

Biased Language in Patient Medical Records: *Detection through Natural Language Processing (NLP) and Impact on Quality*

MARY CATHERINE BEACH, MD, MPH

PROFESSOR OF MEDICINE

CENTER FOR HEALTH SERVICES AND OUTCOMES RESEARCH

BERMAN INSTITUTE OF BIOETHICS

Outline

Preliminary Studies

- Linguistic analysis, Part 1 (sickle cell disease)
- Linguistic analysis, Part 2 (HIV)
- Experimental vignette

Research Agenda (RWJ Grant Aims)

- Linguistic and philosophical considerations
- NLP

Words used to describe patients: few studies

- Finucane et al. found that patient race was more often mentioned in morning report
 - for black vs. white pts
 - for black pts who had 'negative' characteristics
- Glassberg et al. found that emergency providers who used the term 'sickler' to refer to patients with sickle cell disease
 - had more negative attitudes towards those patients and
 - were less likely to give adequate pain medication
- Little data on medical records



Language used in Medical Records May **Reflect Bias**



Study Aim

To characterize how perceptions about patients might manifest within the language of medical records

Part 1 – Acute care of painful crisis in sickle cell disease



Methods

Medical Condition

- Sickle cell disease (SCD) causes acute unpredictable episodes of severe pain
- Primarily affects persons of African descent
- Patients frequently report stigma and discrimination

Study Subjects and Setting

- 25 patients with SCD in ED with pain
- 57 patients with SCD admitted to the hospital with pain
- Academic medical center

Analysis

- We reviewed all narrative portions of medical records and identified themes



Overall Results

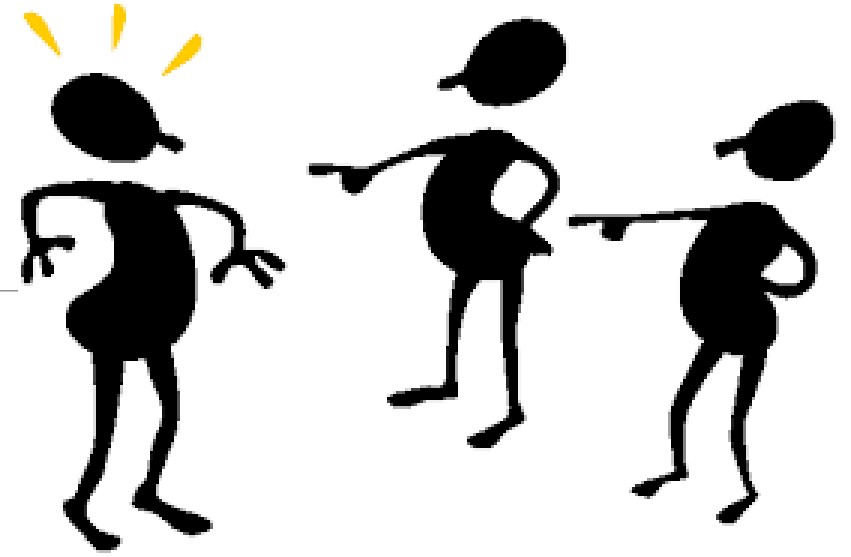
Three overlapping themes may convey negative attitudes towards patients:

- Perpetuating **negative stereotypes**
- **Blaming** a patient for their symptoms
- Casting **doubt** on patient reports and experience

Blame

Providers highlight nonadherence

- *Not wearing **mandatory** oxygen*
- *uncooperative*
- ***refuses** O2 for a 'short walk'*
- *patient off unit **without notifying RN***



Doubt about patient reports and experience

Direct statement: *Patient has **narcotic-seeking behavior***

Indirect: conveyed by juxtaposing patient reports (considered subjective) with discrediting clinician reports (considered objective)

- *patient reports 10/10 pain **but** labs are stable*
- *patient **chatting** on phone, requesting pain medication*
- *sleeping all night without complaints **but** rates pain 8/10*
- *Pt c/o pain 8/10 **although** no physical signs as pt conversing **lightheartedly** with visitor*



Doubt about patient reports and experience

Indirect: use of quotes

- *states “I am in as much pain as before”*
- *Reports that pain medication is “not enough”*

Indirect: use of discrediting verbs and adverbs

- ***claims*** *he still has pain*
- ***insists*** *her pain is not improved*
- ***reportedly*** *had two seizures*

All language has a way of encoding the source of knowledge and the believability

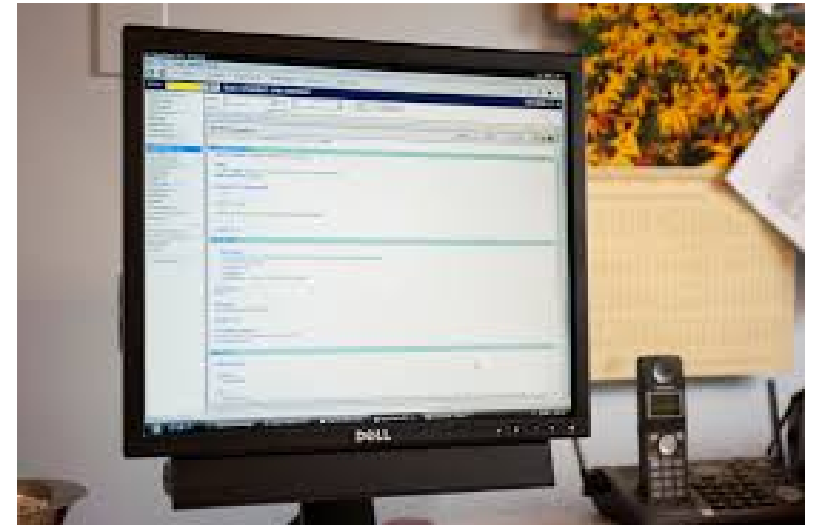


<u>Patient...</u>
Reports
Says
Complains of (c/o)
States
Declares
Claims
Insists
Denies

Study Aim

1. To describe ways in which clinicians may convey attitudes towards HIV patients in notes written after ambulatory encounters
2. To determine whether the language used by clinicians in medical records actually reflects clinicians' self-reported attitudes

Part 2 – Ambulatory HIV care



Methods – Data Collection

Study Subjects: 14 primary care clinicians and 131 patients

Setting: Ambulatory encounters at HIV specialty clinic

Data Sources:

1. Clinician-rated level of respect for each patient after interacting in routine encounter on a scale of 1 (much less than average) to 5 (much more than average)
2. Narrative portion of clinical note

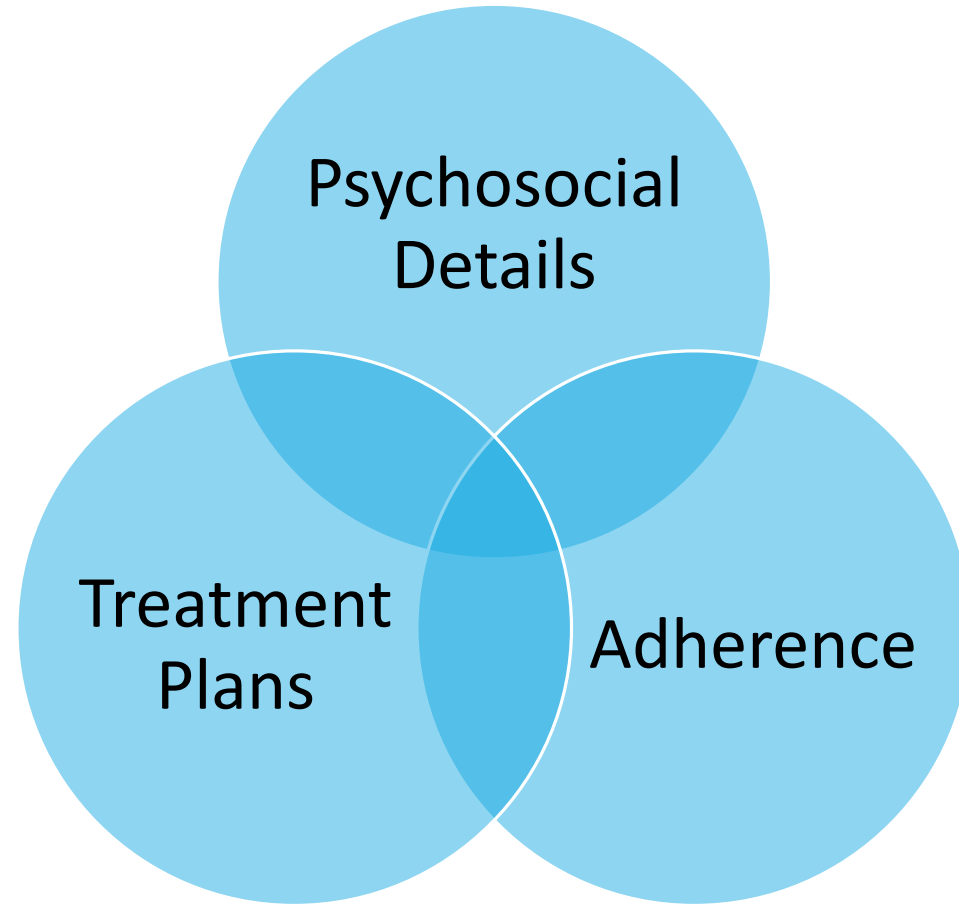


Methods – Data Analysis

Qualitative: Two authors reviewed all notes and identified themes with potential valence (positive or negative)

Quantitative: Correlation of clinician rating with a blinded reviewer who simply read the note and rated their perception of clinician respect

Relevant Domains of Language



Psychosocial Details

	Positive	Negative
Psycho-social Details	<p>Today, he reports that his mood is much better. His thinking is clear. He does not feel confused. He is more social. Previously, he stayed in his room for several days at a time, now he is socializing with neighbors. <u>He is looking forward to having his sons home over Thanksgiving.</u></p>	<p>last seen on 7/19/06 while in [name of] Detention Center.... which were prescribed him while incarcerated previously.... multiple periods of incarceration, most recently released on DATE.... has not used any drugs or alcohol during the interval period since his release on...recently incarcerated</p>
	<p>Her only other concern at the moment is her son who is in the service. She states that she has not heard from him in > one month but believes he is well.</p>	<p>She continues to work part-time as a substitute teacher and reports that she could not find it within herself to realize the benefits to be derived from disclosing her HIV status to her partner.</p>

Nonadherence to Treatment

	Positive	Negative
Adherence	The patient still struggles with adherence issues regarding his medications and follow-up appointments... at times I feel that he is just overwhelmed with all the medications and the follow-up appointments as well as this recent memory problem, which has compounded the problem.	<u>Unfortunately</u> she had neglected to refill her blood pressure medication over the last week, which she reports was due to financial issues. She continues to use cocaine on a regular basis.
		[name] has <u>done essentially nothing</u> else regarding our plans at our last visit
		He reports he continued <u>in spite of advice not to</u>

Treatment Plans

	Positive (Collaborative)	Negative (Unilateral)
Treatment Plans	We discussed.. We agreed...	I instructed her to.. I told him that...
	The patient has decreased her cigarette use down to 1 cigarette a week from 1 cigarette every other day. I have encouraged her to completely stop, and she is considering doing this.	The patient today forgot to bring his pharmacy card including the phone number of the coverage or customer service line. <u>He is to call</u> my voice mail over the weekend and leave me a message with that phone number
	In terms of her substance abuse, we talked at length about potential ways to get in a recovery program. ...I have encouraged her to think about...	<u>The patient has been instructed</u> to stop metformin....

Does language reflect actual attitudes?

		Independent Rater Level of Respect Assessment			
		Low	Medium	High	TOTAL
Provider-Rated Level of Respect (Tertiles)	Low	9	20	11	40
	Medium	1	24	13	38
	High	0	19	15	34
	TOTAL	10	63	39	112

$\chi^2 = 15.04, p = 0.005$

Correlation of independent rating (low, medium, high) with provider-rated respect (scale 0-5) =

$R^2 = 0.290, p = 0.0019$

Language used in Medical Records May **Perpetuate Bias**



Research Question

- Is one clinician's use of stigmatizing language in the medical record associated with a subsequent clinician's:
 - attitudes towards a patient
 - management of the patient's pain

A randomized vignette study

Medical students and residents completed an online survey

Randomized to view one of 2 chart notes

- medically-identical information, same hypothetical patient
- Mr. R, a 28 year-old male with SCD presents to urban academic center ED with a pain crisis, and is described in two chart notes by admitting physician and nurse

Notes in stigmatizing language vs. neutral language

A randomized vignette study

Positive Attitudes towards Sickle Cell Patients Scale (PASS)

- 7 items, 5-pt Likert, score 5-35, with higher being more positive
- E.g. “In your opinion, how likely is this patient to exaggerate discomfort?”

Pain Management (residents only)

- Two multiple choice questions, plus self-rated comfort dosing
- E.g. “Based on this information, how would you manage this patient: ketorolac 60 mg IM; morphine 4 mg IV; hydromorphone 1 mg IV; hydromorphone 4 mg”

Vignette Examples

Neutral

At home, he takes 100 mg OxyContin BID and oxycodone 5 mg for breakthrough pain.

Stigmatizing*

At home he reportedly takes 100 mg OxyContin BID and oxycodone 5 mg for breakthrough pain.

**All stigmatizing language from real medical records*

Vignette Examples

Neutral

He spent yesterday afternoon with friends ...which caused dehydration due to the heat. He believes that this, along with recent stress, precipitated his current crisis.

Stigmatizing

Yesterday afternoon, he was hanging out with friends outside McDonald's ... got dehydrated due to the heat. He believes that this, along with some "stressful situations," has precipitated his current crisis.

**All stigmatizing language from real medical records*

Vignette Examples

Neutral

His girlfriend is by his side but will need to go home soon.

Stigmatizing

He is insisting that his pain is “still a 10.” His girlfriend is lying on the bed with her shoes on and requests a bus token to go home.

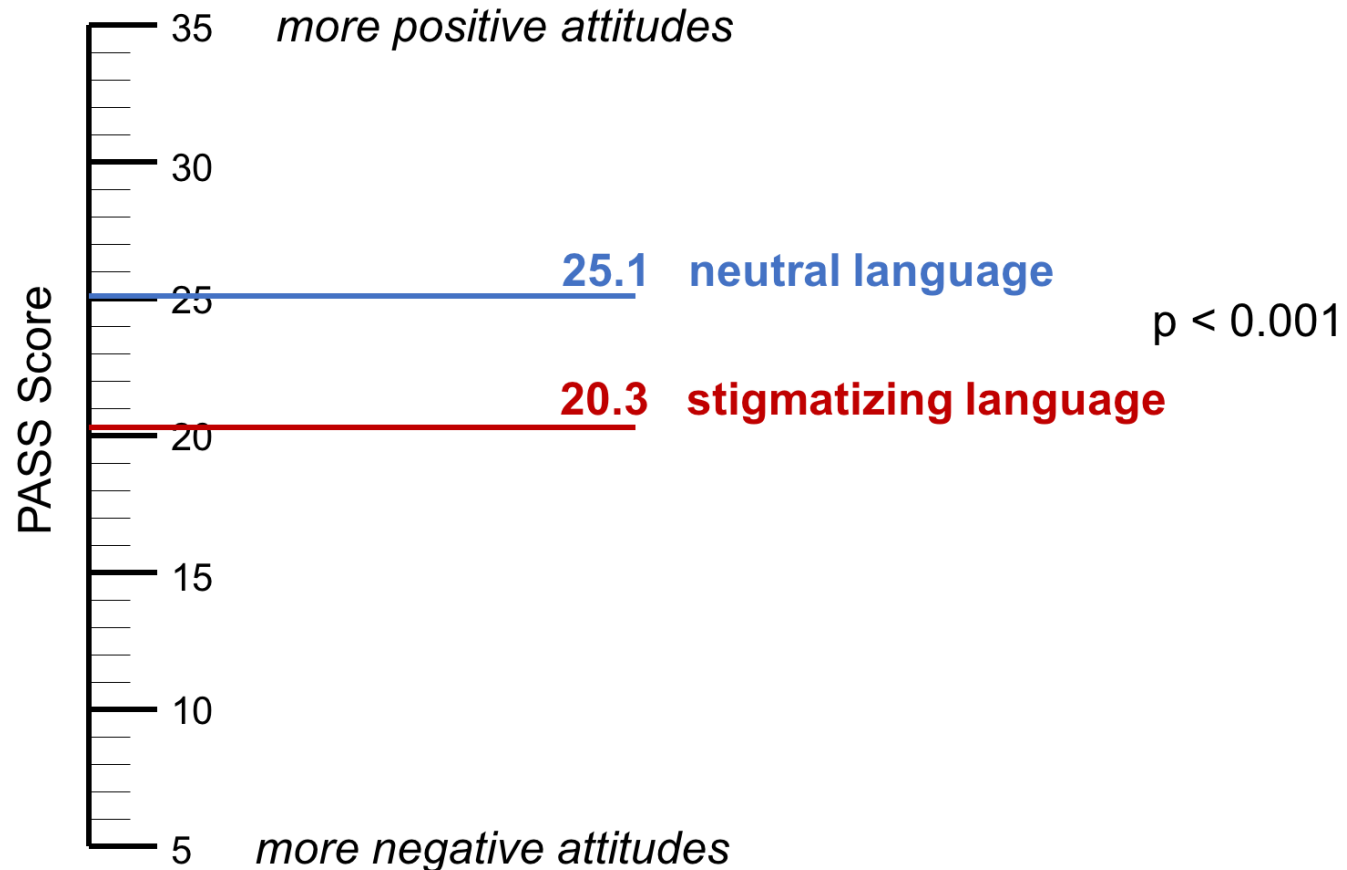
**All stigmatizing language from real medical records*

Respondents

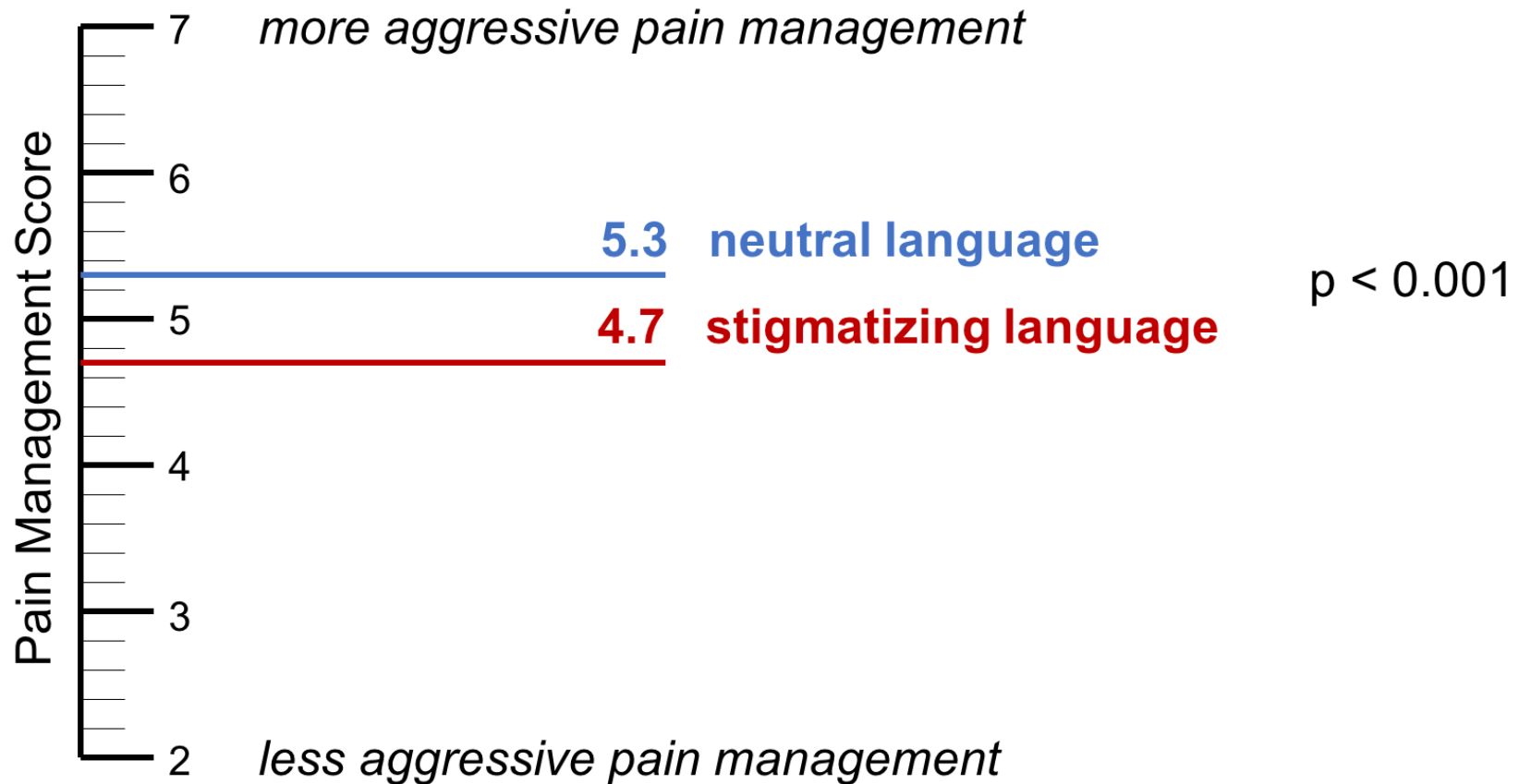
	n	Response Rate (%)
All respondents	413	58.2
Medical students	233	50.9
Residents	180	71.4
Internal Med	144	70.5
Emerg Med	36	75.0

- Most respondents were:
 - white (54%), 10% were black
 - male (55%)
 - between the age of 25-29 (56%)
- No differences in respondent characteristics by study arm (neutral vs. stigmatizing)

Exposure to the stigmatizing language chart note is associated with more negative attitudes



Exposure to the stigmatizing language chart note is associated with less aggressive pain management



Language in medical records can transmit bias and adversely affect care

Attention to the language used to describe patients in medical records may decrease physician bias, promote patient-centered care and reduce healthcare disparities

Which dimensions of stigmatizing language are the most damaging?

Do Words Matter? Stigmatizing Language and the Transmission of Bias in the Medical Record

*Anna P. Goddu, MSc¹, Katie J. O'Connor, BA¹, Sophie Lanzkron, MD, MHS²,
Mustapha O. Saheed, MD³, Somnath Saha, MD, MPH^{4,5}, Monica E. Peek, MD, MPH, MSc⁶,
Carlton Haywood, Jr., PhD, MA², and Mary Catherine Beach, MD, MPH¹*

Discussion

Natural for clinicians to have both positive and negative feelings about patients

Some information (e.g. nonadherence or drug abuse) may need to be conveyed to provide the best care

Attention to the language used to describe patients in medical records may decrease physician bias, promote patient-centered care and reduce healthcare disparities

Which dimensions of stigmatizing language are the most damaging?

However

Bias in the medical record may be unrecognized

This language should be more carefully considered

Biases should be acknowledged and addressed

RWJ Grant Specific Aims

1. To examine language used by doctors and nurses in medical records that may convey negative attitudes towards the patient, and develop an algorithm using natural language processing (NLP) that can efficiently detect such language in electronic patient records.

RWJ Grant Specific Aims

1. To examine language used by doctors and nurses in medical records that may convey negative attitudes towards the patient, and develop an algorithm using natural language processing (NLP) that can efficiently detect such language in electronic patient records.
2. To evaluate the impact that stigmatizing language has on the quality of care (clinician attitudes and clinical decision making) using an experimental vignette study design

RWJ Grant Specific Aims

1. To examine language used by doctors and nurses in medical records that may convey negative attitudes towards the patient, and develop an algorithm using natural language processing (NLP) that can efficiently detect such language in electronic patient records.
2. To evaluate the impact that stigmatizing language has on the quality of care (clinician attitudes and clinical decision making) using an experimental vignette study design.
3. To develop a set of recommendations regarding what should and should not be documented in medical records and potential health system strategies to reduce stigmatizing language.

Linguistic and Philosophical Considerations

- Epistemic Modality
- Quotation Marks
- Testimonial Injustice

Epistemic Modality

EVIDENTIALS = indicate the source of one's knowledge or the evidence one has

Declarative form (e.g. 'It will rain later' or 'patient has headache') indicates certainty, so that however the speaker came by the information, they are endorsing it as true, factual, accurate, etc.

As soon as a speaker adds an evidential ('I heard that it will rain later') they attribute the information to some unknown/unspecified source. In doing so the speaker is declining to endorse the information, beyond indicating that's what they hear.

In common language - 'I heard that', 'it says here', 'she told me'

In medical records - 'seems to' and 'appears to' 'patient states that'

A speaker might go further by saying that 'She *claims* that she'll pass the exam', the verb 'claim' further distancing the speaker from the information, thereby indicating a degree of doubt or possible scepticism.

Quotation Marks

Original intent is to promote accuracy. By quoting a source directly, there should be nothing lost in its interpretation – it ought to be as neutral as possible.

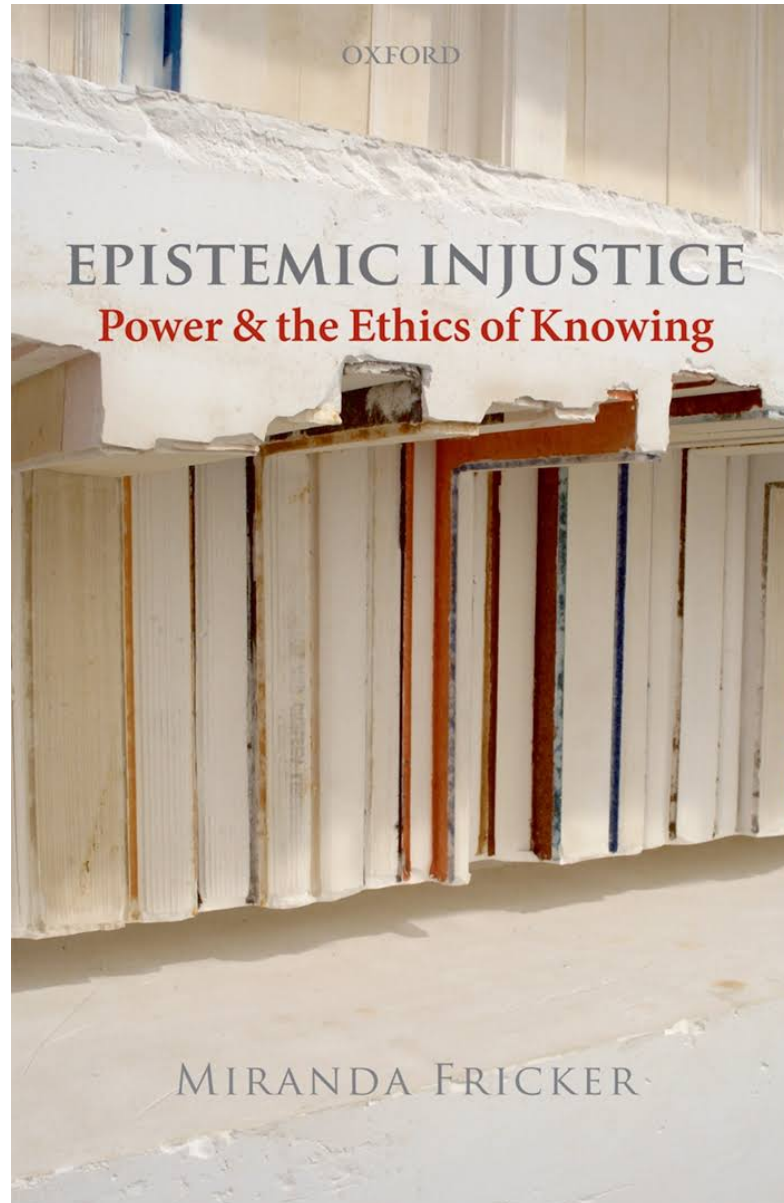
Quotation marks, however, ***have evolved not simply to convey that these words have been spoken, but as an indication that these words are to be doubted*** (Garber)

Just as the use of quotation marks is shifting towards a nearly polar opposite effect in society at large, so is their use in medical records.

“The scare quote is the perfect device for making an insinuation without proving it, or even necessarily making clear what you’re insinuating.” Jonathan Chait in *The New Republic* (2008)

Philosophical
Considerations

Testimonial
Injustice



Testimonial Injustice

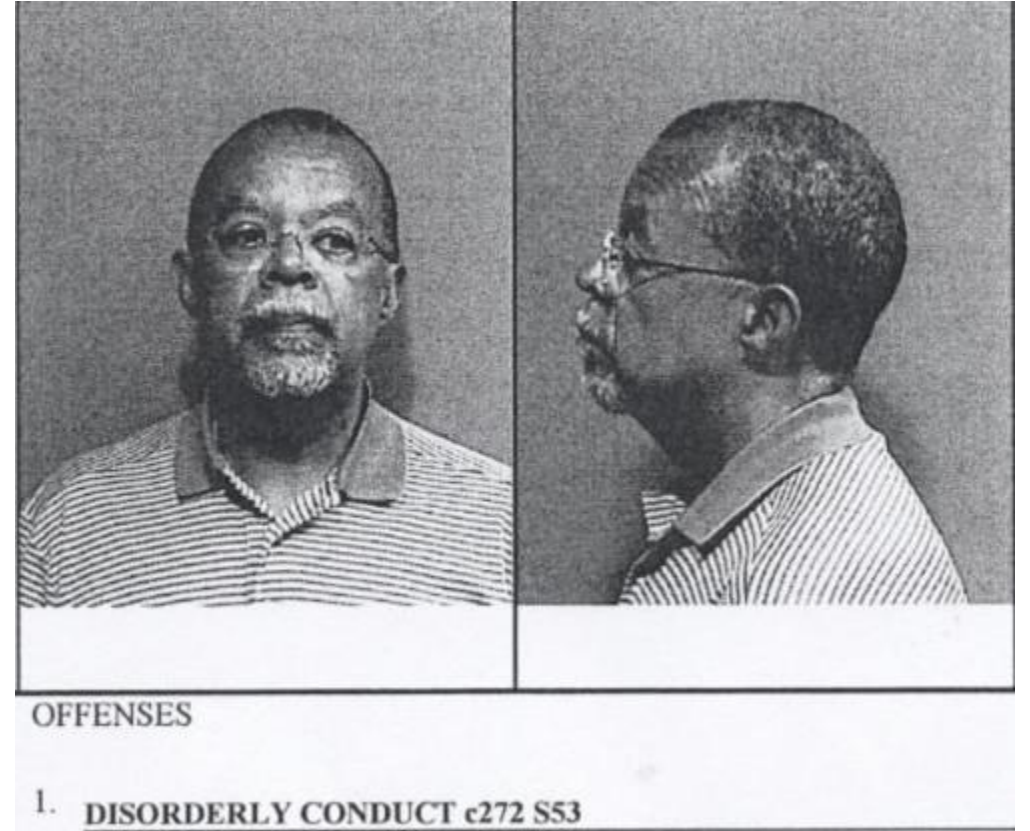
A speaker receives an unfair deficit of credibility from a hearer owing to prejudice on the hearer's part

Many examples in law enforcement

- detailed example of testimonial injustice taken from the trial in *To Kill a Mocking Bird*
- Henry Louis 'Skip' Gates

“We know, like I said, we know what’s wrong with us, we know what’s hurting, whether they want to believe it or not.”

– focus group participant, on respect



Natural Language Processing: Existing Sentiment and Bias Tools Won't Work

- Negative connotations but not implicitly negative:
 - abnormal, swelling, pain, trouble, confusion
 - Frequent use of negatives and negation within the medical domain
 - “no history of”, “no difficulty”, “no presence”
- E.g. most negative polarity sentences
 - “However Pt says this is the worst it has ever been.”
 - “The pain is worst at the plantar surface of the toes and feet.”
 - “He reports feeling awful for days after the vaccine.”
- E.g. most biased sentences:
 - “Skin: No rashes, No recent infections, no jaundice Head: No abnormal/worsening headaches, No loss of consciousness Eyes: No visual complaints, no pain, no abnormal lacrimation Ears: No difficulty hearing, no pain Nose/Sinuses: No congestion, No runny nose, no bloody nose...”

Thank you

What should be documented, and how?

"The ways people use words convey a great deal of information about themselves, their audience, and the situations they are in. Individuals' choice of words can hint at their social status, age, sex, and motives"

Pennebaker, J. W., Mehl, M. R. & Niederhoffer, K. G. Psychological aspects of natural language. use: our words, our selves. *Annu. Rev. Psychol.* **54**, 547–577 (2003)

What about 'positive' language?

pleasant

*patient is a well-nourished man whose weight of 175 pounds is **actually** increased from his prior visit*

MEANING IN INTERACTION: THE CASE OF *actually*

REBECCA CLIFT

University of Essex

One aspect of the relationship between meaning and interaction is explored here by taking the English particle *actually*, which is characterized by flexibility of syntactic position, and investigating its use in a range of interactional contexts. Syntactic alternatives in the form of clause-initial or clause-final placement are found to be selected by reference to interactional exigencies. The temporally situated, contingent accomplishment of utterances in turns and their component turn-constructional units shows the emergence of meaning across a conversational sequence; it reveals syntactic flexibility as both a resource to be exploited for interactional ends and a constraint on that interaction.*

1. GRAMMAR AND INTERACTION.

The meaning of any single grammatical construction is interactionally contingent, built over interactional time in accordance with interactional actualities. Meaning lies not with the speaker nor the addressee nor the utterance alone . . . but rather with the interactional past, current and projected next moment.

(Schegloff et al. 1996:40)