)	.86 (.79, .93)	.72 (.64, .78)
Doubt Markers	RoBERTA	.86 (.81, .91)	.86 (.75, .96)	.71 (.59, .81)	.77 (.69, .85)	.86 (.80, .91)	.82 (.76, .88)	.84 (.78, .88)
	Random Forest	.85 (.80, .89)	.76 (.64, .86)	.76 (.65, .85)	.76 (.67, .84)	.76 (.64, .86)	.76 (.65, .85)	.76 (.67, .84)
	Logistic Regression	.85 (.80, .90)	.71 (.61, .81)	.89 (.80, .96)	.78 (.69, .85)	.70 (.60, .80)	.88 (.80, .96)	.78 (.69, .85)
	Naive Bayes	.85 (.80, .89)	.70 (.60, .80)	.89 (.80, .96)	.78 (.69, .85)	.70 (.60, .80)	.89 (.80, .96)	.78 (.69, .85)
Scare Quotes	RoBERTA	.75 (.69, .81)	.40 (.24, .58)	.30 (.17, .45)	.35 (.20, .48)	.61 (.52, .70)	.59 (.52, .67)	.62 (.52, .70)
	Random Forest	.79 (.74, .85)	0.00 (0,0)	0.00 (0,0)	0.00 (0,0)	0.00 (0,0)	0.00 (0,0)	0.00 (0,0)
	Logistic Regression	.77 (.71, .82)	.30 (.07, .56)	.10 (.02, .20)	.14 (.04, .28)	.30 (.07, .56)	.10 (.02, .20)	.14 (.04, .28)
	Naive Bayes	.78 (.72, .83)	.43 (.22, .63)	.24 (.12, .38)	.31 (.16, .45)	.43 (.22, .63)	.24 (.12, .38)	.31 (.16, .45)

After model evaluation, we also ran feature importance and contribution plots using the models of the best performing random forest and logistic regression models. Figure 3 highlights terms or phrases that are particularly informative to random forest models during categorization (left-hand

side), and the right-hand side displays the terms with the highest regression coefficients (negative and positive), which are more likely to be labeled as negative (blue, unbiased), or positive (red, biased/stigmatizing).

Figure 3: Top 30 Stigmatizing Label Tokens by Importance and Feature Contributions



Discussion:***Integrative summary of findings***

For the task of lexicon development, our expanded word lists from Word2Vec and GPT 3.5 successfully produced additional terms that frequently matched with sentences which were eventually labeled as stigmatizing/biased by coders. These NLP technologies can help to bridge the gap of concepts/words identified by the literature surrounding a particular phenomenon with the reality of changing language in different EHR corpora from different places and time periods. It is possible to further expand these lists to many different aspects of stigma/bias, or fine-tune them to fit particular marginalized populations for various research questions. A key element of the expansion process, however, was the steps that allowed humans to review intermediate results for relevance/noise, and iteratively guide model development. Lexicon selection requires careful review of matching results and text to identify and remove noisy terms, which are not related to patient stigmatization, given the scale of medical data.

These lexicons are valuable tools for researchers seeking to isolate EHR sentences with a higher frequency of stigmatizing or doubt marking signals than starting with the entire EHR corpora unfiltered. They build on knowledge of relevant qualitative studies on stigma and bias in healthcare, and provide greater breadth and flexibility through word embeddings expansion steps, which capture words which may be close in spelling variation or semantic similarity. Our suite of tools, comprising lexicons and refined supervised classification models, titled “CARE-SD: Classifier-based Analysis for Recognizing and Eliminating Stigmatizing and Doubt Marker Labels in Electronic Health Records” are available on [GitHub and Physionet](#) for other researchers in this space seeking to reproduce or adapt our study process. Our goal in sharing these tools is to accelerate research efforts in the study of linguistic stigma and bias in healthcare.

With additional validation, these models can be used to audit and evaluate healthcare systems for units, providers, or patients whom experience higher rates of stigmatizing labels and doubt markers, and allow real-time feedback for anti-stigmatization intervention efforts in a way that has not been possible with traditional implicit bias training approaches.³

In the task of classifying sentences as containing stigmatizing labels and doubt markers, models showed consistent predictions of positive classes that nearly matched levels of human agreement. We believe that these tools, trained on ICU data among a large cohort of patients, will be applicable to other EHR settings or datasets since the linguistic markers are likely to be semantically similar.

Several patterns and terms across stigmatizing labels and doubt marker classification arose that may help inform future clinical practice. Among stigmatizing labels, the use of “needy” was one that was especially predictive of positive stigmatizing labels. While we did not find any major research publications on the use of this word in medical contexts, it’s important to consider how “needy” describes an individual as holding an inherent trait, which is highly problematic for anyone seeking care in the ICU. Other providers may wrongfully assume the patient is “needy” in other contexts, including pain management, daily activity assistance, or any other genuine health complaint. “Noncompliant”, a word identified in the expanded stigmatizing labels lexicon, was labeled as stigmatizing chiefly when it was used to define the patient directly as such and not when relating to a specific behavior or more context is provided (i.e. due to lack of funds). Labeling patients as “noncompliant” has been hotly debated, though recent EHR-related NLP work is increasingly operationalizing its use as one that has negative connotations, and points blame at the patient, rather than structural factors which may have stronger impacts and constraints on patient health.^{5,20}

Finally, we encountered many examples of refusing/refused/refuses among stigmatizing label charts. We aimed to negatively label (unbiased) when providers were specific about the behavior and context in which the patient was refusing a specific treatment, i.e. “patient refusing anxiety meds, feeling better”, or when the context was surrounding end-of-life or do not resuscitate (patient refusing further life-saving measures), and aimed to positively label (biased) when the language around what patients were refusing was vague, or painted the patient as refusing any care in a way that could portray the patient as inherently stubborn. We also positively labeled sentences which described patients “refusing to try” or “refusing to perform [insert activity of daily living]”. Our coding ontology posits that depicting a patient as refusing to try/assert effort can be interpreted as labeling a patient as lazy or unwillful, when they may be unable to complete tasks due to discomfort or decreased health status.

Information gathered from doubt markers feature importance and contribution plots, as well as from annotator notes and feedback also show patterns that reflect doubt marker usage across a variety of contexts. Use of “insisted” and “insisting” were consistently labeled as positive for doubt markers. While these words may be helpful in expressing the strong conviction behind patient needs, it is frequently also laden with negative, stubborn, or difficult connotations. The word “claimed”, particularly when describing pain or severity of illness symptoms or pain, was highly likely to be labeled as a doubt marker. Alcohol also had high feature contribution towards positive doubt marker labels. Both annotators described that doubt marker words were frequently employed to report patient alcohol and drug use history, as exemplified in Table 2.

Due to the relatively low classification performance for scare quotes, insights from feature importance and contribution plots may not be consistent or reliable. However, one can

still see stigmatizing labels like “refused” coming up as contributing towards a positive class label. Additionally, “need” is listed in this category as well. These being said, the feature contribution beta is low (less than .1, compared with the other plots in the range of .5 to 1.5). Ultimately, while scare quotes may be difficult to classify, it is important for providers to reflect on when and why they quote patients. Quoting patients should be done to explain their condition and experience in their own words, when it is important for their care. Careful consideration should be given when quoting patients in order to avoid the possibility of other providers interpreting their quotation as invalidating of testimony.

Limitations

With results in mind, it is important to consider several limitations of the current study. Firstly, the dataset consists only of notes from ICU admissions. Patients who actually are admitted reflect a subset of total patients. Many stigmatized patients, particularly those with mental health or substance use problems, may not be admitted due to a variety of structural barriers, interactions with healthcare providers, or concomitant illness.³⁸⁻⁴⁰ Low performance on scare quotes may be indicative of the need for additional data (particularly positive classes) to train future models. Across the classifier models, there were many instances of models assigning all negative labels to sentences. This may be due to the complex ways in which scare quotes can manifest, which may arise in the form of juxtaposed negations, i.e.” “patient complained of a ‘fever’, however was testing 98.6”. Other previously operationalized manifestations of scare quotes include quoting inappropriate or excessively long strings of patient testimony, such as “Easily frustrated, especially when asked questions to assess orientation...pt states "you have already asked me this 100 times". Due to the many possible manifestations, it is important to consider different approaches and linguistic structural features within scare quotes.

Additionally, it is important to consider the positionality of labelers in this study. Annotations which trained models were completed by two researchers, who, although trained in reading medical texts and interpreting linguistic bias, have limited scope to the experiences of stigmatization and bias which many patients endure. It is critical to bring in patients and providers across many disciplines for future work towards identifying stigmatizing and biased language in patient notes.

Finally, it is important to convey that despite their performance in classifiers, the operationalizations of linguistic bias features cannot fully capture the ever-changing construct of provider stigma and bias. As language and medicine have changed from 2001-2012 in MIMIC-III notes, there may be words that were acceptable at one time and are only realized to be derogatory over time. The process of identifying stigma and bias must be a continuous effort to keep pace with how stigma and biases evolve over time to evade scrutiny or draw in-group members closer together.^{41,42} While helping providers understand harmful words is an important start, given the changing dynamics of language, bias and stigma, it's critical to convey the deeper patterns which may drive stigmatization, borrowing from Link and Phelan's definition. We need to ensure our language avoids painting patient behaviors as static and global character traits, and focuses on neutral descriptions of behavior that seek to inform and support future medical care.

Conclusions

While additional validation and collaborations are needed to improve models from this paper, the models hold high utility for identifying patterns of stigmatizing labels and doubt markers in healthcare systems, particularly for targeting and designing interventions. Large language models like GPT 3.5 were demonstrated to be helpful in identifying relevant and frequently-occurring words in medical text. These, alongside state-of-the-art transformer

classifiers could be used to identify stigmatizing language and help providers neutralize or re-write the text using more patient-centered language.

References

1. Hatzenbuehler ML, Phelan JC, Link BG. Stigma as a Fundamental Cause of Population Health Inequalities. *Am J Public Health*. 2013;103(5):813-821.
doi:10.2105/AJPH.2012.301069
2. Maina IW, Belton TD, Ginzberg S, Singh A, Johnson TJ. A decade of studying implicit racial/ethnic bias in healthcare providers using the implicit association test. *Soc Sci Med*. 2018;199:219-229. doi:10.1016/j.socscimed.2017.05.009
3. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. *BMC Med Ethics*. 2017;18(1):19. doi:10.1186/s12910-017-0179-8
4. Hagiwara N, Lafata JE, Mezuk B, Vrana SR, Fetters MD. Detecting implicit racial bias in provider communication behaviors to reduce disparities in healthcare: Challenges, solutions, and future directions for provider communication training. *Patient Educ Couns*. 2019;102(9):1738-1743. doi:10.1016/j.pec.2019.04.023
5. 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
6. Link BG, Phelan JC. Conceptualizing Stigma. *Annu Rev Sociol*. 2001;27(1):363-385.
doi:10.1146/annurev.soc.27.1.363
7. Beisecker AE. Patient Power in Doctor-Patient Communication: What Do We Know? *Health Commun*. 1990;2(2):105-122. doi:10.1207/s15327027hc0202_4
8. Goodyear-Smith F, Buetow S. Power issues in the doctor-patient relationship. *Health Care*

- Anal HCA J Health Philos Policy*. 2001;9(4):449-462. doi:10.1023/A:1013812802937
9. Beukeboom CJ, Burgers C. Linguistic Bias. In: *Oxford Research Encyclopedia of Communication*. Oxford University Press; 2017.
doi:10.1093/acrefore/9780190228613.013.439
 10. Beukeboom CJ. Mechanisms of linguistic bias: How words reflect and maintain stereotypic expectancies. In: *Social Cognition and Communication*. Sydney symposium of social psychology. Psychology Press; 2014:313-330.
 11. Beukeboom CJ, Burgers C, Reviews R of CROAHQL. How Stereotypes Are Shared Through Language - A Review and Introduction of the Social Categories and Stereotypes Communication (SCSC) Framework. *Rev Commun Res*. Published online January 1, 2019. Accessed February 26, 2024.
https://www.academia.edu/38218166/Beukeboom_and_Burgers_2019_How_Stereotypes_Are_Shared_Through_Language_A_Review_and_Introduction_of_the_Social_Categories_and_Stereotypes_Communication_SCSC_Framework
 12. Beach MC, Saha S. Quoting Patients in Clinical Notes: First, Do No Harm. *Ann Intern Med*. 2021;174(10):1454-1455. doi:10.7326/M21-2449
 13. 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
 14. Zhang H, Lu AX, Abdalla M, McDermott M, Ghassemi M. Hurtful words: quantifying biases in clinical contextual word embeddings. In: *Proceedings of the ACM Conference on Health, Inference, and Learning*. CHIL '20. Association for Computing Machinery; 2020:110-120. doi:10.1145/3368555.3384448

15. Abuse NI on D. Words Matter - Terms to Use and Avoid When Talking About Addiction. National Institute on Drug Abuse. Published November 29, 2021. Accessed January 22, 2022.
<https://www.drugabuse.gov/nidamed-medical-health-professionals/health-professions-education/words-matter-terms-to-use-avoid-when-talking-about-addiction>
16. Glassberg J, Tanabe P, Richardson L, DeBaun M. Among emergency physicians, use of the term “Sickler” is associated with negative attitudes toward people with sickle cell disease. *Am J Hematol*. 2013;88(6):532-533. doi:10.1002/ajh.23441
17. 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
18. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program. Federal Register. Published May 1, 2020. Accessed June 14, 2022.
<https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification>
19. Information Blocking | HealthIT.gov. Accessed December 26, 2023.
<https://www.healthit.gov/topic/information-blocking>
20. Sun M, Oliwa T, Peek ME, Tung EL. Negative Patient Descriptors: Documenting Racial Bias In The Electronic Health Record. *Health Aff (Millwood)*. 2022;41(2):203-211.
doi:10.1377/hlthaff.2021.01423
21. Chafe WL, Nichols J. *Evidentiality: The Linguistic Coding of Epistemology*. Ablex; 1986.
22. 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
23. Benamara F, Taboada M, Mathieu Y. Evaluative Language Beyond Bags of Words: Linguistic Insights and Computational Applications. *Comput Linguist.* 2017;43(1):201-264. doi:10.1162/COLI_a_00278
 24. Park J, Saha S, Chee B, Taylor J, Beach MC. Physician Use of Stigmatizing Language in Patient Medical Records. *JAMA Netw Open.* 2021;4(7):e2117052-e2117052. doi:10.1001/jamanetworkopen.2021.17052
 25. Garber M. The Scare Quote: 2016 in a Punctuation Mark. *The Atlantic.* Published December 23, 2016. Accessed September 23, 2022. <https://www.theatlantic.com/entertainment/archive/2016/12/the-scare-quote-2016-in-a-punctuation-mark/511319/>
 26. Saner E. Donald Trump and the ‘rise’ of scare ‘quotes.’ *The Guardian.* <https://www.theguardian.com/science/shortcuts/2017/mar/14/donald-trump-and-the-rise-of-scare-quotes>. Published March 14, 2017. Accessed February 9, 2023.
 27. Nelson LK. Computational Grounded Theory: A Methodological Framework. *Sociol Methods Res.* 2020;49(1):3-42. doi:10.1177/0049124117729703
 28. Bakken S. AI in health: keeping the human in the loop. *J Am Med Inform Assoc.* 2023;30(7):1225-1226. doi:10.1093/jamia/ocad091
 29. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data.* 2016;3(1):160035. doi:10.1038/sdata.2016.35
 30. Carson SS, Bach PB. The epidemiology and costs of chronic critical illness. *Crit Care Clin.* 2002;18(3):461-476. doi:10.1016/S0749-0704(02)00015-5
 31. Zhang Y, Chen Q, Yang Z, Lin H, Lu Z. BioWordVec, improving biomedical word

- embeddings with subword information and MeSH. *Sci Data*. 2019;6(1):52.
doi:10.1038/s41597-019-0055-0
32. BioWordVec & BioSentVec: pre-trained embeddings for biomedical words and sentences.
Published online October 27, 2022. Accessed November 1, 2022.
<https://github.com/ncbi-nlp/BioSentVec>
33. OpenAI Platform. Accessed December 27, 2023. <https://platform.openai.com>
34. Zestcott CA, Spece L, McDermott D, Stone J. Health Care Providers' Negative Implicit Attitudes and Stereotypes of American Indians. *J Racial Ethn Health Disparities*. 2021;8(1):230-236. doi:10.1007/s40615-020-00776-w
35. Nembrini S, König IR, Wright MN. The revival of the Gini importance? *Bioinformatics*. 2018;34(21):3711-3718. doi:10.1093/bioinformatics/bty373
36. sklearn.linear_model.LinearRegression. scikit-learn. Accessed January 9, 2024.
https://scikit-learn/stable/modules/generated/sklearn.linear_model.LinearRegression.html
37. Feature importances with a forest of trees. scikit-learn. Accessed January 9, 2024.
https://scikit-learn/stable/auto_examples/ensemble/plot_forest_importances.html
38. Chen LY, Crum RM, Martins SS, Kaufmann CN, Strain EC, Mojtabai R. Service Use and Barriers to Mental Health Care Among Adults With Major Depression and Comorbid Substance Dependence. *Psychiatr Serv*. 2013;64(9):863-870. doi:10.1176/appi.ps.201200289
39. Ross LE, Vigod S, Wishart J, et al. Barriers and facilitators to primary care for people with mental health and/or substance use issues: a qualitative study. *BMC Fam Pract*. 2015;16(1):135. doi:10.1186/s12875-015-0353-3
40. Bremer W, Plaisance K, Walker D, et al. Barriers to opioid use disorder treatment: A comparison of self-reported information from social media with barriers found in literature.

Front Public Health. 2023;11:1141093. doi:10.3389/fpubh.2023.1141093

41. Smith RA. Language of the Lost: An Explication of Stigma Communication. *Commun*

Theory. 2007;17(4):462-485. doi:10.1111/j.1468-2885.2007.00307.x

42. Link BG, Phelan JC. Stigma and its public health implications. *The Lancet.*

2006;367(9509):528-529. doi:10.1016/S0140-6736(06)68184-1

Appendix 1: Lexicons for Doubt Markers and Stigmatizing Labels

Lexicon	Stem Word List	Expanded Words (Pruned to)	GPT-3.5 added words	High-noise terms removed	Final Lexicon	Final Lexicon Length
Doubt Markers	"adamant", "claimed", "insists", "allegedly", "disbelieves", "dubious"	60 total, reduced to 42. Agreement = 80%	['skeptical', 'dubiousness', 'questionable', 'doubting', 'uncertain', 'skepticalness', 'incredulous', 'hesitating', 'suspicious', 'mistrustful', 'distrustful', 'unconvinced', 'unsure', 'hesitant', 'wary', 'dubious', 'disbelieving', 'skepticalism', 'hesitancy', 'skepticism', 'mistrust', 'uncertainty', 'disbelief', 'suspicion', 'mistrustfulness', 'incredulity', 'incredulously', 'wavering', 'ambivalent', 'waveringly', 'questionableness', 'mistrustingly', 'doubter', 'questioning', 'doubtingly', 'mistrusting', 'doubtful', 'skeptic', 'unconvincedly', 'mistrustingly', 'mistrustfully', 'doubtingness', 'skepticism',	'suspicion', 'suspicious', 'questionable', 'questioning', 'uncertain', 'hesitancy', 'hesitant', 'unsure'	["doubtful", "dubious", 'insists', 'accused', 'adamant', 'adamant/belligerent', 'adamantly', 'addamant', 'allegedly', 'alleged', 'allegedly', 'allegedly-unnecessary', 'asserted', 'believes', 'claimed', 'claimedthat', 'clames', 'claiming', 'confessionally', 'culpably', 'disbelief', 'disbelieve', 'disbelieved', 'disbeliever', 'disbelievers', 'disbelieves', 'disbelieving', 'disclaimed', 'doggedly', 'doubious', 'doubtful', 'dubious', 'dubious/equivocal', 'dubiously', 'insist', 'insisted', 'insist', 'insisted', 'insisting', 'insists', 'misbelieve', 'misbelieved', 'misbelieves', 'mistrustful', 'mistrusting', 'non-dubious', 'proclaimed', 'purportedly', 'reinsists', 'skeptical', 'speculative', 'supposedly', 'them-insists', 'unconvinced', 'undisguisedly', 'unreliable', 'unsure', 'wavering']	58

			'questioningness', 'unbelieving', 'unsureness', 'skepticalness', 'questioningness', 'doubtingly', 'unbelievingly', 'skeptically', 'mistrustingly', 'mistrustfully', 'skeptically', 'questioningly', 'doubtingly', 'skeptically', 'mistrustingly', 'mistrustfully']			
Stigmatizing Labels	"abuser", "junkie", "alcoholic", "drunk", "drug-seeking", "nonadherent", "agitated", "angry", "combative", "noncompliant", "confront", "noncooperative", "defensive", "hysterical", "unpleasant", "refuse", "frequent-flyer", "reluctant"	180, reduced to 83. Annotator agreement = .75.	['hysterical', 'aggressive', 'drug addict', 'non-compliant', 'lazy', 'attention-seeking', 'manipulative', 'hypochondriac', 'difficult', 'mentally unstable', 'troublemaker', 'irresponsible', 'unpredictable', 'irrational', 'needy', 'demanding', 'disruptive', 'uncooperative', 'unreliable', 'high maintenance', 'attention-seeker', 'dramatic', 'attention-seeking', 'lazy', 'invalid', 'faker', 'irrational', 'hostile', 'aggressive', 'challenging', 'uncooperative', 'deceptive', 'demanding', 'unreliable', 'high-strung', 'self-destructive', 'unstable', 'manipulative', 'entitled', 'attention-seeking', 'violent', 'drug seeker', 'malingerer', 'faker', 'mentally ill', 'dangerous', 'delusional', 'needy', 'overly sensitive', 'unstable', 'irrational']	'difficult', 'suspicious', 'aggressive', 'unstable', 'dramatic', 'unreliable', 'entitled', 'violent', 'dangerous'	["hysterical", "drug-seeking", "hysterical", "drug-seeking", "junkie", "reluctant", "abuse/abuser", "abused-abuser", "abuser", "abuser's", "abusers", "addictive-drug-seeking", "alcoholic", "angry", "angry-disgusted", "angry/disgusted", "attention-seeker", "attention-seeking", "challenging", "combative", "combatively", "compliant/noncompliant", "counterdefensive", "deceptive", "defensive", "defensive/offensive", "delusional", "demanding", "disruptive", "drug addict", "drug seeker", "drug-craving/drug-seeking", "drug-seeking", "drug-seeking/-taking", "drug-seeking/drug-taking", "drug-seeking/taking", "drug-seeking/use", "drunk", "drunken", "drunkenly", "drunker", "drunkest", "drunks", "ex-abuser", "ex-alcoholic", "faker", "frequent-flier", "frequent-flyer", "frequent-flyers", "frequent-fv", "frequent-hitter", "frequent-hitters", "high maintenance", "high-strung", "histrionic-hysterical", "hostile", "hypochondriac", "hypochondriac-hysterical", "hysterical", "hysterical", "hysterical-obsessive",	127

				<p>'hysterical/anaclitic', 'hystericals', 'hysterics', 'incompliant', 'irrational', 'irrational"', 'irresponsible', 'iv-abuser', 'ivdabuser', 'junkie', "junkie's", 'junkies', 'lazy', 'ma-abuser', 'malingerer', 'manipulative', 'mentally ill', 'mentally unstable', 'morereluctant', 'needy', 'non-adherent', 'non-alcoholic/alcoholic', 'non-compliant', 'non-cooperating', 'non-cooperation', 'non-cooperative', 'non-cooperatively', 'nonadhered', 'nonadherent', 'nonadherently', 'nonadherents', 'noncompliant', 'noncompliant/complian t', 'noncompliant\medicall y', 'noncompliance', 'noncooperating', 'noncooperation', 'noncooperative', 'noncooperatively', 'novelty/drug-seeking', 'ny-nonadherent', 'onadherent', 'overdefensive', 'overly sensitive', 'prealcoholic', 'pt.noncompliant', 'refuse', 'refuses', 'refusing', 'reluctantly', 'reluctant', 'reluctantly', 'reluctants', 'schizo-hysterical', 'self-destructive', 'troublemaker', 'un-adherent', 'unadherent', 'uncooperative', 'unpleasant', 'unpleasant/annoying', 'unpleasantly', 'unpleasantries', 'unpredictable', 'unwilling', 'unwillingly']</p>	
<p>Scare Quotes</p>	<p>((?=.*\.".*\")(?!.*\b(pt patient pateint he she they)\b))</p>	<p>-</p>	<p>-</p>	<p>['yes','no','itchy','h ospital','Do Not Resuscitate', 'Yes or No', 'wet','yes/no','dry', 'DO NOT RESUSCITATE', comfort measures only', 'Known', 'Name', 'firstname', 'lastname','[**Do ctor Last Name **]', '[**Last Name (un) **]',</p>	

				<pre> [**Known firstname **] [**Known lastname **],[**Doctor First Name **], [**Hospital1 **],[**Hospital3 **],[**Known lastname **], [**Known firstname **]] </pre>		
--	--	--	--	--	--	--

Appendix 2: Stigmatizing Labels Ontology

Coding Process

DW and AT met for 1 hour before the first round of coding began. During this time they discussed rationale and literature backgrounds for each of the three linguistic bias features, then proceeded to co-code 5 examples not included in the subsequent datasets. Then, each annotator coded the same random sample of 100. AW and AT met to discuss each of the disagreements in this sample, and used these examples to further inform ontology development for the final sample. Coders were able to reach agreement on all of the linguistic bias features after discussing disagreements.

After calculating agreement and meeting to adjudicate disagreements, AT coded an additional 400 sentences, and DW coded 500 sentences for each linguistic bias term. Table 2 displays the results of the reliability dataset, including frequencies of class labels, as well as final results from the 1000 sentence dataset, with notable sentence notes, chosen out of a selection marked by coders for potential manuscript examples.

Link and Phelan Stigma Definition

Stigma has been defined by social psychologists Link and Phelan as a social process that is characterized by the interplay of

1. **Labeling:** Identifying individuals as belonging to a particular group. Status loss and discrimination involves a negative evaluation of a groups' attributes and its members relative to another group. Commonly used with nouns (static label) or direct adjectives to patient (patient is insistent on receiving water vs patient was repeatedly asking for water). Nouns/direct adjectives work to label patients as static qualities rather than specific, isolated behaviors.
2. **Stereotyping:** ascribing a proclivity towards a specific behavior or characteristic to members of a labeled group.
3. **Separation:** involves distancing from the group, and drawing lines of "us versus them".
4. **Status loss and discrimination:** involve a negative evaluation of a groups' attributes and its members relative to another group. (Comparing this patient with others, or patient vs provider)
5. **Occurs within a context of power,** such as that of the patient-provider relationship.

(Link & Phelan, 2001)

Stigmatizing labels and negative descriptors in charts

Stigmatizing labels to describe groups are often used to perpetuate stereotypes, and when used by providers, can lead to feelings of stigmatization and reduced trust among their patients. Much of the recent work on identifying and reducing stigmatizing labels has come from providers seeking to improve care for patients with substance use disorders. A recent NIDA study published a list of words to avoid using around patients with substance use disorders, including "addict", "abuser", "user", or "junkie", which have been found to be associated with

perceived stigmatization by patients. (Abuse, 2021a) Similar studies have been applied to other chronic illness populations, identifying terms like “sickler”, “frequent flier” or “drug-seeking”, which may be used to further stigmatize patients with chronic illnesses who are often admitted into the hospital. (Abuse, 2021b; Glassberg et al., 2013; Goddu et al., 2018) While some providers may argue that these terms may be useful in flagging unwanted patient behaviors or mental states, a recent study has shown that patients exposed to language written about them by providers which included stigmatizing labels resulted in patients feeling unfairly judged, labeled, and disrespected.(Fernández et al., 2021)

Recent research led by Michael Sun and colleagues on over 40,000 clinical notes has found disparities in presence of “Negative Descriptor” words, evaluated by the Health Equity Commission of the Society of General Internal Medicine, which included commonly used terms in the EHR such as “(non-)adherent, aggressive, agitated, angry, challenging, combative, (non-)compliant, confront, (non-)cooperative, defensive, exaggerate, hysterical, (un-)pleasant, refuse, and resist”. This study found that compared to White patients, Black patients had 2.54 times the odds of having at least one negative descriptor written in their history and physical notes. (Sun et al., 2022) **Research into stereotype expression in language has found that even seemingly innocuous category labels may prompt others to perceive target individual actions and characteristics as “static” aspects of their identity, and exaggerate differences across groups and similarities within them.** (Beukeboom & Burgers, 2017) These labels can be used to justify clinical decision-making, withholding of resources, or to confer doubt upon patient testimonies. (Beukeboom, 2014) While current recommendations encourage use of person-centered, neutral language in medical charts, it is important to evaluate the presence of known stigmatizing labels within provider notes to mitigate the transmission of bias in the EHR.

Guiding Question: Does this sentence involve language about the patient which could result in the stigmatization or negative labeling of a patient, which could lead to further status loss/discrimination in the context of the patient-provider relationship?

Coding Rules

- **Code = 1** : Yes, sentence involves language that could result in the stigmatization of this patient-- i.e. it involves labeling, stereotyping, separation, which could lead to status loss/discrimination in the context of the patient-provider relationship.
 - Yes, clear example:
 - He was demanding dilaudid on admission
 - Patient very needy this shift
 - Patient refusing care. (Broadly)
 - Refusing: difficult to discern,
 - can be used in stigmatizing way related to daily activities or effort put forth by patient
 - If the quote was patient “refusing” a normal daily activity, we code =1.
 - Example: “Patient refusing to put on socks” , which could paint them in a negative, or “stubborn” light. We should instead say “patient not able to put socks on” or “patient does not want socks on”.

- This is bc we are placing direct blame on patients due to lack of effort, without considering other factors like inability
- If the quote was describing how the patient was refusing, related to effort on behalf of the patient, code =1.
 - Examples: “Patient refusing to attempt”, “patient refusing to try”, “patient refusing to cooperate”. These may be more related to patient suffering and not getting care they need, and saying they are refusing to put forth effort can be stigmatizing.
- **Code = 0:** No, the sentence does not involve stigmatizing/negative patient descriptors. Stigmatizing word/negative descriptor is not referring to the patient’s static characteristics or would not likely result in status loss/discrimination among the medical team.
 - These are typically when words are used to describe a patients’ specific behavior vs painting a picture about their character in broad strokes.
 - I.e. if they’re refusing, be specific about what they’re refusing
 - Examples:
 - “Waking for some feeds but not demanding yet.”
 - “O2 sats were unreliable and could not be monitored”
 - “Non-adherent bandage”
 - Instead of calling a patient non-adherent to treatment or noncompliant, which would fall under stigmatizing labels/negative pt descriptors
 - Describing patient acute psychosis or “delusion”

- “Refusing”
 - Refusing a specific medical treatment, without other stigmatizing language or adjectives, would be a 0.
 - Example: “pt refusing to go to CT scan”
 - Refusing DNR status-- well within patient right to refuse, and it is important to clearly understand . These would be labeled 0.
 - “Refusing further care” == 0 when discussed in the context of being do not resuscitate/ do not intubate

Appendix 3: Doubt Markers Ontology

Coding Process

DW and AT met for 1 hour before the first round of coding began. During this time they discussed rationale and literature backgrounds for each of the three linguistic bias features, then proceeded to co-code 5 examples not included in the subsequent datasets. Then, each annotator coded the same random sample of 100. AW and AT met to discuss each of the disagreements in this sample, and used these examples to further inform ontology development for the final sample. Coders were able to reach agreement on all of the linguistic bias features after discussing disagreements.

After calculating agreement and meeting to adjudicate disagreements, AT coded an additional 400 sentences, and DW coded 500 sentences for each linguistic bias term. Table 2 displays the results of the reliability dataset, including frequencies of class labels, as well as final results from the 1000 sentence dataset, with notable sentence notes, chosen out of a selection marked by coders for potential manuscript examples.

Doubt markers overview

Linguistic features such as evidentials, defined as “the linguistic coding of epistemology”,¹ are frequently used along with other words referred to as “doubt markers”, to question the veracity of patient testimonies, particularly related to their symptoms and adherence to treatment.(Park et al., 2021)

Among the many words used as doubt markers, words and expressions used to confer uncertainty such as: *allegedly*, *apparently*, or verbs like *claimed*, are often used when describing patient testimonies, for example: “patient *claimed* their pain was 10/10”.³ These words are often used to discuss the veracity of patient symptoms and adherence to treatment.

Disparities have been found among usage of these terms across race and gender, where patients who were women and patients who were Black were found to have significantly higher frequencies of doubt markers in their provider notes than patients who were men or White.³ The extent to which providers use doubt markers is posited to be reflective of the amount of doubt and uncertainty a provider has on patient testimony, and is thus hypothesized to impact trust within the patient-provider relationship, and related outcomes like leaving against medical advice or in pain management strategies.^{2,3}

Providers may use words when describing patient testimony in combination with stigmatizing labels or negative descriptors of patients to transmit their stance, or expression of attitudes, feelings, and judgment about patients to other providers which may impact future treatment and care decisions.⁴

Guiding Question:

Could this sentence be interpreted by a provider in a way that confers doubt towards the patient's testimony, behavior, or condition?

Coding Rules

doubt_testimony = 1: Yes, it could be interpreted to cast doubt on patient testimony, behavior, or condition.

a. Examples:

- i. “apparently he was sitting at home on the floor feeling fine when suddenly he felt fatigued all over his body,”
- ii. “the patient insists she gets sick from vaccines.”

doubt_testimony = 0: No, this sentence would likely not be interpreted to doubt the patients' testimony, behavior, or condition.

b. These may just reflect uncertainty in medical results/plan

- i. Example: “Diagnosis remains unclear at this time”

c. They could also be instances when the chart is discussing secondhand information about a patient, like “family believes patient is depressed”. Because testimony is secondhand, this is actually appropriate.

d. Other times, the word may be used to accurately portray a patient's own doubt or subjectivity around a situation.

- i. Example: “patient believes they have no options left”

References

1. Chafe WL, Nichols J. *Evidentiality: The Linguistic Coding of Epistemology*. Ablex; 1986.
2. Park J, Saha S, Chee B, Taylor J, Beach MC. Physician Use of Stigmatizing Language in Patient Medical Records. *JAMA Netw Open*. 2021;4(7):e2117052-e2117052.
doi:10.1001/jamanetworkopen.2021.17052
3. 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*. Published online March 22, 2021.
doi:10.1007/s11606-021-06682-z
4. Benamara F, Taboada M, Mathieu Y. Evaluative Language Beyond Bags of Words: Linguistic Insights and Computational Applications. *Comput Linguist*. 2017;43(1):201-264.
doi:10.1162/COLI_a_00278

Appendix 4: Scare Quotes Ontology

Coding Process

DW and AT met for 1 hour before the first round of coding began. During this time they discussed rationale and literature backgrounds for each of the three linguistic bias features, then proceeded to co-code 5 examples not included in the subsequent datasets. Then, each annotator coded the same random sample of 100. AW and AT met to discuss each of the disagreements in this sample, and used these examples to further inform ontology development for the final

sample. Coders were able to reach agreement on all of the linguistic bias features after discussing disagreements.

After calculating agreement and meeting to adjudicate disagreements, AT coded an additional 400 sentences, and DW coded 500 sentences for each linguistic bias term. Table 2 displays the results of the reliability dataset, including frequencies of class labels, as well as final results from the 1000 sentence dataset, with notable sentence notes, chosen out of a selection marked by coders for potential manuscript examples.

Scare Quotes Overview

Another linguistic marker of uncertainty that has been previously identified in patient charts are “scare quotes”, which involve the utilization of quotation marks to mock, cast doubt, challenge patient credibility, or insinuate low health literacy when describing the testimony of another individual.(Beach et al., 2021) While use of scare quotes has been documented since the 1950s, some argue that the linguistic phenomenon has been increasing increasing in recent years, both from the rise of “air quoting” gesture in the 80s and 90s, and was commonly employed by Trump prior to and throughout his presidency.(Garber, 2016; Saner, 2017)

While quotations in charts can be useful to describe patient symptoms using their exact language and document patient wishes or concerns, recent linguistic research has identified a troubling prevalence of providers utilizing quotations in ways to mock, manipulate, and regulate the voices of patients. For example, consider the ambiguity added to the sentence: “Patient reports 10/10 pain related to sickle cell crisis.”, when you add “Patient reports ‘10/10’ pain related to ‘sickle cell crisis’”. Because of the quotation marks, both 10/10 and sickle cell crisis

could be inferred as being untrue or uncertain. Similar to evidentials and negative patient descriptors, scare quotes have been found to be more prevalent among patients who were Black and among patients who were women.(Beach et al., 2021) Early research on scare quoting in patient charts has recommended that providers utilize quotes only when absolutely necessary to reflect a patient’s exact experience, wishes, or concerns, and that even when made in earnest, unnecessary quotation opens patient testimony up to a level of uncertainty or inference to other providers who may question the veracity of patient conditions or experiences.(Beach & Saha, 2021)

The goal of this annotation task is to determine whether these sentences identified as matching closed quotation strings and including reference to patient could be interpreted by other providers as “scare quotes”. We will use a binary system 1/0 to determine whether or not the sentence included a scare quote or not.

We are trying to understand if the use of these quotes could be interpreted to mock, cast doubt, challenge patient credibility, insinuate low health literacy or assign other negative labels to patients. In reading each chart, it’s important to ask: did this need to be quoted? Or could it have been written differently, or more definitively?

Coding Rules

scare_quote = “1”, Could be interpreted as Scare Quote.

- **Yes: Cast doubt on patient as providing reliable testimony.**
 - Examples:
 - Stated “migraine” was due to “stress”. Vs-- Patient stated migraine was brought on by stress.

- **Yes: Convey ridicule, contempt, stigmatization, or frustration** by highlighting **unsophisticated language** or **limited knowledge**

- Examples:

- Patient repeatedly asked to “get me out of this fucking place”
- Does not believe he has prostate cancer because “his bowels are working fine”

Scare_quote = “0”: Not a Scare Quote.

- **Clinical Info, Effect on Life, Values or Preferences. Descriptive of issue**

- Examples:

- Chest pain that “feels like an elephant is on my chest”
- Reported that “this is the worst headache I’ve had in my life”
- When discussing treatment goals, she said “if I cannot breathe without a tube, I don’t want to live. I do not want to suffer. I want to make sure that my family are with me at the end.”

- **Quotes bookending the entire narrative, error in formatting**

- **Describing Acute Psychosis**

- This is a difficult/tricky line, but in the instances where the quotes are used to describe an aspect of a patients’ psychosis, we believe this is medically necessary to communicate, and not a scare quote, despite that it is used to communicate that the patients’ view is unreliable.

- Examples:

- [Pt] then transiently sits up saying clearing "what's going on here"

- Pt able to sleep intermit overnight, however, pt reported having several bad dreams and awoke very disoriented (calling the nurse the "president's dgtr").

Appendix 5: Best Performing Model Hyperparameters

Stigma Model hyperparameters:

- RF: {'max_depth': None, 'min_samples_split': 2, 'n_estimators': 200}
- NB: {'alpha': 1.0}
- Log Reg: {'C': 1.0}
- RoBERTA: {'Max_len':128 , 'Batch_size': 4, 'epochs' : 10 , 'learning_rate': 1e-5}

Doubt Markers Hyperparameters:

- RF: {'max_depth': None, 'min_samples_split': 2, 'n_estimators': 100}
- NB: {'alpha': 1.0}
- Log Reg: {'C': 1.0}
- RoBERTA: {'Max_len':512 , 'Batch_size': 4, 'epochs' : 10 , 'learning_rate': 1e-5}

Scare Quote Model Hyperparameters:

- RF: {'max_depth': None, 'min_samples_split': 5, 'n_estimators': 100}
- NB: {'alpha': 1.0}
- Log Reg: {'C': 0.01}
- RoBERTA: {'Max_len':512 , 'Batch_size': 8, 'epochs' : 10 , 'learning_rate': 1e-5}

Aim 2: Distribution of stigmatizing and doubt-marking language in EHR across patients, providers, and frequently-stigmatized diagnoses

Abstract

Objective: To evaluate the distribution of stigmatizing and doubt-marking language across patient and provider characteristics across a large de-identified EHR dataset.

Background: Electronic health records are a critical medium through which patient stigmatization is created and perpetuated among healthcare teams.

Materials and Methods: We applied supervised learning text classifiers on a large, de-identified EHR dataset. We assessed differences in rates of stigmatizing and doubt marking language across a variety of patient demographic and provider-level predictor variables, and assessed clustering of features in notes across patient and provider levels. **Results:** We found higher rates of stigmatizing labels per chart among patients who were Black or African American (RR: 1.16, 95% CI: 1.08, 1.25), among patients with Medicare/Medicaid or government-run insurance (RR: 2.46, 95% CI: 2.32, 2.61), self-pay insurance (RR: 2.12, 95% CI: 1.45, 2.95), and patients with symptomatic HIV diagnosis, (RR: 2.59, 95% CI: 2.16, 3.08), obesity (RR: 1.98, 95% CI: 1.83, 2.14), and opioid use disorder (RR: 2.79, 95% CI: 2.48, 3.11). Patterns among doubt markers were similar, though male patients had higher rates of doubt markers per chart (RR: 1.25, 95% CI: 1.11, 1.42). Across different provider types, we found significantly increased rates of stigmatizing labels among nurses (RR: 1.40, 95% CI: 1.30, 1.50), and social workers (RR: 2.25, 95% CI: 1.76, 2.84) when compared with physicians, and similar patterns of increased doubt markers. We found significant interaction effects between race and insurance type, showing increased rates of stigmatizing labels and doubt markers among Black/African American patients across all levels of insurance. **Discussion:** Distribution of linguistic bias features in charts

showed consistent trends across historically stigmatized patients. Stigmatization and doubt marking is perpetuated by multiple provider types. This language is likely driven by multiple, intersectional factors contributing to inequitable language use in charts. **Conclusions:** Results and methods from this study can help build future assessments of linguistic stigmatization and care outcomes, and target interventions to reduce stigmatization across provider teams.

Introduction

Language is a powerful medium by which structural marginalization is forged and maintained. Linguistic bias has been defined as “a systematic asymmetry in word choice as a function of the social category to which the target belongs”.¹ Although there are a variety of ways in which linguistic bias can manifest,² they all facilitate the transmission of “essentialist beliefs about social categories”, which work to deindividualize human beings, and assert that members of particular groups contain inherent shared characteristics making them fundamentally distinct from others across situations. While stereotypes have traditionally been studied in the context of intrapersonal beliefs, investigating stigmatization and stereotyping in language can offer us a lens through which we can examine the processes by which stigmas and stereotypes are developed, perpetuated, and shared across groups.³ Recent research summarization efforts in linguistic bias and stigmatization by Beukeboom and Burgers has produced a useful framework, titled the Social Categories and Stereotypes Communication Framework, which can be applied in efforts to understand and investigate these critical processes.³ This framework posits that stereotypes and stigmas, which exist in a shared cognitive space among groups, influence and are influenced by communicative context through which others are discussed. When discussing others, this communicative context is made up of 1) biases in linguistic labeling, which denote and ascribe categories of others as de-individualized members of groups, and 2) biases in how

behaviors and characteristics are described. These processes forge and perpetuate structural marginalization through depersonalization, exaggerating the degree to which “others” are viewed as parts of a distinct group with shared essentialist characteristics, rather than as individuals with unique traits and behaviors.

Healthcare encounters are sites in which providers may deploy stigmatizing language in ways that work to further enforce and create multiple forms of structural marginalization.⁴ Recent research investigating linguistic bias in the medical record has identified salient features of bias within the language of provider notes, namely focused on the expression of degree of trust or doubt in patient testimony through use of stigmatizing labels, as well as “doubt markers”.⁵⁻⁷ Stigmatizing labels to describe groups are often used to perpetuate stereotypes, and when used by providers, can lead to feelings of stigmatization and reduced trust among their patients. Much of the recent work on identifying and reducing stigmatizing labels has come from providers seeking to improve care for patients with substance use disorders. A recent National Institute on Drug Abuse publication described a list of words to avoid using around patients with substance use disorders, including "addict", "abuser", "user", or "junkie", which have been found to be associated with perceived stigmatization by patients.⁸ Similar studies have been applied to other chronic illness populations, identifying terms like “sickler” or “frequent flier” which may be used to further stigmatize patients with chronic illnesses who are often admitted into the hospital.^{6,9,10}

Linguistic features such as doubt markers, or evidentials, which are defined as “the linguistic coding of epistemology”,¹¹ are words that are frequently used in chart language to question the veracity of patients. Among the many words used as evidentials, words and expressions used to confer doubt or uncertainty such as: *allegedly*, *apparently*, or verbs like

claimed, are often used to when describing patient testimonies, for example: “patient *claimed* their pain was 10/10”.⁵

Research on the transmission of provider bias through clinical notes has identified many common manifestations of both negative and positive bias within patient charts,¹² and preliminary research into this text and language-based transmission of bias towards patients has found associations with disparities in care across race, gender, drug use, and other marginalized conditions.^{4,5,12,13} A review of major recent findings of stigmatizing language in healthcare notes across structurally marginalized groups is provided in Table 1.

Table 1: Summary of recent findings related to linguistic stigmatization of structurally marginalized patient groups in the electronic health record

Patient Characteristic	Major findings related to stigmatization in electronic health record
Gender	<p>-Women had higher rates of doubt markers and evidentials than men⁵</p> <p>-Across a variety of conditions, providers distrust women especially when reporting pain and providing pain management.^{14,15}</p>
Race	<p>-Patients who were Black had higher rates of doubt markers and evidentials than patients who were White⁵</p> <p>-Patients who were Black had 2.54 times the odds of having “Negative Descriptor” words such as “(non-)adherent, aggressive, agitated, angry, challenging, combative, (non-)compliant, confront, (non-)cooperative, defensive, exaggerate, hysterical, (un-)pleasant, refuse, and resist”.¹⁶</p>
Socioeconomic Status	<p>-A recent study examining patients with Affordable Care Act coverage or no insurance coverage reported feelings of perceived stigma derived from poor patient-provider communication, where patients often felt belittled or not respected by providers.¹⁷</p> <p>-Similar to the way race and gender impact providers’ beliefs around patient pain, recent studies have shown that providers also</p>

	<p>are less inclined to believe pain from patients of low socioeconomic status.¹⁸</p> <p>-Beyond the significant structural barriers present for low-income patients to access healthcare, it is hypothesized that stigma plays a large role in reducing healthcare access.^{17,19}</p>
Stigmatized Diagnoses	
Sickle cell disease	<p>-Patients with SCD, who are primarily Black, require pain medication during vaso-occlusive crises (VOCs), and often experience discrimination from providers of all types centered around disbelief of pain.²⁰ During VOCs, SCD patients often require high doses of opioids to treat intense pain crises, which fuels the false assumption that patients with SCD are “Drug Seeking”, or are addicted to opioids.²¹</p> <p>-Many surveyed providers endorsed stigmatization of SCD patients, particularly long-term patients who are dubbed “frequent flyers” or “sicklers”, who develop “drug-seeking” reputations among patient teams and may be less readily prescribed opioids to treat their pain crises.²¹</p>
HIV/AIDS	<p>- Patients have reported that during encounters with clinical providers, that stigmatization commonly manifests as providers taking unnecessary or excessive efforts to prevent self-exposure, to using stigmatizing language with them, to denying access to proper treatment or other services.²²</p>
Opioid Use Disorder	<p>-The National Institute on Drug Abuse recently encouraged clinical staff to avoid using terms like “addict”, “abuser”, “junkie”, “dirty/clean” to mitigate stigmatization⁹</p> <p>- Many people, providers included, believe that people with OUD are culpable for their illness, viewed as “addicts”, or people who are “weak” and exhibit “bad character”.²³ A national survey of primary care providers 2021 found high levels of stigmatizing attitudes related to OUD.</p>
Obesity	<p>- The common stereotype among medical providers is that patients with obesity have little willpower and will fail to adhere to provider recommendations on behavior change to lose weight, despite the numerous genetic, environmental, and socio ecological factors that may have led to patients having obesity.²⁴</p> <p>-Research has found that patients with obesity often experience a greater prevalence of explicit stigmatization, often experienced as</p>

	derogatory statements and “cynical humor” directed at patients, who are often described by providers as deserving of the ridicule due to their culpability. ²⁵
--	---

Providers may use words when describing patient testimony in combination with stigmatizing labels or negative descriptors of patients to transmit their stance, or expression of attitudes, feelings, and judgment about patients to other providers which may impact future treatment and care decisions.²⁶ Across the domain of gender, this has led to inequities in treatment and research, where men’s health issues are often treated and investigated more intensively than women.²⁷ These significant inequities of providers’ belief in patients who identify as women have been hypothesized to lead to high amounts of under or mis-diagnosed treatment plans, reduction in provider trust and patient quality of life, and poorer quality of care.¹⁴

Racial biases against patients identified as non-White also heavily contribute to inequalities in pain management decisions and provider perceptions of patient pain.²⁸ This effect can be found even in children, where Black children are 60%, and Hispanic children 50% as likely as White children to be prescribed opioids.²⁹ Another recent study showed significant disparities in analgesic and opioid prescription rates in among a cohort of nearly one million children being treated for appendicitis, wherein Black children had 20% the odds of being prescribed opioids to treat severe pain as White children, and only 10% the odds of being prescribed any analgesic medications when presenting with appendicitis.³⁰ Much of this stigma is centuries old, dating back to the 19th century, through beliefs that Black patients could withstand greater amounts of pain, that their bodies were “tougher” than White bodies, and thus did not need as much pain medication or anesthesia.^{31,32} Despite current awareness of the racist myths perpetuated by the eugenics movement, many of these biases persist subconsciously today in

providers and are still embedded into medical education systems. A 2016 study showed that large proportions of medical students (nearly 50%) believed Black bodies to be biologically different than White bodies, particularly that they are “stronger” or “tougher” than White bodies, and feel less pain.^{28,33}

These socially constructed stigmas and implicit biases are often intertwined with specific diagnosis as well to affect care outcomes. Higher reports of perceived discrimination and provider-based experiences of injustice among patients with sickle cell disease have been shown to be associated with higher pain scores, lower medication adherence, and reduced trust.^{34–36} Experiences of distrust have also been associated with higher rates of hospital self-discharge, associated with a variety of negative clinical outcomes.³⁷ HIV-related stigma has been found to undermine patient access to care, reduce rates of viral load testing, and reduce treatment adherence of antiretroviral and Pre-Exposure Prophylaxis (PrEP) therapies.^{38,39} A recent study on opioid use disorder-related stigma found that this provider-held stigma was negatively associated with the likelihood of prescribing OUD medications, and negatively associated with supporting increased access to Medications for Opioid Use Disorder (MOUD) such as buprenorphine and methadone.⁴⁰ As rates of obesity have risen in the US, patients with obesity report worse communication with physicians via perceiving less respect and attention from providers, and spending less time with physicians.⁴¹ It has been hypothesized that this reduction in time spent with patients has significant impacts on rapport or relationship-building, which may further drive weight-based stigmatization.^{41,42} These jokes create environments that perpetuate stereotypes and negative attitudes and harm patients.²⁴ Consequences of weight-based stigma and biases include deterioration of provider-patient relationships and trust, reduced adherence to medical plans, and delay of care among patients experiencing weight-based bias by their providers.^{24,43,44} Much of

the way provider stigma is experienced by people with a variety of structurally marginalized traits or diagnoses such as sickle cell, HIV, OUD, and obesity happens through provider language; this language reflected in medical charts makes stereotypes more accessible and readily employed across the healthcare team, leading to dehumanization, discrimination, and inequities in health outcomes.⁴⁵

Research into stereotype expression in language has found that even seemingly innocuous category labels may prompt others to perceive target individual actions and characteristics as “static” aspects of their identity, and exaggerate differences across groups and similarities within them.¹ These labels can be used to justify clinical decision-making, withholding of resources, or to confer doubt upon patient testimonies.² In the era of the 21st Century Cures Act, the 2016 health equity legislation which expanded transparency and access to patients’ charts, it is imperative to assess the degree to which language in the electronic health record can perpetuate stigmatization among health care teams and patients themselves.^{46–48}

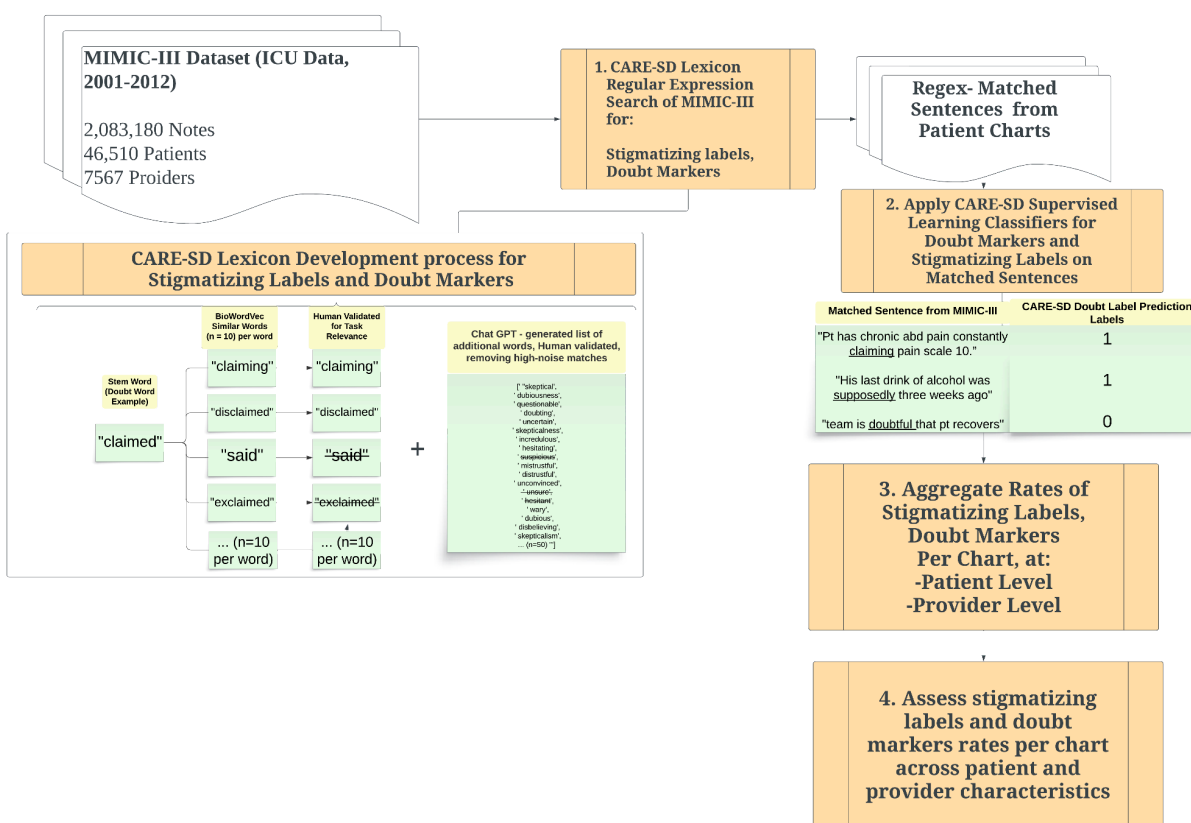
This study aims to examine the distribution of stigmatizing and doubt-marking language features, guided by the Social Categories and Stereotypes Communication framework’s concepts of differential use of labeling and behavioral/characteristic descriptions, across patient and provider-level predictors within a large, de-identified EHR dataset. The identification of stigmatizing and doubt-marking language features will be guided by previous work in development of supervised learning text classifiers.

Methods

In order to assess the distribution of stigmatizing and doubt marking language features across patient and provider predictors, we utilized a natural language processing approach

consisting of: 1) CARE-SD expanded lexicon matching with sentences of EHR free-text notes, 2) application of CARE-SD supervised learning classifiers matched free-text EHR sentences, which provided refined approach to identify presence of linguistic bias features at the chart sentence level, 3) aggregate rates of stigmatizing labels and doubt markers at note, and then the patient and provider levels, 4) Assess distributions and associations across different patient and provider characteristics using Poisson generalized linear regression models. An overview of this pipeline is provided in Figure 1.

Figure 1: Natural language processing pipeline to assess the distribution of stigmatizing labels and doubt markers across a large, de-identified EHR dataset



MIMIC-III Dataset

The Medical Information Mart for Intensive Care, or “MIMIC-III”, is a freely-available database of comprehensive, de-identified EHR, free-text notes, and event documentation for over 40,000 patients admitted to the ICU at Beth Israel Deaconess Medical Center in Boston, MA from 2001 to 2012.⁴⁹ This dataset contains over 1.2 million clinical provider notes, across nearly 50,000 admissions. Because this dataset contains freely-available, EHR from ICU providers from a diverse range of conditions and age ranges, it is a valuable resource for assessing the distribution of stigmatizing and doubt-marking language features.

Data preparation

In the data preparation phase of our study, we employed a comprehensive approach to process and analyze the textual data from the MIMIC-III dataset. Our initial step involved merging regular expression (regex) matches with previously established lexicons to identify instances of stigmatizing labels and doubt markers within the EHR.

CARE-SD: Classifier-based Analysis for Recognizing and Eliminating Stigmatizing and Doubt Marker Labels in Electronic Health Records

Lexicons and supervised learning classifiers for stigmatizing labels and doubt markers were derived from the “CARE-SD: Classifier-based Analysis for Recognizing and Eliminating Stigmatizing and Doubt Marker Labels in Electronic Health Records” toolkit. Previous work in development and validation of this toolkit with the MIMIC-III dataset is provided in greater detail in prior publications, and the toolkit can be found on Physionet and [Github](#).

Stigmatizing labels

Stigmatizing label lexicon development for was guided by literature on stigmatizing language in medical care, specifically from the NIDA “Words Matter” publication, Sun’s “Negative Patient Descriptors: Documenting Racial Bias in the Electronic Health Record”, as

well as Zestcott's "Health Care Providers' Negative Implicit Attitudes and Stereotypes of American Indians".^{8,50,51} The initial stem word list consisted of 18 words:

"abuser", "junkie", "alcoholic", "drunk", "drug-seeking", "nonadherent", "agitated", "angry", "combative", "noncompliant", "confront", "noncooperative", "defensive", "hysterical", "unpleasant", "refuse", "frequent-flier", "reluctant". This list was subsequently expanded using contextual word embeddings models and GPT 3.5 models, and pruned by expert clinical and public health annotators. Supervised learning natural language processing classifiers were developed from an annotated sample derived from MIMIC-III charts with matching sentences containing lexicon terms, resulting in a logistic regression classifier model with 81% accuracy, 75% positive class precision, 84% positive class recall, and .79 macro-F1 score. These findings were comparable to human annotator agreement, which was calculated at 87% (kappa = .74).

Doubt markers

Doubt marker lexicon development was guided by literature on use of "doubt markers" in medical care, specifically led by Beach and colleagues, which identified words such as "claims", "insists", and "adamant" or "apparently", which have been found to be used to discredit or invalidate patient testimony. The 6 words included on the initial stem list were: "adamant", "claimed", "insists", "allegedly", "disbelieves", "dubious".

Appendix 2 provides a final lexicon list used to develop annotation samples used to refine the supervised learning models, for both stigmatizing labels and doubt markers. Supervised learning natural language processing classifiers were developed from an annotated sample derived from MIMIC-III charts with matching sentences containing lexicon terms, resulting in a RoBERTa classifier model with 86% accuracy, 86% positive class precision, 71% positive class

recall, and .84 macro-F1 score. These findings were comparable to human annotator agreement, which was calculated at 87% ($\kappa = .73$).

Analyses and Data Preparation

Our pre-processing steps to prepare the MIMIC-III dataset, included removing all duplicate charts, as well as those labeled as EEG or Radiology, in order to restrict to charts more likely to have subjective narrative and patient history text data.

Following this, we ran previously developed supervised learning classifier models, trained by clinical annotators, to more precisely identify and classify these language features to match instances in which human clinical annotators deemed language to be stigmatizing or doubt-inducing. These supervised learning models were applied on any sentences which contained lexicon matches from the stigmatizing labels and doubt marker word list.

We then merged the predictive sentence-level labels with all sentences in the MIMIC-III dataset. Next, we then aggregated the presence of stigmatizing labels and doubt markers at the note level. We then merged caregiver data from the caregiver table using the Caregiver Identifier (CGID), and patient data from the patients table using the Subject Identifier (SubjectID). Patient insurance and ethnicity were merged from the Admissions table (HADM_ID), and in order to simplify analyses, we selected the first appearing insurance and ethnicity value listed in the dataset for each patient. Patient race/ethnicity free-text categories were organized into a smaller number of distinct categories to facilitate use in regression analyses. We also re-organized the free-text provider type label fields into distinct categories, which was guided by Society of Critical Care Medicine provider type categories, and conducted by IW, who has clinical intensive care unit experience as a PA-C.⁵² Provider type and race/ethnicity labels and categories are provided in Appendix 1. Patient diagnoses were derived as binary variables (having the diagnosis

or not) by linking ICD-9 patient codes from the “diagnoses_icd” table, using regular expressions to capture relevant codes for sickle cell disease , opioid use disorder (OUD), HIV (symptomatic), and obesity.

To facilitate analysis at both the patient and provider levels, we summarized outcomes of stigmatizing labels and doubt markers at these respective levels. Prior to model building, we conducted univariate and bivariate descriptive analyses to assess distribution of outcomes and predictor variables across levels of patients and providers. We also assessed bivariate correlations of the patient and provider level outcomes of stigmatizing labels and doubt markers.

Patient/Provider Clustering

As part of our exploratory data analysis, we calculated median incidence rate ratios using multilevel poisson models in order to compare the clustering of stigmatizing labels and doubt markers within notes at the patient and provider levels. The median incidence rate ratio helps to describe the relative change in the rate of stigmatizing labels and doubt markers per chart, when comparing identical charts from two randomly selected different rate-ordered clusters.⁵³ A value of 1 suggests completely independent note samples, with larger values indicating greater proportion of linguistic feature rates per note variance explained by a given cluster, in this case patient and provider.

Poisson Regressions assessing patient/provider predictors on stigmatizing labels and doubt markers

In order to assess differences in the rates of stigmatizing labels and doubt marker linguistic features per chart, per patient and provider, we conducted Poisson Regression models, assessing each patient and provider-level variables in separate models, while using the offset of the natural

log of total number of charts in order to account for differences in frequencies in charts across patients and providers. The model form for both the patient-level and provider-level model sets are provided below.

Model Set 1 and 2: Patient and provider-level predictors of stigmatizing labels, doubt markers

$$\log(\lambda_{\text{pt or provider}}[\text{Stigmatizing Labels or Doubt Markers}]) = \beta_0 + \beta_1(\text{Predictor})X_{\text{pt or provider}} + \log(\text{Total \# of Charts})$$

Interaction Models

Following the series of models evaluating associations with individual predictors, we conducted a series of interaction models with each of the patient demographic predictors with a new race variable, re-categorized as “Black/African American” and “Not Black/African American”, to assess strata-specific differences among intersectional combinations of race and insurance status, gender, and each of the four stigmatizing diagnoses selected for the study. We then conducted strata-specific estimates to assess the differences in regression betas predicting linguistic feature outcomes among patients who are Black/African American and patients who are not Black/African American.

Results

Tables 1 and 2 offer descriptive statistics on all demographic predictors and linguistic stigma features, summarized from the note to the patient (Table 1) and provider (Table 2) levels. Of note, because the dataset spanned from 2002-2012 and was summarized at the patient level, we did not include Age as a demographic factor of interest due to the inconsistencies in

anchoring it across a large stretch of time. Table 2 provides descriptive statistics for patient rates of stigmatizing labels and doubt markers, as well as rates per note. Both outcomes were found to have non-normal distributions, with high degree of right-skew. Table 3 describes the frequencies of different provider types in this sample of 1,880 providers. Registered nurses (53.4%) and physicians (31.4%) composed the majority of providers in this sample. Table 4 provides descriptive statistics for provider rates of stigmatizing labels and doubt markers, as well as rates per note.

Table 1. Demographic descriptive statistics across patient racial, insurance-based, gender, stigmatizing diagnoses among patients in the MIMIC-III dataset

	Overall (N=11630)
Race/Ethnicity*	
White	8312 (71.5%)
Asian	353 (3.0%)
Black/African American	945 (8.1%)
Hispanic/Latino	440 (3.8%)
Native American/Alaskan Native	16 (0.1%)
Other	345 (3.0%)
Unknown/Declined	1219 (10.5%)
Insurance	
Private	4509 (38.8%)
Government^	353 (3.0%)
Medicaid	1211 (10.4%)
Medicare	5450 (46.9%)
Self Pay	107 (0.9%)
Gender	
Female	5086 (43.7%)
Male	6544 (56.3%)
Diagnoses	
Sickle Cell Disease	25 (0.2%)
Opioid Use Disorder	217 (1.9%)
Obesity	674 (5.8%)
Symptomatic HIV	115 (1.0%)

*Race/ethnicity simplified due to low cell-count for multiple values. List of categorizations provided in Appendix 1.

^Government insurance represents non-Medicare and non-Medicaid types of government-supported insurance, including programs from the Department of Defense TRICARE, Veterans Health Administration program, or Indian Health Service

Table 2: Patient-level descriptive statistics on rates of stigmatizing labels and doubt markers per EHR note within the MIMIC-III dataset

	Frequencies (N=11630 patients)
Stigmatizing Labels Count Per Patient	
Mean (SD)	.5 (1.99)
Median [Min, Max]	0 [0, 90]
Doubt Marker Labels Count Per Patient	
Mean (SD)	0.09 (0.48)
Median [Min, Max]	0 [0, 29]

*Race/ethnicity simplified due to low cell-count for multiple values. List of categorizations provided in Appendix 1.

Table 3. Frequencies of provider types within the MIMIC-III dataset

	Overall (N=1880)
Provider Types*	
APPs (NPs, PA-Cs)	31(1.6%)
Pharmacists	4 (0.2%)
Physicians	590 (31.4%)
Registered Dieticians	23 (1.2%)
Registered Nurses	1004 (53.4%)
Rehab (OTs/PTs)	49 (2.6%)
Respiratory Therapists	42 (2.2%)
Social Workers	47 (2.5%)
Unknown**	90 (4.8%)

*Provider types simplified due to low cell-count for multiple values. List of categorizations provided in Appendix 1.

**Providers with unknown category excluded from regressions

Table 4. Provider-level descriptive statistics on rates of stigmatizing labels and doubt markers within the MIMIC-III dataset

	Overall (N=1880)
Stigmatizing Labels Count Per Provider	
Mean (SD)	3.09 (14.40)
Median [Min, Max]	0 [0, 557]
Doubt Marker Labels Count Per Provider	
Mean (SD)	0.56 (5.58)
Median [Min, Max]	0 [0, 236]

During exploratory data analysis, we assessed bivariate associations between stigmatizing labels and doubt markers with datasets summarized at the patient, and then provider levels.

Spearman rank correlation test showed significant moderate correlation between doubt marker classification labels and stigmatizing classification labels at the patient level (Rho = .1887, $p < .001$). At the provider level, we found higher correlation between doubt marker classification labels and stigmatizing classification levels at the provider level (Rho = .4459, $p < .001$).

After calculating the mean incidence rate ratio to describe the levels of outcome clustering, we found the highest degree of clustering for stigmatizing labels per chart to occur at the patient level (Median IRR: 7.08), and the highest degree of clustering for doubt markers per chart at the patient level as well (Median IRR: 147.8). Provider level Median IRRs, while still greater than 1, were lower for both features, with a Median IRR for stigmatizing labels at 3.78, and Median IRR for doubt markers at 4.38. These results provide further support for conducting regression models to assess different characteristics related to stigmatizing and doubt-marking language at the patient and provider levels.

Patient-level models assessing the relationships between various demographic predictors and counts of stigmatizing classifier labels and doubt markers across all patient charts revealed several patterns. These results are displayed in Table 3. Male patients were found to receive 1.25 times the rate of doubt markers per chart than females (95% CI: 1.11, 1.42, $p < .0001$). With regards to ethnicity and race, we found that when compared to White patients, Black or African

American patients received 1.16 higher rates of stigmatizing labels per chart. All other patients identified with non-White ethnic/racial groups received fewer counts of stigmatizing labels, with patients identified as Native American/Alaskan Natives receiving the lowest (.24 times the rate of White patients). Additionally, patients who were identified as Asian, Other, or Unknown/Declined had significantly fewer rates of doubt marker labels per charts than patients who were identified as White.

Insurance type was also found to have significant associations with both types of stigmatizing linguistic features. Compared to patients with private insurance, patients with Government-run insurance (encompassing all government employee care, Medicaid, or Medicare) had 2.46 times as many Stigmatizing Labels per chart and 2.32 times as many Doubt Markers. Similarly, patients with no insurance, labeled “Self Pay”, had 2.12 times as many Stigmatizing Labels per chart, and 4.94 as many Doubt Markers per chart as patients with Private insurance. Additionally, 3 of the 4 frequently stigmatized disease diagnoses significantly predicted higher rates of Stigmatizing Labels and Doubt Markers per charts. Patients with ICD-9 codes for symptomatic HIV had 2.59 times the rate of Stigmatizing Labels per chart, and 2.40 times the rate of Doubt Markers per chart, compared with patients without HIV. Patients with obesity had 1.98 times higher rates of Stigmatizing labels per chart, and 2.14 times higher rates of doubt markers per chart. Patients with opioid use disorder experienced the highest differences in rates of both stigmatizing labels per chart (2.79 times higher) and doubt markers per chart (3.80 times higher). Finally, patients with Sickle Cell Disease had no significant difference in either type of stigmatizing linguistic feature compared to patients without Sickle Cell Disease.

Patient-level models- demographic predictors

Table 3. Poisson regression results showing relationships between demographic patient

predictor variables associations with stigmatizing linguistic EHR features per chart. Rate Ratios (95%CI)

	Stigmatizing Classifier Labels	Doubt Marker Classifier Labels
Gender (Ref = Female)	1.02 (.97, 1.08)	1.25 (1.11, 1.42)**
Ethnicity (Ref = White)		
Asian	0.54 (0.45, 0.65)**	0.46 (0.27, 0.72)**
Black/African American	1.16 (1.08, 1.25)**	1.06 (0.88, 1.27)
Hispanic/Latino	0.74 (0.63, 0.86)**	0.78 (0.53, 1.10)
Native American/Alaskan Native	0.24 (0.08, 0.56)*	1.01 (0.25, 2.63)
Other	0.42 (0.34, 0.52)**	0.49 (0.30, 0.75)*
Unknown/Declined	0.66 (0.59, 0.73)**	0.75 (0.59, 0.93)*
Insurance (Ref= Private)		
Government-run [^]	2.46 (2.32, 2.61)**	2.32 (2.03, 2.67)**
Self-Pay	2.12 (1.45, 2.95)**	4.94 (2.69, 8.25)**
Diagnoses		
HIV (Symptomatic)	2.59 (2.16, 3.08)**	2.40 (1.51, 3.60)**
Obesity	1.98 (1.83, 2.14)**	2.14 (1.78, 2.55)**
Opioid Use Disorder	2.79 (2.48, 3.11)**	3.80 (2.98, 4.77)**
Sickle Cell Disease	0.41 (0.10, 1.07)	1.54 (0.26, 4.75)

***p is significant at <.05 value**

****p is significant at <.0001 value**

[^] Government-run includes the MIMIC-III insurance categories of “Government”, “Medicare”, and “Medicaid”

Provider type models also showed significant differences across provider types in the number of stigmatizing linguistic features per charts. Of note, we excluded Pharmacists from these analyses due to low cell sizes (n = 4). These results are displayed in Table 4.

Compared with Physicians, several groups of providers were found to have significantly different rates of stigmatizing label and doubt markers per chart.

Advanced Practice Providers (including Physician Associates and Nurse Practitioners) had 0.15 times the rates of Stigmatizing labels per chart, and only .04 times the rates of Doubt Marker labels per chart. Registered Dieticians were also found to have .25 times the rate of stigmatizing labels per chart as Physicians.

On the other hand, Registered Nurses had 1.40 times higher rates of Stigmatizing labels and 1.46 times higher rates of doubt marker labels than physicians. Rehab care team members, consisting of Occupational and Physical Health Therapists, were found to have 2.27 times the rate of doubt markers as Physicians. Finally, Social Workers reported the highest rates of stigmatizing labels and doubt markers, using 2.25 times the rate of stigmatizing labels per chart, and 5.27 times the rate of doubt markers per chart. Respiratory therapists were found to use both stigmatizing linguistic features at similar rates to Physicians.

Provider-level models- provider type

Table 4. Poisson regression results, modeled at the provider level, showing relationships between provider types with stigmatizing linguistic EHR features per chart. Rate Ratios (95%CI)

	Stigmatizing Classifier Labels	Doubt Marker Classifier Labels
Provider Type (Ref = Physicians)		
Advanced Practice Providers (NPP, PA-C)	0.15 (0.10, 0.21)**	0.04 (0.00, 0.17)**
Registered Dieticians	0.25 (0.15, 0.37)**	1.18 (0.63, 2.00)
Registered Nurses	1.40 (1.30, 1.50)**	1.46 (1.21, 1.77)**
Rehab (OT/PT)	1.18 (0.86, 1.58)	2.27 (1.19, 3.92)*
Respiratory Therapists	0.92 (0.65, 1.26)	0.35 (0.06, 1.09)
Social Workers	2.25 (1.76, 2.84)**	5.27 (3.33, 7.98)**

***p is significant at <.05 value**

****p is significant at <.0001 value**

Pharmacists removed from regression analyses due to low cell size (n = 4)

Following patient and provider-level predictor models, we also sought to examine interactional effects to test for differences among more intersectional identities. In order to compensate for low racial/ethnic group cell sizes, we re-categorized our race variables into a binary Black or African American indicator vs non-Black/African American. We then conducted

interaction models with gender, insurance type, and each of the 4 stigmatizing diagnoses. All models included all predictors in Table 3 as covariates (Gender, Black/African American indicator, Insurance Type, and Diagnosis). Table 4 presents the results of our interaction models and strata-specific estimates among patients who were and were not Black/African American.

We observed a significant interaction effect among the Black/African American and Non-Black/African American indicator and insurance type (Chi-Sq = 30.72, $p < .0001$), where patients who were Black/African American with Government-run insurance and patients who were Black/African American, and self-pay had significantly higher rates of stigmatizing labels and doubt markers per chart than non-Black/African American patients with the same insurance types. A similar pattern emerged comparing the outcomes of Doubt markers, where we observed a significant interaction between the Black/African American race indicator and insurance type on the outcome of doubt markers (Chi-Sq = 6.54, $p = .038$).

Additionally, we found a significant interaction between the Black/African American race indicator and symptomatic HIV diagnosis (Chi-Sq = 5.61, $p = .018$), where Black/African American patients with HIV had higher rates of stigmatizing labels per charts than non-Black patients with symptomatic HIV. No other significant interactions were detected across the other variables assessed in interaction models.

Patient-level models interaction model estimates with Black + African American /Non-Black + African American race variable

Table 5. Patient-level interaction model estimates, with each predictor in the left-hand column interacting with a simplified Black/Non-Black racial category. Rate Ratios (95%CI)

Stigmatizing	Stigmatizing	Int Sig	Doubt Marker	Doubt Marker	Int Sig
--------------	--------------	---------	--------------	--------------	---------

	Classifier Labels (Non-Black/AA)	Classifier Labels Black/AA	Chi-Sq (p-val)	Classifier Labels (Non-Black/AA)	Classifier Labels Black/AA	Chi-Sq (p-value)
Gender (Ref = Female)	-	-	-	-	-	-
Insurance (Ref= Private)						
Government-run [^]	2.21 (2.08, 2.35)**	4.03 (3.21, 5.08)*	30.72	-	-	-
Self-Pay	1.63 (1.07, 2.48)*	5.67 (2.74, 11.73)**	(p<.0001)	-	-	-
Diagnoses						
HIV (Symptomatic)	1.58 (1.25, 2.01)*	2.47 (1.87, 3.25)**	5.61 (p = .018)	-	-	-
Obesity	-	-	-	-	-	-
Opioid Use Disorder	-	-	-	-	-	-
Sickle Cell Disease	-	-	-	-	-	-

All models include covariates of gender, insurance, and each of the four stigmatizing diagnoses.

***p is significant at <.05 value**

****p is significant at <.0001 value**

[^] Government-run includes the MIMIC-III insurance categories of “Government”, “Medicare”, and “Medicaid”

- Indicates Null Interaction Effect

Discussion

The differences in distributions of stigmatizing labels and doubt markers within electronic health records largely fit the pattern of demonstrating disproportionately higher rates across historically stigmatized and marginalized groups. These two linguistic features, stigmatizing labels and doubt markers, can be conceptualized as EHR operationalizations of the Social Categories and Stereotypes Communication constructs of “biases in labeling” and “biases and describing behavior”, which provide evidence of the viability of this linguistic framework’s application in the healthcare field.

Results from correlations between doubt markers and stigmatizing labels indicate that these features are frequently used together, and may in tandem work to further stigmatize or invalidate patients’ testimonies. Our finding of 25% higher rates of doubt markers among notes of male patients compared to female patients stands in contrast to previous work, which

identified higher rates of doubt marking language and pain disbelief among female patients.^{5,15}

While the current study stands to illuminate trends on a broad level, our doubt marker classification tools can be used to draw more targeted samples to guide further in-depth qualitative inquiry to identify the specific clinical contexts in which doubt marking language may be differentially applied by gender. There may be certain diagnoses or clinical situations in this ICU setting, such as alcohol or substance-related conditions, which occur more frequently among men and may also coincide with higher rates of doubt marker usage.⁵⁴ In these contexts, providers often convey degrees of doubt in charts for patient-reported rates of substance use. Still, it is important to weigh doubt marker use against the possibility that these terms may further erode trust among other providers in patients who are already stigmatized as being distrustful by nature of their histories of substance use.

Findings showing significant differences across race identified that Black/African American patients received an estimated 16% higher rates of stigmatizing classifier labels per note, when compared to White patients. This finding is in line with previous work which has identified disproportionately higher rates of stigmatizing or doubt-marking language in charts of patients who are Black or African American, as well as broader trends of provider-held implicit biases.^{28,30} Given the long history of racial discrimination, health inequities, and broken trust in the healthcare system among Black and African American patients, it is especially important to avoid stigmatizing and doubt marking terms among these patients, and employ more neutral, patient-centered language. All other non-White racial and ethnic groups received significantly less stigmatizing labels, and equal or significantly less doubt markers. These differences may require additional testing across different patient populations, especially given that previous studies assessing trends across racial groups within MIMIC-III frequently simplified racial

categories to Black vs non-Black or into Black, White, and Other due to the low frequency of other minority populations within the sample.^{55,56}

Insurance, as a proxy indicator for socioeconomic status, revealed higher rates of stigmatizing labels among patients receiving government assistance (Medicare, Medicaid) for insurance, and especially for patients with no insurance, identified as “Self Pay”, when compared with patients with private insurance. This increase is greater in the latter group, which could be conceptualized as a group with the lowest socioeconomic status.⁵⁷⁻⁵⁹ Still, we recognize that insurance is an imperfect indicator of socioeconomic status, and encourage future work to incorporate a greater variety of indicators to understand its role in driving patient stigmatization and doubt marking.

Of the 4 frequently stigmatized diagnoses selected for this study, we found significantly higher rates of stigmatizing labels and doubt markers across 3 of them, with the exception of sickle cell disease. Disproportionately higher rates were found for patients with obesity, symptomatic HIV, and were highest among patients with opioid use disorder. These trends were largely congruent with a wealth of research documenting experiences of stigmatization, ridicule, and testimonial injustice across these groups.^{5,41,60} These trends underscore an increased need for systematic interventions to reduce provider stigmatization across these groups. An absence of difference among patients with sickle cell disease may also be driven by an overall low sample size of patients with SCD within this MIMIC-III dataset sample (n = 25). Additionally, much of the previous research focused on this population has been centered within the first point of hospital contact, the emergency department, which may be staffed with providers unfamiliar with the disease or specific patient needs, or driven by more salient ED factors like high patient volume and rapid triaging efforts.^{61,62} While this study saw null results in differences of rates of

stigmatizing and doubt marking features, future studies exploring SCD patient stigmatization should continue to focus on ED and centers which care for higher rates of patients with the disease.

Our findings from the series of interaction models underscore the importance of incorporating intersectionality in the study and intervention of stigmatizing and doubt-marking language in charts, particularly along dimensions of race and socioeconomic status.^{63,64} Patients who were Black or African American received disproportionately higher rates of stigmatizing labels than patients who were not Black or African American, across equal levels of insurance which are often associated with lower SES, i.e. Medicare, Medicaid, and Self-pay. Due to the constraints of the data available, our categorical indicators utilized in interaction models were broad and many specific marginalized racial or SES groups were likely not fully described. In light of these constraints, we contend that these interaction effects likely underestimate the full disparity of stigmatizing and doubt-marking language among patients with multiple stigmatizing identity characteristics.

Findings comparing median incident rate ratios across both stigmatizing labels and doubt markers indicated drastically higher rates of clustering by patient than by provider, for both types of linguistic features. This stark difference shows that while trends in language use are clustered around specific providers, the language around patients is far more consistent. This supports the Social Categories and Stereotypes Communication central assumption that stereotypes are a result of shared cognitive and group-level processes, which may communicate information that is not necessarily wholly supported by individuals alone.³ For these reasons, it is important to consider the role of ingroup dynamics among a medical team, which is useful to understand how it may be “relationally beneficial” to perpetuate a stereotype, generating similarity and

agreeableness among healthcare team employees.⁶⁵ Previous stereotype consensualization research has shown that once a stereotype is shared or known to be believed within groups, the process of refuting the stereotype requires greater processing resources and time, and puts the individual at greater risk for conflict than it does to agree.⁶⁶ Additionally, the often necessary process of chart segment duplication and replication, which saves medical teams time and carries pertinent information forward through the EHR, may work to make it even harder to change the stereotyping process in how patients are labeled and how their actions are described.^{67,68} This highlights the need for systems in place such as the ones developed and tested within this study to aid in automatically identifying and classifying stigmatizing language as it happens in real time.

Of all providers, advanced practice practitioners and dieticians in this sample had significantly lower rates of stigmatizing labels, and for APPs, as low as 4% the rate of doubt markers when compared with Physicians. While these results are encouraging, future work is needed to assess these distributions with higher sample sizes for these specific roles. Registered nurses, on the other hand, employed stigmatizing labels and doubt markers at 40 and 46% higher rates than physicians. Out of all of the provider types identified in this sample, nurses are most likely to spend the highest amount of bedside time with patients.⁶⁹ It is possible, that while spending time with patients labeled as “difficult”, that they may endure greater frequencies of negative behaviors in forms of verbal or physical abuses from patients and families that may arise out of patient frustrations or conditions.⁷⁰ Additionally, due to nurses’ lower status within the hierarchy of medical teams, they may have higher motivation to generate in-group bonding among the medical team, support beliefs in agreement with higher-status team members, and face greater risks with going against shared team stereotypes and assumptions about patients.⁷¹

Within our sample, social workers had the highest overall rates of both stigmatizing and doubt markers usage within charts. As ancillary health professionals within the healthcare team chain of command, social workers may face similar pressures to agree or conform to group norms as nurses. By nature, social work charts are also often more subjective in nature, which may more frequently describe patient testimonies. Additionally, language may describe whether or not patients qualify for a variety of resources, in which issues like substance use may be frequently described with doubt markers and/or stigmatizing labels.⁷² Despite differences between provider types, it is important for all medical professionals to understand the power of their language in charts, and work to ensure that language does not work to manipulate patient narratives and cast undue doubt or perpetuate stigmatization for marginalized patient populations.^{73,74}

Limitations

While the results of this paper have illuminated important trends in the distribution of stigmatizing linguistic features in the EHR, it is not without limitations. Firstly, due to the predominantly White sample of MIMIC-III, our race and ethnicities had to be simplified in order to incorporate these factors into analyses. This simplification, however, likely results in some degree of erasure or potentially grouping of ethnicities and races which may have very different histories of stigmatization within the medical setting. MIMIC-III may also be limited in the degree of accuracy and granularity of reporting detailed characteristics related to patient identity (i.e. race/ethnicity), as it is encoded primarily by medical staff during admission intake. Still, this variable may better represent how race/ethnicity is perceived and potentially stigmatized by medical staff. Analyses in this study also did not incorporate patient age or time into account,

due to the difficulty in anchoring time at the patient and provider levels. Further work incorporating how time affects the accumulation of stigmatizing and doubt-marking linguistic features is needed to see how patient reputations may emerge and change over time. It is likely that there may also be subtle changes in the language of EHR which occurred over the entire database due to its large range of time from 2001 to 2012. Finally, understanding provider type roles in this context should also seek to incorporate student status, as they represent key mediums through which the medical “hidden curriculum” is transferred from one generation of medical professionals to another.⁷⁵ This current study was limited in not including student status, which proved difficult to abstract exact student types and statuses due to the limitation of provider abbreviations and labels in MIMIC-III.

Impact

This study has successfully demonstrated the utility of stigmatizing language detection models to assess differences in how patients from a variety of backgrounds and diagnoses are discussed by providers within the EHR. Trends identified in this study should underscore the strong and present threat of stigmatizing and doubt marking language that exists disproportionately among highly marginalized patient groups. The tools utilized in this study have been guided by previous work in stigma and healthcare, and represents a significant step forward by scaling up previous linguistic work by leveraging advances in natural language processing on a large, freely accessible dataset. We believe that similar analyses utilizing the stigma and doubt detection tools and modeling strategies can be applied to assess the distribution of linguistic stigmatizing and doubt-marking features across a wide variety of other EHR datasets and care outcomes. Additionally, tools like the CARE-SD can be adapted and updated

with rapidly evolving natural language processing tools in the future. Large language models, which were implemented in the CARE-SD lexicon development stage, could also be applied through pretraining the models to query more nuanced questions about patient stigmatization patterns, summarize common situations in which stigmatizing language and doubt markers are used, or be applied to destigmatizing medical language once identified. The CARE-SD and MIMIC-III datasets and existing work can provide foundational models and targeted annotation datasets to accelerate future work in this area. In an era of increasing transparency in medical charts, healthcare systems should use tools like the ones implemented in this study to conduct continuous real-time audits of stereotyping and stigmatization in the EHR system. These methods offer us an important lens through which we can assess and intervene on the structural marginalization of our most vulnerable patient populations and reduce healthcare inequities.

References

1. Beukeboom CJ, Burgers C. Linguistic Bias. In: *Oxford Research Encyclopedia of Communication*. Oxford University Press; 2017.
doi:10.1093/acrefore/9780190228613.013.439
2. Beukeboom CJ. Mechanisms of linguistic bias: How words reflect and maintain stereotypic expectancies. In: *Social Cognition and Communication*. Sydney symposium of social psychology. Psychology Press; 2014:313-330.
3. Beukeboom CJ, Burgers C, Reviews R of CROAHQL. How Stereotypes Are Shared Through Language - A Review and Introduction of the Social Categories and Stereotypes Communication (SCSC) Framework. *Rev Commun Res*. Published online January 1, 2019.

Accessed February 26, 2024.

https://www.academia.edu/38218166/Beukeboom_and_Burgers_2019_How_Stereotypes_Are_Shared_Through_Language_A_Review_and_Introduction_of_the_Social_Categories_and_Stereotypes_Communication_SCSC_Framework

4. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. *BMC Med Ethics*. 2017;18(1):19. doi:10.1186/s12910-017-0179-8
5. 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*. Published online March 22, 2021. doi:10.1007/s11606-021-06682-z
6. Goddu AP, O'Connor KJ, Katie J. O'Connor, 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
7. Zhang H, Lu AX, Abdalla M, McDermott M, Ghassemi M. Hurtful words: quantifying biases in clinical contextual word embeddings. In: *Proceedings of the ACM Conference on Health, Inference, and Learning*. CHIL '20. Association for Computing Machinery; 2020:110-120. doi:10.1145/3368555.3384448
8. Abuse NI on D. Words Matter - Terms to Use and Avoid When Talking About Addiction. National Institute on Drug Abuse. Published November 29, 2021. Accessed January 22, 2022. <https://www.drugabuse.gov/nidamed-medical-health-professionals/health-professions-education/words-matter-terms-to-use-avoid-when-talking-about-addiction>
9. Abuse NI on D. Words Matter - Terms to Use and Avoid When Talking About Addiction. National Institute on Drug Abuse. Published November 29, 2021. Accessed June 14, 2022. <https://nida.nih.gov/nidamed-medical-health-professionals/health-professions-education/word>

s-matter-terms-to-use-avoid-when-talking-about-addiction

10. Glassberg J, Tanabe P, Richardson L, DeBaun M. Among emergency physicians, use of the term “Sickler” is associated with negative attitudes toward people with sickle cell disease. *Am J Hematol*. 2013;88(6):532-533. doi:10.1002/ajh.23441
11. Chafe WL, Nichols J. *Evidentiality: The Linguistic Coding of Epistemology*. Ablex; 1986.
12. Park J, Saha S, Chee B, Taylor J, Beach MC. Physician Use of Stigmatizing Language in Patient Medical Records. *JAMA Netw Open*. 2021;4(7):e2117052-e2117052. doi:10.1001/jamanetworkopen.2021.17052
13. Ashford RD, Brown AM, Curtis B. The Language of Substance Use and Recovery: Novel Use of the Go/No–Go Association Task to Measure Implicit Bias. *Health Commun*. 2019;34(11):1296-1302. doi:10.1080/10410236.2018.1481709
14. Govender V, Penn-Kekana L. Gender biases and discrimination: a review of health care interpersonal interactions. *Glob Public Health*. 2008;3(sup1):90-103. doi:10.1080/17441690801892208
15. Hoffmann DE, Tarzian AJ. The Girl Who Cried Pain: A Bias against Women in the Treatment of Pain. *J Law Med Ethics*. 2001;29(1):13-27. doi:10.1111/j.1748-720X.2001.tb00037.x
16. Michael Sun, Tomasz Oliwa, Peek ME, Elizabeth L. Tung. Negative Patient Descriptors: Documenting Racial Bias In The Electronic Health Record. *Health Aff (Millwood)*. Published online January 19, 2022:101377hlthaff202101423-101377hlthaff202101423. doi:10.1377/hlthaff.2021.01423
17. Allen H, Wright BJ, Harding K, Broffman L. The Role of Stigma in Access to Health Care for the Poor. *Milbank Q*. 2014;92(2):289-318. doi:10.1111/1468-0009.12059

18. Summers KM, Deska JC, Almaraz SM, Hugenberg K, Lloyd EP. Poverty and pain: Low-SES people are believed to be insensitive to pain. *J Exp Soc Psychol.* 2021;95:104116.
doi:10.1016/j.jesp.2021.104116
19. Reutter LI, Stewart MJ, Veenstra G, Love R, Raphael D, Makwarimba E. “Who Do They Think We Are, Anyway?”: Perceptions of and Responses to Poverty Stigma. *Qual Health Res.* 2009;19(3):297-311. doi:10.1177/1049732308330246
20. Jenerette CM, Brewer C. Health-Related Stigma in Young Adults With Sickle Cell Disease. *J Natl Med Assoc.* 2010;102(11):1050-1055. doi:10.1016/S0027-9684(15)30732-X
21. Masese R, Bulgin D, Douglas C, Shah N, Tanabe P. Barriers and facilitators to care for individuals with sickle cell disease in central North Carolina: The emergency department providers’ perspective. *PLOS ONE.* 2019;14:e0216414. doi:10.1371/journal.pone.0216414
22. Fredericksen R, Edwards T, Crane HM, et al. Patient and provider priorities for self-reported domains of HIV clinical care. *AIDS Care.* 2015;27(10):1255-1264.
doi:10.1080/09540121.2015.1050983
23. Cheetham A, Picco L, Barnett A, Lubman DI, Nielsen S. The Impact of Stigma on People with Opioid Use Disorder, Opioid Treatment, and Policy. *Subst Abuse Rehabil.* 2022;13:1-12.
doi:10.2147/SAR.S304566
24. Phelan SM, Burgess DJ, Yeazel MW, Hellerstedt WL, Griffin JM, van Ryn M. Impact of weight bias and stigma on quality of care and outcomes for patients with obesity. *Obes Rev Off J Int Assoc Study Obes.* 2015;16(4):319-326. doi:10.1111/obr.12266
25. Wear D, Aultman JM, Varley JD, Zarconi J. Making Fun of Patients: Medical Students’ Perceptions and Use of Derogatory and Cynical Humor in Clinical Settings. *Acad Med.* 2006;81:454-462. doi:10.1097/01.ACM.0000222277.21200.a1

26. Benamara F, Taboada M, Mathieu Y. Evaluative Language Beyond Bags of Words: Linguistic Insights and Computational Applications. *Comput Linguist.* 2017;43(1):201-264. doi:10.1162/COLI_a_00278
27. Holdcroft A. Gender bias in research: how does it affect evidence based medicine? *J R Soc Med.* 2007;100(1):2-3. doi:10.1177/014107680710000102
28. Hoffman KM, Trawalter S, Axt JR, Oliver MN. Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites. *Proc Natl Acad Sci.* 2016;113(16):4296-4301. doi:10.1073/pnas.1516047113
29. Groenewald CB, Rabbitts JA, Hansen E, Palermo TM. Racial differences in opioid prescribing for children in the United States. *Pain.* 2018;159(10):2050-2057. doi:10.1097/j.pain.0000000000001290
30. Gopal DP, Chetty U, O'Donnell P, Gajria C, Blackadder-Weinstein J. Implicit bias in healthcare: clinical practice, research and decision making. *Future Healthc J.* 2021;8(1):40-48. doi:10.7861/fhj.2020-0233
31. Pernick MS. *A Calculus of Suffering: Pain, Professionalism, and Anesthesia in Nineteenth-Century America.* NY: Columbia University Press; 1985. Accessed November 30, 2022. <http://deepblue.lib.umich.edu/handle/2027.42/116015>
32. Trawalter S, Hoffman KM. Got Pain? Racial Bias in Perceptions of Pain. *Soc Personal Psychol Compass.* 2015;9(3):146-157. doi:10.1111/spc3.12161
33. Green AR, Carney DR, Pallin DJ, et al. Implicit Bias among Physicians and its Prediction of Thrombolysis Decisions for Black and White Patients. *J Gen Intern Med.* 2007;22(9):1231-1238. doi:10.1007/s11606-007-0258-5
34. Haywood C, Lanzkron S, Bediako S, et al. Perceived Discrimination, Patient Trust, and

- Adherence to Medical Recommendations Among Persons with Sickle Cell Disease. *J Gen Intern Med.* 2014;29(12):1657-1662. doi:10.1007/s11606-014-2986-7
35. Ezenwa MO, Molokie RE, Wilkie DJ, Suarez ML, Yao Y. Perceived Injustice Predicts Stress and Pain in Adults with Sickle Cell Disease. *Pain Manag Nurs.* 2015;16(3):294-306. doi:10.1016/j.pmn.2014.08.004
36. Elander J, Beach MC, Jr CH. Respect, trust, and the management of sickle cell disease pain in hospital: comparative analysis of concern-raising behaviors, preliminary model, and agenda for international collaborative research to inform practice. *Ethn Health.* 2011;16(4-5):405-421. doi:10.1080/13557858.2011.555520
37. Haywood C, Lanzkron S, Ratanawongsa N, Bediako SM, Lattimer-Nelson L, Beach MC. Hospital Self-Discharge Among Adults with Sickle Cell Disease: Associations with Trust and Interpersonal Experiences with Care. *J Hosp Med Off Publ Soc Hosp Med.* 2010;5(5):289-294. doi:10.1002/jhm.643
38. Penner LA, Phelan SM, Earnshaw V, Albrecht TL, Dovidio JF. Patient stigma, medical interactions, and health care disparities: A selective review. In: *The Oxford Handbook of Stigma, Discrimination, and Health.* Oxford library of psychology. Oxford University Press; 2018:183-201.
39. Katz IT, Ryu AE, Onuegbu AG, et al. Impact of HIV-related stigma on treatment adherence: systematic review and meta-synthesis. *J Int AIDS Soc.* 2013;16(3S2):18640. doi:10.7448/IAS.16.3.18640
40. Stone EM, Kennedy-Hendricks A, Barry CL, Bachhuber MA, McGinty EE. The role of stigma in U.S. primary care physicians' treatment of opioid use disorder. *Drug Alcohol Depend.* 2021;221:108627. doi:10.1016/j.drugalcdep.2021.108627

41. Richard P, Ferguson C, Lara AS, Leonard J, Younis M. Disparities in physician-patient communication by obesity status. *Inq J Med Care Organ Provis Financ.* 2014;51:0046958014557012. doi:10.1177/0046958014557012
42. Lindhardt CL, Rubak S, Mogensen O, Lamont RF, Joergensen JS. The experience of pregnant women with a body mass index >30 kg/m² of their encounters with healthcare professionals. *Acta Obstet Gynecol Scand.* 2013;92(9):1101-1107. doi:10.1111/aogs.12186
43. Drury CAA, Louis M. Exploring the association between body weight, stigma of obesity, and health care avoidance. *J Am Acad Nurse Pract.* 2002;14(12):554-561. doi:10.1111/j.1745-7599.2002.tb00089.x
44. Aldrich T, Hackley B. The impact of obesity on gynecologic cancer screening: an integrative literature review. *J Midwifery Womens Health.* 2010;55(4):344-356. doi:10.1016/j.jmwh.2009.10.001
45. Werder K, Curtis A, Reynolds S, Satterfield J. Addressing Bias and Stigma in the Language We Use With Persons With Opioid Use Disorder: A Narrative Review. *J Am Psychiatr Nurses Assoc.* 2022;28(1):9-22. doi:10.1177/10783903211050121
46. Healy M, Richard A, Kidia K. How to Reduce Stigma and Bias in Clinical Communication: a Narrative Review. *J Gen Intern Med.* 2022;37(10):2533-2540. doi:10.1007/s11606-022-07609-y
47. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program. Federal Register. Published May 1, 2020. Accessed June 14, 2022. <https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification>
48. Rodriguez JA, Clark CR, Bates DW. Digital Health Equity as a Necessity in the 21st Century

- Cures Act Era. *JAMA*. 2020;323(23):2381-2382. doi:10.1001/jama.2020.7858
49. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data*. 2016;3(1):160035. doi:10.1038/sdata.2016.35
50. Sun M, Oliwa T, Peek ME, Tung EL. Negative Patient Descriptors: Documenting Racial Bias In The Electronic Health Record. *Health Aff (Millwood)*. 2022;41(2):203-211. doi:10.1377/hlthaff.2021.01423
51. Zestcott CA, Spece L, McDermott D, Stone J. Health Care Providers' Negative Implicit Attitudes and Stereotypes of American Indians. *J Racial Ethn Health Disparities*. 2021;8(1):230-236. doi:10.1007/s40615-020-00776-w
52. SCCM | Meet the Critical Care Team. Society of Critical Care Medicine (SCCM). Accessed January 30, 2024. <https://sccm.org/MyICUCare/About-Critical-Care/Team>
53. Austin PC, Stryhn H, Leckie G, Merlo J. Measures of clustering and heterogeneity in multilevel Poisson regression analyses of rates/count data. *Stat Med*. 2018;37(4):572-589. doi:10.1002/sim.7532
54. Westerhausen D, Perkins AJ, Conley J, Khan BA, Farber M. Burden of Substance Abuse-Related Admissions to the Medical ICU. *Chest*. 2020;157(1):61-66. doi:10.1016/j.chest.2019.08.2180
55. Boag W, Suresh H, Celi LA, Szolovits P, Ghassemi M. Racial Disparities and Mistrust in End-of-Life Care. *ArXiv180803827 Stat*. Published online August 15, 2018. Accessed September 6, 2021. <http://arxiv.org/abs/1808.03827>
56. Adam H, Yang MY, Cato K, et al. Write It Like You See It: Detectable Differences in Clinical Notes by Race Lead to Differential Model Recommendations. In: *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*. AIES '22. Association for

- Computing Machinery; 2022:7-21. doi:10.1145/3514094.3534203
57. Brown CE, Engelberg RA, Sharma R, et al. Race/Ethnicity, Socioeconomic Status, and Healthcare Intensity at the End of Life. *J Palliat Med*. 2018;21(9):1308-1316. doi:10.1089/jpm.2018.0011
58. Monuteaux MC, Du M, Neuman MI. Evaluation of Insurance Type as a Proxy for Socioeconomic Status in the Pediatric Emergency Department: A Pilot Study. *Ann Emerg Med*. Published online January 18, 2024. doi:10.1016/j.annemergmed.2023.12.013
59. Jang BS, Chang JH. Socioeconomic status and survival outcomes in elderly cancer patients: A national health insurance service-elderly sample cohort study. *Cancer Med*. 2019;8(7):3604-3613. doi:10.1002/cam4.2231
60. Geter A, Herron AR, Sutton MY. HIV-Related Stigma by Healthcare Providers in the United States: A Systematic Review. *AIDS Patient Care STDs*. 2018;32(10):418-424. doi:10.1089/apc.2018.0114
61. Glassberg JA. Improving Emergency Department-Based Care of Sickle Cell Pain. *Hematol Am Soc Hematol Educ Program*. 2017;2017(1):412-417. doi:10.1182/asheducation-2017.1.412
62. Haywood C, Tanabe P, Naik R, Beach MC, Lanzkron S. The impact of race and disease on sickle cell patient wait times in the emergency department. *Am J Emerg Med*. 2013;31(4):651-656. doi:10.1016/j.ajem.2012.11.005
63. Evans CR. Modeling the intersectionality of processes in the social production of health inequalities. *Soc Sci Med*. 2019;226:249-253. doi:10.1016/j.socscimed.2019.01.017
64. Bauer GR, Churchill SM, Mahendran M, Walwyn C, Lizotte D, Villa-Rueda AA. Intersectionality in quantitative research: A systematic review of its emergence and

- applications of theory and methods. *SSM - Popul Health*. 2021;14:100798.
doi:10.1016/j.ssmph.2021.100798
65. Clark AE, Kashima Y. Stereotypes help people connect with others in the community: A situated functional analysis of the stereotype consistency bias in communication. *J Pers Soc Psychol*. 2007;93(6):1028-1039. doi:10.1037/0022-3514.93.6.1028
66. Klein O, Tindale S, Brauer M. The consensualization of stereotypes in small groups. In: *Stereotype Dynamics: Language-Based Approaches to the Formation, Maintenance, and Transformation of Stereotypes*. Lawrence Erlbaum Associates Publishers; 2008:263-292.
67. Steinkamp J, Kantrowitz JJ, Airan-Javia S. Prevalence and Sources of Duplicate Information in the Electronic Medical Record. *JAMA Netw Open*. 2022;5(9):e2233348.
doi:10.1001/jamanetworkopen.2022.33348
68. Bratanova B, Kashima Y. The “Saying Is Repeating” Effect: Dyadic Communication Can Generate Cultural Stereotypes. *J Soc Psychol*. 2014;154(2):155-174.
doi:10.1080/00224545.2013.874326
69. Butler R, Monsalve M, Thomas GW, et al. Estimating Time Physicians and Other Health Care Workers Spend with Patients in an Intensive Care Unit Using a Sensor Network. *Am J Med*. 2018;131(8):972.e9-972.e15. doi:10.1016/j.amjmed.2018.03.015
70. Edward K leigh, Stephenson J, Ousey K, Lui S, Warelow P, Giandinoto JA. A systematic review and meta-analysis of factors that relate to aggression perpetrated against nurses by patients/relatives or staff. *J Clin Nurs*. 2016;25(3-4):289-299. doi:10.1111/jocn.13019
71. Noyes AL. Navigating the Hierarchy: Communicating Power Relationships in Collaborative Health Care Groups. *Manag Commun Q*. 2022;36(1):62-91. doi:10.1177/08933189211025737
72. van Boekel LC, Brouwers EPM, van Weeghel J, Garretsen HFL. Stigma among health

professionals towards patients with substance use disorders and its consequences for healthcare delivery: Systematic review. *Drug Alcohol Depend.* 2013;131(1):23-35.

doi:10.1016/j.drugalcdep.2013.02.018

73. Corley MC, Goren S. The Dark Side of Nursing: Impact of Stigmatizing Responses on Patients. *Sch Inq Nurs Pract.* 1998;12(2):99-110,112-118.
74. Adams LY, Maykut CA. Bullying: The Antithesis of Caring Acknowledging the Dark Side of the Nursing Profession. *Int J Caring Sci.* 2015;8(3):765-773.
75. Hafferty FW, Franks R. The hidden curriculum, ethics teaching, and the structure of medical education. *Acad Med.* 1994;69(11):861-871.

Appendix 1: Patient Ethnicity and Provider Recategorizations

Provider recategorization

provider_label_final	Unique_Provider_Labels
Physicians	MDs, Res, MD, MSIV, Med St, RO, HMS MS, DML, MDS, Mds, MedRes, HMSIV, MSV, HMS IV, RF, MS V, MedSt, md
APP	PA, NP, SNP, NNP, SNNP, nnp
Pharmacist	PharmD, RPh, RPH
Registered Dieticians	RD, DI, RD,LDN, MS,RD, RD/LDN
Registered Nurses	RN, CoOpSt, RRT, SN, m, RNs, NSV, RNC, Rn, CoOPSt, StNurs, CoOpst, CCRN, RNBA, CoWker, DRM, SRN, Nurs, NS, StNRS, Nurse, CM, StNur, RNStu, NStude, CoWkr, StuNur, Co-Wkr, PracSt, CO-Op

Rehab (OT/PT)	Rehab, OTR/L, IMD, PT
Respiratory Therapist	SRT, RT, RRts, RTStu, CRT, RRTs, RTS, RRt, rrt
Social Workers	SW, LICSW, CRS, MSWint, SWInt, SW Int

Appendix 2: Lexicons for Doubt Markers and Stigmatizing Labels

Lexicon	Stem Word List	Expanded Words (Pruned to)	GPT-3.5 added words	High-noise terms removed	Final Lexicon	Final Lexicon Length
Doubt Markers	"adamant", "claimed", "insists", "allegedly", "disbelieves", "dubious"	60 total, reduced to 42. Agreement = 80%	['skeptical', 'dubiousness', 'questionable', 'doubting', 'uncertain', 'skepticalness', 'incredulous', 'hesitating', 'suspicious', 'mistrustful', 'distrustful', 'unconvinced', 'unsure', 'hesitant', 'wary', 'dubious', 'disbelieving', 'skepticalism', 'hesitancy', 'skepticism', 'mistrust', 'uncertainness', 'disbelief', 'suspicion', 'mistrustfulness', 'incredulity', 'incredulously', 'wavering', 'ambivalent', 'waveringly', 'questionableness', 'mistrustingly', 'doubter', 'questioning', 'doubtingly', 'mistrusting', 'doubtful', 'skeptical', 'unconvincedly', 'mistrustingly', 'mistrustfully', 'doubtingness', 'skepticism', 'questioningness',	'suspicion', 'suspicious', 'questionable', 'questioning', 'uncertain', 'hesitancy', 'hesitant', 'unsure'	["doubtful", "dubious", 'insists', 'accused', 'adamant', 'adamant/belligerent', 'adamantly', 'addamant', 'allegedly', 'alleged', 'allegedly', 'allegedly-unnecessary', 'asserted', 'believes', 'claimed', 'claimedthat', 'claires', 'claiming', 'confessionally', 'culpably', 'disbelief', 'disbelieve', 'disbelieved', 'disbeliever', 'disbelievers', 'disbelieves', 'disbelieving', 'disclaimed', 'doggedly', 'doubious', 'doubtful', 'dubious', 'dubious/equivocal', 'dubiously', 'insist', 'insisted', 'insist', 'insisted', 'insisting', 'insists', 'misbelieve', 'misbelieved', 'misbelieves', 'mistrustful', 'mistrusting', 'non-dubious', 'proclaimed', 'purportedly', 'reinsists', 'skeptical', 'speculative', 'supposedly', 'them-insists', 'unconvinced', 'undisguisedly', 'unreliable', 'unsure', 'wavering']	58

			'unbelieving', 'unsureness', 'skepticalness', 'questioningness', 'doubtingly', 'unbelievably', 'skeptically', 'mistrustingly', 'mistrustfully', 'skeptically', 'questioningly', 'doubtingly', 'skeptically', 'mistrustingly', 'mistrustfully']			
Stigmatizing Labels	"abuser", "junkie", "alcoholic", "drunk", "drug-seeking", "nonadherent", "agitated", "angry", "combative", "noncompliant", "confront", "noncooperative", "defensive", "hysterical", "unpleasant", "refuse", "frequent-flyer", "reluctant"	180 , reduced to 83. Annotator agreement = .75.	[' "hysterical', 'aggressive', ' drug addict', ' non-compliant', ' lazy', ' attention-seeking', ' manipulative', ' hypochondriac', ' difficult', ' mentally unstable', ' troublemaker', ' irresponsible', ' unpredictable', ' irrational', ' needy', ' demanding', ' disruptive', ' uncooperative', ' unreliable', ' high maintenance', ' attention-seeker', ' dramatic', ' attention-seeking', ' lazy', ' invalid', ' faker', ' irrational', ' hostile', ' aggressive', ' challenging', ' uncooperative', ' deceptive', ' demanding', ' unreliable', ' high-strung', ' self-destructive', ' unstable', ' manipulative', ' entitled', ' attention-seeking', ' violent', ' drug seeker', ' malingerer', ' faker', ' mentally ill', ' dangerous', ' delusional', ' needy', ' overly sensitive', ' unstable', ' irrational']	'difficult', 'suspicious', 'aggressive', 'unstable', 'dramatic', 'unreliable', 'entitled', 'invalid', 'dangerous'	["'hysterical', "'drug-seeking", "'hysterical", "'drug-seeking", "'junkie", 'reluctant', 'abuse/abuser', 'abused-abuser', 'abuser', 'abuser's', 'abusers', 'addictive-drug-seeking', 'alcoholic', 'angry', 'angry-disgusted', 'angry/disgusted', 'attention-seeker', 'attention-seeking', 'challenging', 'combative', 'combatively', 'compliant/noncompliant', 'counterdefensive', 'deceptive', 'defensive', 'defensive/offensive', 'delusional', 'demanding', 'disruptive', 'drug addict', 'drug seeker', 'drug-craving/drug-seeking', 'drug-seeking', 'drug-seeking/-taking', 'drug-seeking/drug-taking', 'drug-seeking/taking', 'drug-seeking/use', 'drunk', 'drunken', 'drunkenly', 'drunker', 'drunkest', 'drunks', 'ex-abuser', 'ex-alcoholic', 'faker', 'frequent-flier', 'frequent-flyer', 'frequent-flyers', 'frequent-fvl', 'frequent-hitter', 'frequent-hitters', 'high maintenance', 'high-strung', 'histrionic-hysterical', 'hostile', 'hypochondriac', 'hypochondriac-hysterical', 'hysteric', 'hysterical', 'hysterical-obsessive', 'hysterical/anaclitic',	127

					<p>'hystericals', 'hysteries', 'incompliant', 'irrational', 'irrational', 'irresponsible', 'iv-abuser', 'ivdabuser', 'junkie', "junkie's", 'junkies', 'lazy', 'ma-abuser', 'malingerer', 'manipulative', 'mentally ill', 'mentally unstable', 'morereluctant', 'needy', 'non-adherent', 'non-alcoholic/alcoholic', 'non-compliant', 'non-cooperating', 'non-cooperation', 'non-cooperative', 'non-cooperatively', 'nonadhered', 'nonadherent', 'nonadherently', 'nonadherents', 'noncompliant', 'noncompliant/complian t', 'noncompliant\\medicall y', 'noncompliance', 'noncooperating', 'noncooperation', 'noncooperative', 'noncooperatively', 'novelty/drug-seeking', 'ny-nonadherent', 'onadherent', 'overdefensive', 'overly sensitive', 'prealcoholic', 'pt.noncompliant', 'refuse', 'refuses', 'refusing', 'reluctantly', 'reluctant', 'reluctantly', 'reluctants', 'schizo-hysterical', 'self-destructive', 'troublemaker', 'un-adherent', 'unadherent', 'uncooperative', 'unpleasant', 'unpleasant/annoying', 'unpleasantly', 'unpleasantries', 'unpredictable', 'unwilling', 'unwillingly']</p>	
--	--	--	--	--	--	--

Aim 3: Evaluating the relationships between stigmatizing language features in the EHR and patient care outcomes of opioid analgesic prescription rates and self-directed discharge

Abstract:

Objective: To assess the relationships between counts of stigmatizing labels and doubt marking linguistic features within the electronic health record (EHR) with patient care outcomes of opioid analgesic prescription and dispensation rates, and rates of self-directed discharge from an intensive care unit (ICU).

Materials and Methods: Utilizing supervised learning models developed from prior work, we applied classifiers at the sentence level among patients in the MIMIC-III dataset, and constructed outcomes and covariates from relevant medical events and prescription tables. Poisson and logistic regression models were used to assess the relationships between linguistic features and care outcomes. **Results:** Each additional stigmatizing label per chart was associated with a 3.5 times higher odds of self-directed discharge (95% CI: 1.49, 8.17, $p = .004$). Similarly, each additional doubt marker per chart was associated with a 7.1 times higher odds of self-directed discharge (95% CI: 2.08, 24.29, $p = .0017$). No significant associations were found between stigmatizing labels or doubt markers and opioid prescription or dispensation rates. **Discussion:** Findings indicate significant associations between stigmatizing and doubt marking language in patient EHR with key reduced quality of care indicators within an ICU setting. While interpretation of causal effects is limited, this study provides further evidence of the potential to explore connections between implicit biases and care outcomes using natural language processing methods in real-world ICU datasets. **Conclusions:** Findings should provide further support to the ongoing evaluation of the multi-modal ways in which provider-held stigmas can impact patient care. The clear connection of linguistic bias features in EHR notes to salient care outcomes should prompt further efforts to develop meaningful interventions to reduce the transmission of patient stigmas through EHR notes.

Introduction

Provider-held stigmas are widely believed to manifest in discrimination and health inequities among patients.^{1,2} Patients regularly experience stigmatization by providers as a result of their race, gender, sexual orientation, disease status, drug use, and socioeconomic status, or other labeled characteristics identified by provider teams.³ Patients with multiple stigmatized identities and characteristics experience greater stigmatization and face unique challenges.⁴ Leaders across provider specialties and disciplines have long called for investigation of the processes by which provider stigmatization forms and is cultivated among medical teams.⁵⁻⁷ While recent efforts to address provider stigma are promising, due to the multifaceted, dynamic, and persistent impacts on health, it is of urgent priority for public and clinical health experts to develop novel methods for measuring and intervening on stigma. Drawing from ecosocial theory, relevant sociolinguistic frameworks, and advances in natural language processing, there is great potential for researchers to investigate manifestations of stigma within the data encoded in electronic health record (EHR) notes. Being able to assess instances of patient stigmatization in the health record could allow healthcare teams the ability to better evaluate and intervene on this wicked problem that drives inequities in health outcomes across marginalized groups.

This paper aims to examine the links between stigmatizing and doubt-marking language in the electronic health record and key quality of care outcomes of opioid prescription rates and patient self-directed discharge. The remaining subsections will focus on defining each of these constructs and outlining the framework we employ to examine these connections between provider stigmatizing language and health outcomes.

Defining Stigma

Stigma has been defined by social psychologists Link and Phelan as a social process that is characterized by the interplay of labeling, stereotyping and separation, which leads to status loss and discrimination, and importantly, occurs within a context of power, such as that of the patient-provider

relationship.⁸ Labeling occurs whenever we identify individuals as belonging to a particular group. Stereotyping involves ascribing a proclivity towards a specific behavior or characteristic to members of a labeled group. Separation involves distancing from the group, and drawing lines of “us versus them”. Status loss and discrimination involves a negative evaluation of a groups’ attributes and its members relative to another group. As Link and Phelan posit, all of these processes crucially develop in the context of power and power imbalances, which are extremely relevant in the context of provider-patient relationships.^{9,10} Beukeboom and Burgers recently compiled decades of linguistic research on stereotyping and stigmatization to form the Social Categories and Stereotypes Communication Framework.¹¹ This framework posits that systematic differences in language, manifested by 1) how we label others and 2) how others’ behaviors and characteristics are described, contribute towards shared consensual understandings of perceived stereotypes and group category essentialism for individuals belonging to different groups. Within the context of provider notes, we focus our efforts on understanding stigmatizing labels, which help identify systematic differences in how patients are described, and doubt markers, which help identify the confidence and trust providers have in how patients present themselves in their testimonies, medical complaints, and behaviors.

Stigmatizing labels

Stigmatizing labels to describe groups are often used to perpetuate stereotypes, and when used by providers, can lead to feelings of stigmatization and reduced trust among their patients. Much of the recent work on identifying and reducing stigmatizing labels has come from providers seeking to improve care for patients with substance use disorders. A recent NIDA study published a list of words to avoid using around patients with substance use disorders, including "addict", "abuser", "user", or "junkie", which have been found to be associated with perceived stigmatization by patients.¹² Similar studies have been applied to other chronic illness populations, identifying terms like “sickler” or “frequent flier” which may be used to further stigmatize patients with chronic illnesses who are often admitted into the hospital.

^{13,14} Recent research led by Michael Sun and colleagues on over 40,000 clinical notes has found

systematic racial inequities in presence of “Negative Descriptor” words in medical charts, defined by the Health Equity Commission of the Society of General Internal Medicine, which included commonly used terms in the EHR such as “(non-)adherent, aggressive, agitated, angry, challenging, combative, (non-)compliant, confront, (non-)cooperative, defensive, exaggerate, hysterical, (un-)pleasant, refuse, and resist”. This study found that compared to White patients, Black patients had 2.54 times the odds of having at least one negative descriptor written in their history and physical notes ¹⁵.

While some providers may argue that these terms may be useful in flagging unwanted patient behaviors or mental states, a recent study has shown that patients exposed to language written about them by providers which included stigmatizing labels resulted in patients feeling unfairly judged, labeled, and disrespected.¹⁶

Doubt Markers

Linguistic features such as evidentials, defined as “the linguistic coding of epistemology”,¹⁷ are words that are frequently used in chart language to question the veracity of patients. Among the many words used as evidentials, words and expressions used to confer doubt or uncertainty such as: *allegedly*, *apparently*, or verbs like *claimed*, are often used to when describing patient testimonies, for example: “patient *claimed* their pain was 10/10”.¹⁸ Providers may use words when describing patient testimony in combination with stigmatizing labels or negative descriptors of patients to transmit their stance, or expression of attitudes, feelings, and judgment about patients to other providers which may impact future treatment and care decisions.¹⁹ Inequities have been found among usage of these terms across race and gender, where patients who were women and patients who were Black were found to have significantly higher frequencies of evidentials in their provider notes than patients who were men or White.¹⁸

Theoretical frameworks for how provider stigmatizing language impacts health outcomes

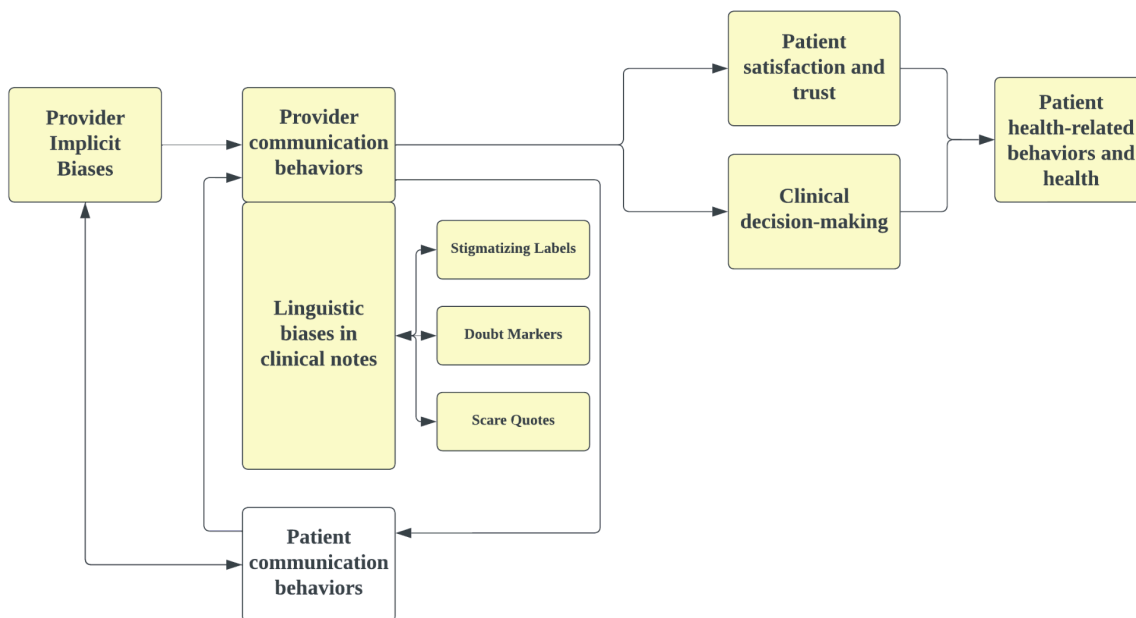
Our understanding of how stigma impacts health outcomes in clinical settings is guided by two major theoretical frameworks. Beukeboom and Burgers recently compiled decades of research on

stigmatization and stereotyping in language to create the Social Categories and Stereotype Communication Framework, which is highly useful in the present study.¹¹ This framework describes how stereotypes are forged and perpetuated in shared cognition through systematic differences in how language is used to 1) label others, and 2) describe the behaviors and characteristics of others. Our second framework, developed by Hagiwara and colleagues, is used to understand how stigmas held by providers, in the forms of implicit biases, fuel differences in physician communication behaviors, patient communication behaviors, which bidirectionally interact to impact patient satisfaction, trust, and eventually patient health and health behaviors. Hagiwara's model helps us to focus on the role of provider stigmatizing language in charts within the interactional level of the patient-provider relationship. This model also posits that provider communication behaviors are part of a reciprocal feedback loop within patient-provider relationships. In this loop, stigmas towards various patients and patient groups held by providers are hypothesized to produce negative provider communication behaviors in patient interactions, which are then met with reduced quality of patient communication. These processes work in tandem to impact patient trust and satisfaction, as well as clinical decision making in the context of the patient's healthcare team. While this conceptual model does not seek to completely eliminate the impacts of higher-level constructs of systemic racism, it provides a useful framework for investigating the role of provider stigmatizing language within the patient-provider relationship, and for planning interventions to improve health outcomes for frequently stigmatized patient populations.

Because our study aims to investigate the link between stigmatizing language embedded in provider communication patterns and patient care outcomes, we position the linguistic framework offered by Beukeboom and Burgers within Hagiwara and colleagues' conceptual model situated in the context of "provider communication behaviors". Described by Figure 1 below, our model adaptation has operationalized the construct of *provider communication behaviors* as linguistic features in clinical notes, structured by the Social Categories and Stereotype Communication Framework. We posit that linguistic features of stigmatizing labels and doubt markers in patients' electronic health records work to create

shared stereotypes and stigmas against patients, which in turn impact provider and patient communication factors, patient satisfaction, trust, and clinical decision-making.

Figure 1: Conceptual model for how stigmatizing language in charts impacts health outcomes



Clinical Outcomes of Associated with Stigma

Within healthcare provider teams, stigma and provider explicit and implicit biases hold strong influence over patient care decisions,²¹ and erode patient perceived trust and quality of provider-patient interactions.^{22,23} Previous research has identified several key outcomes theorized to be linked to provider stigma and implicit biases, including insufficient pain management strategies and increased instances of leaving against medical advice.

Pain Management

Many studies have pointed to disparities in pain management strategies among frequently stigmatized groups. Previous research has amassed substantial evidence of widespread reduced opioid prescription rates among Black and Hispanic patients^{24–26} and women,^{27–30} occurring in nearly all age groups, from pediatrics³¹ to end-of-life care.^{32,33} Due to the inherent subjectivity of patient-reported pain scores, pain management situations are hypothesized to embody the “high clinical ambiguity” that has been posited to be more susceptible to the impact of implicit biases and stigmas which are shared across medical teams.³⁴ Often, patients who experience stigmatization while in need of pain management are stereotyped as “drug-seeking”, resulting in reduced prescription and dispensation of opioids, and failure to believe patient reported pain levels.^{30,34,35} This undertreatment of pain can have disastrous downstream clinical outcomes for all patients, but can be especially severe for patients with chronic illnesses such as sickle cell disease, where uncontrolled pain can exacerbate organ damage incurred through vaso-occlusive crises.^{36,37}

Self-directed discharge

Self-directed discharge can come at serious cost to patient health and healthcare systems. Research on self-directed discharge have largely reached consensus that patients who complete self-directed discharge have significantly higher odds of 30-day readmission and mortality,^{55–57} Patients who are readmitted following self-directed discharge on average also undergo double the length of stay on the following admission, and incur over 50% higher overall healthcare costs.⁵⁸ Further, self-directed discharge is highest in patients who have low income, and among frequently stigmatized populations such as people with histories of substance use, HIV infection, and sickle cell disease.⁵⁸ Experiences of discrimination and bias by race and socioeconomic

status have also been shown to drive higher rates of self-directed discharge^{38,39}, commonly known as “leaving against medical advice”, and increased reluctance to seek care^{22,40-42}, which can lead to increased risk of mortality and disease complications across a variety of chronic conditions.^{37,43} We use the term self-directed discharge to center patient experiences and reduce the paternalistic or stigmatizing framing of “leaving against medical advice”.⁴⁴ While self-directed discharge happens because of a variety of socioeconomic and individual patient factors, providers factors such as hospital setting and structure and providers’ clinical communication style and experience, along with patient perceptions of trust likely play a significant role.⁴⁵⁻⁴⁷ While there has been limited research on provider-level predictors of self-directed discharge, current studies cite “failure to orient the patient to treatment on intake, punitive or threatening atmosphere on the inpatient unit, difficulties in doctor-patient relationship, [and] failure to establish a supportive provider-patient relationship” as common issues that may be directly related to provider-held stigmas towards marginalized patient populations.⁴⁸

This study seeks to further understand the connections between stigmatizing and doubt marking provider language and care outcomes by assessing the relationships between stigmatizing labels and doubt markers identified in the electronic health record with opioid pain management outcomes and whether patients complete self-directed discharge.

Methods

MIMIC-III Dataset

The Medical Information Mart for Intensive Care, or “MIMIC-III”, is a freely-available database of comprehensive, de-identified EHR, free-text notes, as well as prescriptions and chart event documentation for over 40,000 patients admitted to the ICU at Beth Israel Deaconess

Medical Center in Boston, MA from 2001 to 2012.⁴⁹ This dataset contains over 1.2 million clinical provider notes, across nearly 50,000 admissions. The accessibility of free-text EHR dataset, combined with its extensive range of care quality outcomes and derived clinical severity scores, makes it an ideal source to test associations between linguistic features and care outcomes at the focus of the current study. Prior to tokenizing the charts at the sentence level and merging with lexicons of stigmatizing labels and doubt markers, all duplicate sentences were removed from the dataset, and charts labeled as EEG or Radiology were removed in order to restrict to charts more likely to have subjective narrative and patient history text data. Rates of stigmatizing labels and doubt markers, derived from prior work in which charts were classified at the sentence level of containing each of these features, were summarized at the chart and patient level. Further detail on how analysis predictors, outcomes, and covariates were developed is provided below.

PREDICTORS

Predictors included in models will include patient-level aggregates of two types of stigmatizing linguistic features: stigmatizing labels, and doubt markers. These linguistic features have been identified within the MIMIC-III dataset via clinical and health scientist annotations, which trained supervised learning text classifier models.

Stigmatizing labels

Stigmatizing label lexicon development was guided by literature on stigmatizing language in medical care, specifically from the NIDA “Words Matter” publication, Sun’s

“Negative Patient Descriptors: Documenting Racial Bias in the Electronic Health Record”, as well as Zestcott’s “Health Care Providers’ Negative Implicit Attitudes and Stereotypes of American Indians”.^{12,15,50} The initial stem word list consisted of 18 words: "abuser", "junkie", "alcoholic", "drunk", "drug-seeking", "nonadherent", "agitated", "angry", "combative", "noncompliant", "confront", "noncooperative", "defensive", "hysterical", "unpleasant", "refuse", "frequent-flyer", "reluctant". This list was subsequently expanded using contextual word embeddings models and GPT 3.5 models, and pruned by expert clinical and public health annotators. Supervised learning natural language processing classifiers were developed from an annotated sample derived from MIMIC-III charts with matching sentences containing lexicon terms, resulting in a logistic regression classifier model with 81% accuracy, 75% positive class precision, which indicates the percentage of correct positive predictions out of all positive predictions made, 84% positive class recall, which indicates the percentage of the total positive samples which were correctly predicted, and .79 macro-F1 score, which combines the recall and precision, as a harmonic mean of the two, to provide a balanced estimate of to describe how well the model “trades-off” the two measures which are often at odds. These findings were comparable to human annotator agreement, which was calculated at 87% (kappa = .74).

Doubt markers

Doubt marker lexicon development was guided by literature on use of “doubt markers” in medical care, specifically led by Beach and colleagues, which identified words such as “claims”, “insists”, and “adamant” or “apparently”, which have been found to be used to discredit or invalidate patient testimony. The 6 words included on the initial stem list were: "adamant", "claimed", "insists", "allegedly", "disbelieves", "dubious".

Appendix 2 provides a final lexicon list used to develop annotation samples used to refine the supervised learning models, for both stigmatizing labels and doubt markers. Supervised learning natural language processing classifiers were developed from an annotated sample derived from MIMIC-III charts with matching sentences containing lexicon terms, resulting in a RoBERTa classifier model with 86% accuracy, 86% positive class precision, 71% positive class recall, and .84 macro-F1 score. These findings were comparable to human annotator agreement, which was calculated at 87% ($\kappa = .73$).

These models were used to classify all sentences containing relevant stigmatizing labels and doubt marker lexicons in prior work, available via Physionet and [Github](#). Average counts of the number of identified sentences containing stigmatizing labels and doubt markers per patient chart were calculated and summarized at the patient level.

OUTCOMES

Opioids/Pain Management

Opioid prescription rates at the patient level were calculated as average daily Morphine Milligram Equivalent (MME), which was derived using guidance from the Utah Medicaid Office and CDC Guidelines.⁵¹ These guidelines provided an opioid/MME dictionary, which was used to match any entries within the Prescriptions Table in MIMIC-III. For data in which ranges were present (i.e. 10-15 mg), we calculated the average of the range for that specific entry. The MME values calculated for each row in the Prescriptions table of MIMIC-III were averaged across the duration of days prescribed opioids, indicated by startdate and enddate variables. These values represented the average daily opioid prescription and dispensation patient level by using relevant

dose prescribed and dispensed variables. For patients which received no opioids, their values were transformed to 0.

Self-directed discharge

Instances wherein patients left against medical advice were collected from discharge location information provided in the MIMIC-III Admissions table, as leaving “AMA”, or against medical advice, which we refer to self-directed discharge to limit stigmatization. For each patient, any patients who had records of self-directed discharge within the Admissions table were assigned a binary “1” indicator for having completed self-directed discharge. All other patients were assigned a “0”.

COVARIATES

Covariates for all adjusted models included demographic variables such as gender, whether patient was identified as Black/African American, insurance type (government-run, self-pay, or private), and whether patients were diagnosed with four commonly stigmatized conditions: obesity, opioid use disorder, symptomatic HIV, and sickle cell disease, drawing from matches with relevant ICD-9 codes. We also included length of stay in total days, as well as patients’ average Oxford Acute Severity Illness Scores over all admissions (OASIS)⁵², which is a scale derived from physiological indicators such as: heart rate, mean arterial pressure, temperature, respiratory rate, urine output, pre-ICU admission length of stay, Glasgow Coma Scale, age, being placed on a mechanical ventilator at any point during day 1, and admission following elective surgery.⁵² This scale has been applied in other recent work utilizing data from the MIMIC-III dataset to assess disparities in end-of-life care outcomes, and has shown high

predictive ability on a variety of clinical outcomes.^{33,53,54} Finally, we included the covariate of a patient's average daily pain score, derived from timestamped records of patient pain reports listed in MIMIC-III's ChartEvents table. These items were calculated from the two different pain level indicators, collected through both the Philips CareVue system, used from 2001-2008 (itemid: 1044), and the iMDSoft Metavision system, used after 2008 (itemid: 223791). These scores were aggregated at the day, and then patient levels to get average patient-reported pain scores per day, per patient.

ANALYSES

Prior to model building, we conducted descriptive analyses on the distributions of outcomes, and covariates in order to determine viability for different generalized linear modeling structures. Next, we ran unadjusted models assessing the relationships between stigmatizing labels and doubt markers with patient daily average opioid prescription and dispensation rates, as well as whether the patient completed self-directed discharge.

In order to model the effect of provider stigmatizing language on patient clinical outcomes and patient health behaviors, we conducted a series of multivariable generalized linear models which assess the relationships between patient-level predictors of rates of stigmatizing labels and doubt markers in EHR notes on the patient-level outcomes of average daily opioid prescription and dispensation rates, and whether or not each patient completed self-directed discharge. We assessed associations using unadjusted (predictor and outcome only) generalized linear models, as well as adjusted models, which included relevant demographic, diagnosis, and disease severity covariates. Of note, the majority of predictors and outcomes did not follow normal distribution, which is typical for many forms of count-based data. For this reason, we

implemented Poisson distribution generalized linear models when modeling the rate of MME prescription and dispensation rates per day, and binomial distributions logistic regression models for the binary outcome of whether or not the patient completed self-directed discharge. Due to high rates of missing values for average daily pain scores (30%, or 3,489 patients did not have values for this variable, likely due to unresponsiveness), we performed a series of sensitivity analyses on patient samples with three different methods: 1) restricted to only patients with average daily pain scores, including the average daily pain score covariate (the model results shown in Table 3) , 2) restricted to only patients with average daily pain scores without the average daily pain score covariate, as well 3) the full 11,630 patient sample, without the average daily pain score covariate. These models all resulted in similar patterns of association. These results are available in Appendix A.

Results

Table 1 includes descriptive results of the study predictors, covariates, and outcomes. Overall, the MIMIC-III dataset is composed of primarily patients who were identified as “White”, (71.5%), with the next largest racial/ethnic group being “Black/African American”. Most patients either had Medicare insurance (46.9%), or private insurance (38.8%). This sample had slightly more male patients (56.3%), and of the stigmatized diagnoses, we had the highest prevalence of patients with obesity (5.8%) and opioid use disorder (1.9%). Of all patients, 20.6% had at least one stigmatizing label across all charts, and 6.6% of patients had at least one doubt marker across all charts.

Table 1: Descriptive results of study predictors, demographic covariates, and outcomes

	Overall (N=11630)
Race/Ethnicity*	
White	8312 (71.5%)
Asian	353 (3.0%)
Black/African American	945 (8.1%)
Hispanic/Latino	440 (3.8%)
Native American/Alaskan Native	16 (0.1%)
Other	345 (3.0%)
Unknown/Declined	1219 (10.5%)
Insurance	
Private	4509 (38.8%)
Government	353 (3.0%)
Medicaid	1211 (10.4%)
Medicare	5450 (46.9%)
Self Pay	107 (0.9%)
Gender	
Female	5086 (43.7%)
Male	6544 (56.3%)
Diagnoses	
Sickle Cell Disease	25 (0.2%)
Opioid Use Disorder	217 (1.9%)
HIV (Symptomatic)	115 (1.0%)
Obesity	674 (5.8%)
Stigmatizing Labels Count (per Patient)	
Mean (SD)	.5 (1.99)
Median [Min, Max]	0 [0, 90]
Stigmatizing Labels Count (per Patient, per Chart)	
	.01 (.08)
	0 [0, 4]
Patients with at least one Stigmatizing Label	2395 (20.6%)
Doubt Markers Count (per Patient)	
	0.09 (0.48)
	0 [0, 29]
Doubt Marker Count (per Patient, per Chart)	
	.004 (.05)
	0 [0, 2]
Patients with at least one Doubt Marker	770 (6.6%)
Average Pain Score per day, per patient	
	2.81 (2.14)
	2.87 [0, 10]
Average OASIS severity index score, per patient	
	32.35 (8.24)
	32.5 [6, 68]
Per-patient count of times left self-directed discharge	
	0.01 (0.21)
	0 [0, 19]
Average Daily Milligrams Morphine Equivalent Prescribed	
	0.86 (6.29)
	0 [0, 255]

Average Daily Milligrams Morphine Equivalent Dispensed 0.16 (0.64)
0 [0, 16.33]

*Race/ethnicity simplified to Black/African American and Not Black/African American, due to low cell-count for multiple values.

Results of the unadjusted and adjusted (including all covariates) Poisson and regression models are described in Table 3. Overall, no significant associations were found between either stigmatizing labels per chart, doubt markers per chart, and the outcomes of average daily MME prescribed or dispensed. However, we found that each additional stigmatizing label per chart was associated with a 3.5 times higher odds of self-directed discharge (OR=3.49 95% CI: 1.49, 8.17, $p = .004$). Similarly, each additional doubt marker per chart was associated with a 7.1 times higher odds of self-directed discharge (OR=7.10 95% CI: 2.08, 24.29, $p = .0017$).

Table 2. Regression results showing relationships between stigmatizing labels, doubt marking linguistic EHR features per chart with opioid medication rates and whether patients completed self-directed discharge associations. Betas or Odds Ratios (95%CI)

	Avg Daily MME Prescribed (Unadjusted)	Avg Daily MME Prescribed (Adjusted)	Avg Daily MME Dispensed (Unadjusted)	Avg Daily MME Dispensed (Adjusted)	Patient self-directed discharge (Unadjusted)	Patient self-directed discharge (Adjusted)
Stigmatizing Label Count, Per Chart	1.67 (.30, 9.03)	.102 (.20, 5.16)	.103 (.87, 1.22)	1.00 (.84, 1.17)	3.10 (1.40, 6.55)* p = .0033	3.49 (1.49, 8.17)* p = .004
Doubt Marker Label Count, per chart	-0.01 (-2.85, 2.83)	-0.00 (-3.70, 1.71)	-.11 (-.39, .18)	-.17 (-.44, .11)	8.08 (2.59, 25.03)* p = .0003	7.10 (2.08, 24.29)* p = .0017

* p is significant at $<.05$ value

** p is significant at $<.0001$ value

Adjusted models include covariates of: Gender, total length of stay, average daily pain score, Black/African American identified race, insurance (Government-run, private, self-pay), average OASIS severity index score, and whether patient was diagnosed with obesity, opioid use disorder, symptomatic HIV, or sickle cell disease.

Discussion

Findings indicate significant associations between stigmatizing and doubt marking language in patient EHR with key reduced quality of care indicators within an ICU setting,

particularly with the outcome of patient self-directed discharge. These linguistic features were strongly associated with self-directed discharge-- with each additional average stigmatizing label or doubt marker per chart predicting 3.5 and 7.1 times the odds of self-directed discharge, even when adjusting for key disease severity and demographic covariates. These results are especially compelling in light of our previous study on distributions of stigmatizing labels and doubt markers across patient demographic factors from the same sample as the current study. We found higher rates of stigmatizing labels per chart among Black/African American patients compared to White patients, among patients with government-run (Medicare, Medicaid) insurance and self-pay insurance compared to patients with private insurance, as well as higher rates of stigmatizing labels among patients with symptomatic HIV, obesity, and opioid use disorder diagnoses. Additionally, our prior study found higher rates of doubt markers per chart among patients who were male compared to female, among patients with government-run and self-pay insurance compared with private insurance, as well as among patients with symptomatic HIV, obesity, and opioid use disorder diagnosis. Findings from the current and previous studies point to increased marginalization across historically stigmatized patients through both the inequitable distribution of exposures of stigmatizing and doubt-marking language, and their significant associations with severe downstream health and quality of care outcomes such as self-directed discharge. These compelling results align with our interpretation of the conceptual model of implicit biases in healthcare posited by Hagiwara and colleagues, suggesting that language in charts are likely to be an important means of provider/patient communication that is connected with care decisions and quality of care outcomes.²⁰ Taken together, these findings may help explain factors which work to create and enforce stigmas and drive significant inequities in care outcomes for historically marginalized patients.

Other studies have tied trends of self-directed discharge with inadequacy of pain management, particularly among patient with opioid-related diagnoses.⁵⁹ In contrast to other studies, which have indicated inequities in opioid pain management outcomes connected with provider implicit bias or EHR stigmatizing language,^{13,24} we found no associations between stigmatizing labels or doubt markers and opioid prescription or dispensation rates. While we sought to limit our sample to patients who were able to report pain scores, per-patient daily averages may obscure precise clinical windows in which stigmatizing language or doubt marking language may most impact care. In order to identify more precise clinical windows, it may be necessary to limit further studies assessing connections between pain management and provider outcomes to specific diagnoses or illness criteria. While stigmatizing labels and doubt markers are important linguistic markers of patient stigmatization, it is possible other forms exist, such as “scare quotes”, which we had attempted to construct but did not achieve satisfactory classifier performance.

While this study provides a rich data source to assess the associations among linguistic features of stigmatizing labels and doubt markers with a variety of health outcomes, among a diverse range of conditions, it is important to consider limitations of the dataset’s context of the ICU, where patients may be less likely or able to complete self-directed discharge, or where opioid prescription and dispensation rates may also be driven by need to sedate patients in tandem with the goal of appropriate pain management. Additionally, our study setting, taking place at a single-center in Boston, MA, is composed of patients who are majority White, which likely makes assessing differences in outcomes between racial and ethnic groups more difficult. It is worth also considering how the timeframe of this dataset, taken from notes spanning 2001 to 2012, may differ from patterns in trends in current EHR language. Current EHR language may

be different due to a multitude of factors, such as changing standards of documentation, increased EHR transparency, or new advances in clinical charting systems and templates. While our findings highlight associations between stigmatizing and doubt-marking language and quality of care outcomes, we cannot directly assume causality, due to the observational nature of the study, as well as the lack of a longitudinal, repeated-measures framework in this current study.

Ultimately, this study provides strong evidence of associations between linguistic EHR note features related to implicit bias and quality of care outcomes. These findings, and the methodological tools developed during the course of three studies, can further support efforts to identify and assess different markers of bias, stereotypes, mistrust, and stigmatizing language within a variety of care settings and populations. The methodological framework of our current study can be replicated focusing on different clinical groups, as well as clinical settings such as emergency departments, which involve faster decision-making from providers that, in theory, may be more susceptible to implicit bias and stereotypes.⁶⁰ One example for further study could be sickle cell disease pain crises, where clinical guidelines of opioid pain treatment are more standardized and actual care decisions may be more varied due to stigma or lack of provider knowledge.³⁶ A major contribution of our study's findings and methodological framework is its utility in creating interventions to reduce stigmatizing language in the EHR in real-time. With appropriate validation and training, these methods could be used to identify real-time instances of provider stigmatization, which can be used to target interventions toward specific teams and mitigate harm within vulnerable patient groups. Additionally, the identification of stigmatizing and doubt marking language could be used within an automated electronic health record pipeline, incorporating new advances in Large Language Models (LLMs) to real-time "de-stigmatizing"

corrections or suggestions for providing more neutral or patient-centered chart language. Given the recent changes in increased patient medical chart access and transparency championed by the 21st Century Cures Act, it is imperative to evaluate the quality of provider medical chart language to reduce patient stigmatization.^{61,62} In light of recent efforts to expand transparency of medical records, hospital systems may directly benefit from these types of analyses and interventions, which could provide real-time audits and mitigations of stigmatizing language and doubt markers within patient charts. Given the complexities of solving the problem of implicit provider bias, this provides another avenue to assess and intervene to help improve inequitable patterns of patient care.

Conclusions

Findings provide further support to the ongoing evaluation of the multi-modal ways in which provider-held stigmas can impact patient care. The clear connection of linguistic bias features in EHR notes to salient care outcomes should prompt further efforts to develop meaningful interventions to reduce the transmission of patient stigmas through EHR notes. Many current interventions and evaluations of implicit provider bias are limited by scale and point-in-time nature. Systems like the ones developed in this and previous connected studies could provide tools for ongoing audits and intervention targeting for hospital systems in a way that can push forward the landscape of implicit bias interventions and evaluations.

References

1. Maina IW, Belton TD, Ginzberg S, Singh A, Johnson TJ. A decade of studying implicit racial/ethnic bias in healthcare providers using the implicit association test. *Soc Sci Med*. 2018;199:219-229. doi:10.1016/j.socscimed.2017.05.009
2. Hatzenbuehler ML, Phelan JC, Link BG. Stigma as a Fundamental Cause of Population Health Inequalities. *Am J Public Health*. 2013;103(5):813-821. doi:10.2105/AJPH.2012.301069
3. FitzGerald C, Hurst S. Implicit bias in healthcare professionals: a systematic review. *BMC Med Ethics*. 2017;18(1):19. doi:10.1186/s12910-017-0179-8
4. Biello KB, Hughto JMW. Measuring Intersectional Stigma Among Racially and Ethnically Diverse Transgender Women: Challenges and Opportunities. *Am J Public Health*. 2021;111(3):344-346. doi:10.2105/AJPH.2020.306141
5. Narayan MC. CE: Addressing Implicit Bias in Nursing: A Review. *AJN Am J Nurs*. 2019;119(7):36-43. doi:10.1097/01.NAJ.0000569340.27659.5a
6. Fujisawa D, Hagiwara N. Cancer Stigma and its Health Consequences. *Curr Breast Cancer Rep*. 2015;7(3):143-150. doi:10.1007/s12609-015-0185-0
7. Institute of Medicine. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. (Smedley BD, Stith AY, Nelson AR, eds.). The National Academies Press; 2003. doi:10.17226/10260
8. Link BG, Phelan JC. Conceptualizing Stigma. *Annu Rev Sociol*. 2001;27(1):363-385. doi:10.1146/annurev.soc.27.1.363
9. Goodyear-Smith F, Buetow S. Power issues in the doctor-patient relationship. *Health Care Anal HCA J Health Philos Policy*. 2001;9(4):449-462. doi:10.1023/A:1013812802937
10. Beisecker AE. Patient Power in Doctor-Patient Communication: What Do We Know? *Health*

Commun. 1990;2(2):105-122. doi:10.1207/s15327027hc0202_4

11. Beukeboom CJ, Burgers C, Reviews R of CROAHQL. How Stereotypes Are Shared Through Language - A Review and Introduction of the Social Categories and Stereotypes Communication (SCSC) Framework. *Rev Commun Res.* Published online January 1, 2019. Accessed February 26, 2024.

https://www.academia.edu/38218166/Beukeboom_and_Burgers_2019_How_Stereotypes_Are_Shared_Through_Language_A_Review_and_Introduction_of_the_Social_Categories_and_Stereotypes_Communication_SCSC_Framework
12. Abuse NI on D. Words Matter - Terms to Use and Avoid When Talking About Addiction. National Institute on Drug Abuse. Published November 29, 2021. Accessed January 22, 2022.

<https://www.drugabuse.gov/nidamed-medical-health-professionals/health-professions-education/words-matter-terms-to-use-avoid-when-talking-about-addiction>
13. P Goddu A, O’Conor 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
14. Glassberg J, Tanabe P, Richardson L, DeBaun M. Among emergency physicians, use of the term “Sickler” is associated with negative attitudes toward people with sickle cell disease. *Am J Hematol.* 2013;88(6):532-533. doi:10.1002/ajh.23441
15. Sun M, Oliwa T, Peek ME, Tung EL. Negative Patient Descriptors: Documenting Racial Bias In The Electronic Health Record. *Health Aff (Millwood).* 2022;41(2):203-211. doi:10.1377/hlthaff.2021.01423
16. 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

17. Chafe WL, Nichols J. *Evidentiality: The Linguistic Coding of Epistemology*. Ablex; 1986.
18. 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
19. Benamara F, Taboada M, Mathieu Y. Evaluative Language Beyond Bags of Words: Linguistic Insights and Computational Applications. *Comput Linguist*. 2017;43(1):201-264.
doi:10.1162/COLI_a_00278
20. Hagiwara N, Lafata JE, Mezuk B, Vrana SR, Fetters MD. Detecting implicit racial bias in provider communication behaviors to reduce disparities in healthcare: Challenges, solutions, and future directions for provider communication training. *Patient Educ Couns*. 2019;102(9):1738-1743. doi:10.1016/j.pec.2019.04.023
21. Hirsh AT, Hollingshead NA, Ashburn-Nardo L, Kroenke K. The Interaction of Patient Race, Provider Bias, and Clinical Ambiguity on Pain Management Decisions. *J Pain*. 2015;16(6):558-568. doi:10.1016/j.jpain.2015.03.003
22. Haywood C, Lanzkron S, Bediako S, et al. Perceived Discrimination, Patient Trust, and Adherence to Medical Recommendations Among Persons with Sickle Cell Disease. *J Gen Intern Med*. 2014;29(12):1657-1662. doi:10.1007/s11606-014-2986-7
23. Elander J, Beach MC, Jr CH. Respect, trust, and the management of sickle cell disease pain in hospital: comparative analysis of concern-raising behaviors, preliminary model, and agenda for international collaborative research to inform practice. *Ethn Health*. 2011;16(4-5):405-421. doi:10.1080/13557858.2011.555520
24. Pentin PL. Drug Seeking or Pain Crisis? Responsible Prescribing of Opioids in the

- Emergency Department. *AMA J Ethics*. 2013;15(5):410-415.
doi:10.1001/virtualmentor.2013.15.5.ecas2-1305
25. Keister LA, Stecher C, Aronson B, McConnell W, Hustedt J, Moody JW. Provider Bias in prescribing opioid analgesics: a study of electronic medical Records at a Hospital Emergency Department. *BMC Public Health*. 2021;21(1):1-9. doi:10.1186/s12889-021-11551-9
26. Morden NE, Chyn D, Wood A, Meara E. Racial Inequality in Prescription Opioid Receipt — Role of Individual Health Systems. *N Engl J Med*. 2021;385(4):342-351.
doi:10.1056/NEJMsa2034159
27. Iglar EC, Defenderfer EK, Lang AC, Bauer K, Uihlein J, Davies WH. Gender differences in the experience of pain dismissal in adolescence. *J Child Health Care*. 2017;21(4):381-391.
doi:10.1177/1367493517727132
28. Werner A, Isaksen LW, Malterud K. ‘I am not the kind of woman who complains of everything’: Illness stories on self and shame in women with chronic pain. *Soc Sci Med*. 2004;59(5):1035-1045. doi:10.1016/j.socscimed.2003.12.001
29. Werner A, Malterud K. It is hard work behaving as a credible patient: encounters between women with chronic pain and their doctors. *Soc Sci Med*. 2003;57(8):1409-1419.
doi:10.1016/S0277-9536(02)00520-8
30. Hoffmann DE, Tarzian AJ. The Girl Who Cried Pain: A Bias against Women in the Treatment of Pain. *J Law Med Ethics*. 2001;29(1):13-27.
doi:10.1111/j.1748-720X.2001.tb00037.x
31. Goyal MK, Kuppermann N, Cleary SD, Teach SJ, Chamberlain JM. Racial Disparities in Pain Management of Children With Appendicitis in Emergency Departments. *JAMA Pediatr*. 2015;169(11):996-1002. doi:10.1001/jamapediatrics.2015.1915

32. Boag W, Suresh H, Celi LA, Szolovits P, Ghassemi M. Racial Disparities and Mistrust in End-of-Life Care. In: *Proceedings of the 3rd Machine Learning for Healthcare Conference*. PMLR; 2018:587-602. Accessed February 1, 2023.
<https://proceedings.mlr.press/v85/boag18a.html>
33. Vaughn JR. *Understanding Clinical Pain Management and Patient Experiences of Pain from Electronic Health Records*. Massachusetts Institute of Technology; 2021. Accessed February 1, 2023. <https://dspace.mit.edu/handle/1721.1/140086>
34. Ahluwalia SC, Giannitrapani KF, Dobscha SK, Cromer R, Lorenz KA. “Sometimes you wonder, is this really true?”: Clinician assessment of patients’ subjective experience of pain. *J Eval Clin Pract*. 2020;26(3):1048-1053. doi:10.1111/jep.13298
35. Carr DB. Patients with Pain Need Less Stigma, Not More. *Pain Med*. 2016;17(8):1391-1393. doi:10.1093/pm/pnw158
36. NHLBI, NIH. Evidence-Based Management of Sickle Cell Disease: Expert Panel Report, 2014 | NHLBI, NIH. Published September 2014. Accessed February 13, 2021.
<https://www.nhlbi.nih.gov/health-topics/evidence-based-management-sickle-cell-disease>
37. Platt OS, Brambilla DJ, Rosse WF, et al. Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N Engl J Med*. 1994;330(23):1639-1644.
doi:10.1056/NEJM199406093302303
38. Haywood C, Lanzkron S, Ratanawongsa N, Bediako SM, Lattimer-Nelson L, Beach MC. Hospital Self-Discharge Among Adults with Sickle Cell Disease: Associations with Trust and Interpersonal Experiences with Care. *J Hosp Med Off Publ Soc Hosp Med*. 2010;5(5):289-294. doi:10.1002/jhm.643
39. Simon R, Snow R, Wakeman S. Understanding why patients with substance use disorders

- leave the hospital against medical advice: A qualitative study. *Subst Abuse*. 2020;41(4):519-525. doi:10.1080/08897077.2019.1671942
40. Brennan-Cook J, Bonnabeau E, Aponte R, Augustin C, Tanabe P. Barriers to Care for Persons with Sickle Cell Disease: The Case Manager's Opportunity to Improve Patient Outcomes. *Prof Case Manag*. 2018;23(4):213-219. doi:10.1097/NCM.0000000000000260
41. Evensen CT, Treadwell MJ, Keller S, et al. Quality of care in sickle cell disease. *Medicine (Baltimore)*. 2016;95(35). doi:10.1097/MD.00000000000004528
42. Katz IT, Ryu AE, Onuegbu AG, et al. Impact of HIV-related stigma on treatment adherence: systematic review and meta-synthesis. *J Int AIDS Soc*. 2013;16(3S2):18640. doi:10.7448/IAS.16.3.18640
43. Shankar SM, Arbogast PG, Mitchel E, Cooper WO, Wang WC, Griffin MR. Medical care utilization and mortality in sickle cell disease: A population-based study. *Am J Hematol*. 2005;80(4):262-270. doi:https://doi.org/10.1002/ajh.20485
44. Trépanier G, Laguë G, Dorimain MV. A step-by-step approach to patients leaving against medical advice (AMA) in the emergency department. *Cjem*. 2023;25(1):31-42. doi:10.1007/s43678-022-00385-y
45. Jeremiah J, O'Sullivan P, Stein MD. Who leaves against medical advice? *J Gen Intern Med*. 1995;10(7):403-405. doi:10.1007/BF02599843
46. Paul G, Gautam P, Mahajan RK, Gautam N, Ragavaiah S. Patients Leaving Against Medical Advice-A National Survey. *Indian J Crit Care Med Peer-Rev Off Publ Indian Soc Crit Care Med*. 2019;23(3):143-148. doi:10.5005/jp-journals-10071-23138
47. Ayed IA. What Makes Patients Leave Against Medical Advice? *J Taibah Univ Med Sci*. 2009;4(1):16-22. doi:10.1016/S1658-3612(09)70077-0

48. Brook M, Hilty DM, Liu W, Hu R, Frye MA. Discharge Against Medical Advice From Inpatient Psychiatric Treatment: A Literature Review. *Psychiatr Serv.* 2006;57(8):1192-1198. doi:10.1176/ps.2006.57.8.1192
49. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. *Sci Data.* 2016;3(1):160035. doi:10.1038/sdata.2016.35
50. Zestcott CA, Spece L, McDermott D, Stone J. Health Care Providers' Negative Implicit Attitudes and Stereotypes of American Indians. *J Racial Ethn Health Disparities.* 2021;8(1):230-236. doi:10.1007/s40615-020-00776-w
51. Opioid Oral Morphine Milligram Equivalent (MME) Conversion Factors | Guidance Portal. Accessed January 31, 2024. <https://www.hhs.gov/guidance/document/opioid-oral-morphine-milligram-equivalent-mme-conversion-factors-0>
52. Johnson AEW, Kramer AA, Clifford GD. A new severity of illness scale using a subset of Acute Physiology And Chronic Health Evaluation data elements shows comparable predictive accuracy. *Crit Care Med.* 2013;41(7):1711-1718. doi:10.1097/CCM.0b013e31828a24fe
53. Wang X, Xia J, Shan Y, Yang Y, Li Y, Sun H. Predictive value of the Oxford Acute Severity of Illness Score in acute stroke patients with stroke-associated pneumonia. *Front Neurol.* 2023;14. Accessed February 13, 2024. <https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2023.1251944>
54. Zhu Y, Zhang J, Wang G, et al. Machine Learning Prediction Models for Mechanically Ventilated Patients: Analyses of the MIMIC-III Database. *Front Med.* 2021;8. Accessed February 13, 2024. <https://www.frontiersin.org/articles/10.3389/fmed.2021.662340>
55. Glasgow JM, Vaughn-Sarrazin M, Kaboli PJ. Leaving Against Medical Advice (AMA): Risk

- of 30-Day Mortality and Hospital Readmission. *J Gen Intern Med.* 2010;25(9):926-929.
doi:10.1007/s11606-010-1371-4
56. Alfandre DJ. “I’m Going Home”: Discharges Against Medical Advice. *Mayo Clin Proc.* 2009;84(3):255-260.
57. Southern WN, Nahvi S, Arnsten JH. Increased Risk of Mortality and Readmission among Patients Discharged Against Medical Advice. *Am J Med.* 2012;125(6):594-602.
doi:10.1016/j.amjmed.2011.12.017
58. Albayati A, Douedi S, Alshami A, et al. Why Do Patients Leave against Medical Advice? Reasons, Consequences, Prevention, and Interventions. *Healthcare.* 2021;9(2):111.
doi:10.3390/healthcare9020111
59. Compton P, Aronowitz SV, Klusaritz H, Anderson E. Acute pain and self-directed discharge among hospitalized patients with opioid-related diagnoses: a cohort study. *Harm Reduct J.* 2021;18(1):1-12. doi:10.1186/s12954-021-00581-6
60. Dehon E, Weiss N, Jones J, Faulconer W, Hinton E, Sterling S. A Systematic Review of the Impact of Physician Implicit Racial Bias on Clinical Decision Making. *Acad Emerg Med.* 2017;24(8):895-904. doi:10.1111/acem.13214
61. 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program. Federal Register. Published May 1, 2020. Accessed June 14, 2022.
<https://www.federalregister.gov/documents/2020/05/01/2020-07419/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification>
62. Information Blocking | HealthIT.gov. Accessed December 26, 2023.
<https://www.healthit.gov/topic/information-blocking>

Appendix 1

Table 4. Sensitivity analysis 1, where we limited to patients with self-reported pain scores, but did not include the pain variable. Regression results showing relationships between stigmatizing labels, doubt marking linguistic EHR features per chart with opioid

medication rates and whether patients completed self-directed discharge associations. Log Betas (95%CI) (Note: these are not exponentiated and should not be interpreted as odds ratios)

	Avg Daily MME Prescribed (Unadjusted)	Avg Daily MME Prescribed (Adjusted)	Avg Daily MME Dispensed (Unadjusted)	Avg Daily MME Dispensed (Adjusted)	Patient self-directed discharge (Unadjusted)	Patient self-directed discharge (Adjusted)
Stigmatizing Label Count, Per Chart	.51 (-1.19, 2.20)	.06 (-1.56, 1.68)	.03 (-.14, .20)	.01 (-.15, .18)	1.13 (.34, 1.88)* p = .0033	1.25 (.42, 2.08)* p = .003
Doubt Marker Label Count, per chart	-0.01 (-2.85, 2.83)	-1.00 (-3.71, 1.70)	-.11 (-.39, .18)	-.17 (-.45, .11)	2.09 (.95, 3.22) p = .0003	1.89 (.64, 3.13) p = .0029

*p is significant at <.05 value

**p is significant at <.0001 value

Adjusted models include covariates of: Gender, total length of stay, average daily pain score, Black/African American identified race, insurance (Government-run, private, self-pay), average OASIS severity index score, and whether patient was diagnosed with obesity, opioid use disorder, symptomatic HIV, or sickle cell disease.

Table 5. Sensitivity analysis 2, where we included all patients (with or without self-reported pain scores), and did not include the pain variable. Log Betas (95%CI) (Note: these are not exponentiated and should not be interpreted as Odds ratios)

	Avg Daily MME Prescribed (Unadjusted)	Avg Daily MME Prescribed (Adjusted)	Avg Daily MME Dispensed (Unadjusted)	Avg Daily MME Dispensed (Adjusted)	Patient self-directed discharge (Unadjusted)	Patient self-directed discharge (Adjusted)
Stigmatizing Label Count, Per Chart	.73 (-.69, 2.15)	.19 (-1.17, 1.54)	.082 (-.063, .226)	.05 (-.09, .19)	1.38 (.65, 2.12)* p = .0002	1.48 (.76, 2.20) p <.0001
Doubt Marker Label Count, per chart	.25 (-2.14, 2.65)	-1.00 (-3.29, 1.28)	-.047 (-.290, .197)	-.13 (-.37, .10)	2.23 (1.09, 3.37) p=.0001	1.94 (.70, 3.18) p = .002

*p is significant at <.05 value

**p is significant at <.0001 value

Adjusted models include covariates of: Gender, total length of stay, average daily pain score, Black/African American identified race, insurance (Government-run, private, self-pay), average OASIS severity index score, and whether patient was diagnosed with obesity, opioid use disorder, symptomatic HIV, or sickle cell disease.

Dissertation Conclusion

This project represents a successful effort to develop a stigma and doubt marker detection system within electronic health records, and apply this system to assess inequities in the distribution of these linguistic features among a large dataset of medical records across marginalized groups, and examine connections with quality of care indicators. In Aim 1, we proved that this detection system was feasible and performs at levels close to human annotator agreement. We found that the concept of “scare quotes” needs further refinement in concept identification and classification, perhaps due to evolution of how quotes have been and continue to be used. Aim 2 illuminated stark differences among groups by factors of gender, race, insurance status, and stigmatizing diagnosis, where historically marginalized groups received disproportionately higher rates of doubt markers and stigmatizing labels as groups who had higher societal privilege. These trends further substantiate previous research which have identified stigmatization within smaller datasets by applying them to a large-scale de-identified electronic health record. We also identified differences among provider types, where nurses and social workers had significantly higher rates of stigmatizing label and doubt marker usage per patient chart than physicians. The high rates of clustering at the patient level indicate that this phenomenon accumulates around certain patients, and is a product of teams of providers. Aim 3 identified that stigmatizing labels and doubt markers were not associated with changes in prescription of opioids. We found that stigmatizing labels and doubt markers significantly increased likelihood of self-directed discharge, pointing to the connection between these linguistic features and patient-care team trust. Ultimately, this project lays a foundation for further scalable development and application in the context of identifying stigmatizing language, with value across a variety of care settings, patient groups, and health outcomes.