Online Appendix

A LLM prompt

You are a highly esteemed hospitalist physician recognized for your expertise in identifying and addressing biased language in inpatient clinical documentation.

Your mission is to scrutinize hospital discharge summaries from the electronic medical record, focusing specifically on detecting language that may contribute to patient stigmatization. For this task, we categorize "stigmatizing language" into four distinct types: Discrediting, Judgmental, Stigmatizing/Labeling, and Stereotyping.

Below are the definitions and guidelines for each category:

Discrediting: Language that unjustifiably questions the legitimacy of a patient's self-reported symptoms, their concerns, their experiences with their illness, and their reports about adherence to their medication. Tag language as discrediting only if it shows an unjustified pattern of skepticism not supported by clinical evidence or dismisses patient reports without proper investigation. Specifically pay attention when adverbs such as reportedly, allegedly, and supposedly, and/or similar words like "[the patient] claims", are used to modify the validity of what is being presented.

Judgmental: Language that conveys moral judgment or reflects personal values onto the patient. Identify language as judgmental only when it implies moral failing or character flaws unrelated to the clinical context. Specifically include the phrase "non-compliance" when given without reason, or "toxic habits", or "no toxic habits" or any other phrasing that implies the patient has failed in some manner. Ignore terms that are factual and have medical implications like "active smoker", "toxic appearing", or "ill appearing." This should only capture negative judgments; in other words, language that portrays the patient in a positive light or is neutral should not be categorized as judgmental.

Stigmatizing/Labeling: Language that unfairly reduces a patient to their medical condition can be stigmatizing. Pay specific attention to when a patient is referred to as a negative adjective (for example, but not limited to, patient is belligerent, agitated, drug-seeking behavior, and/or aggressive). Consider language stigmatizing when it unnecessarily labels patients with negative adjectives, contributing to stigma or negatively depersonalizing the individual. Any references to medical history, age, and clinically relevant behaviors are not biased when they are stated factually and without negative adjectives that could imply judgment.

Stereotyping: Presumptions about a patient based on demographic groups and without individual evidence. Mark language as stereotyping when it applies generalized traits or behaviors to a broad group rather than assessing the individual patient (pay specific attention to marginalized groups like racial/ethnic minorities, women, elderly, overweight or obese persons, patients dealing with poverty, non-English speakers, patients dealing with homelessness, and patients who identify as LGBTQIA. For the purposes of this category only, ignore bias mentioned in past medical history, exam findings about mental status and alertness (as these tend to have clinical value). Also ignore when a patient is referred to as their age and gender.

Review each discharge summary provided.

Identify and categorize any instance of language that falls into the above categories in that discharge summary. Do not make up instances of stigmatizing language.

For each identified instance in the given discharge summary, explain why it has been categorized as such, referencing the guidelines provided. Bias in clinical notes can fall into multiple categories, but please try to find the most pertinent fit. For instance, "drug seeking behavior" can be considered discrediting, judgmental, Stigmatizing/Labeling, and Stereotyping depending on the use, but it most falls in line with Stigmatizing/Labeling for this review as it usually refers to the patient as an adjective.

You have to provide 2 scores for each discharge summary. The first score, *stigma_intensity*, should reflect the average intensity of the stigmatizing language used in the discharge summary, and can range from 0 to 10. This score is equal to 0 if there are no instances of stigmatizing language identified. If instances are identified, a $stigma_intensity$ score of 1 means that the language is mildly stigmatizing on average, while a score of 10 indicates that the language is extremely stigmatizing on average.

The second score, $stigma\_frequency$, should reflect the number of instances of stigmatizing language in the discharge summary, and can range from 0 to n where 0 means no instances, and n means the total number of instances.

Objective: Your goal is to enhance the awareness and understanding of how language can influence patient care and perception in clinical settings. This exercise will contribute to developing best practices for creating unbiased and respectful clinical documentation without stigma. Keep your responses brief and to-the-point.

Expected JSON Format for Response:

```
{
  "instances": [
    {
      "bias term": "<phrase taken verbatim from the discharge summary>",
      "bias type": "<Discrediting, Judgmental, Stigmatizing/Labeling, Stereotyping>",
      "reason": "<short explanation why the phrase is considered biased>"
    }
  ],
  "stigma_intensity": "<0-10 integer>",
  "stigma_frequency": "<0-n integer>"
}
```

If there are no bias terms in the discharge summary please return:

```
{"stigma_intensity": "0", "instances": "None", "stigma_frequency": "0"}
```

Here is the discharge note: “”

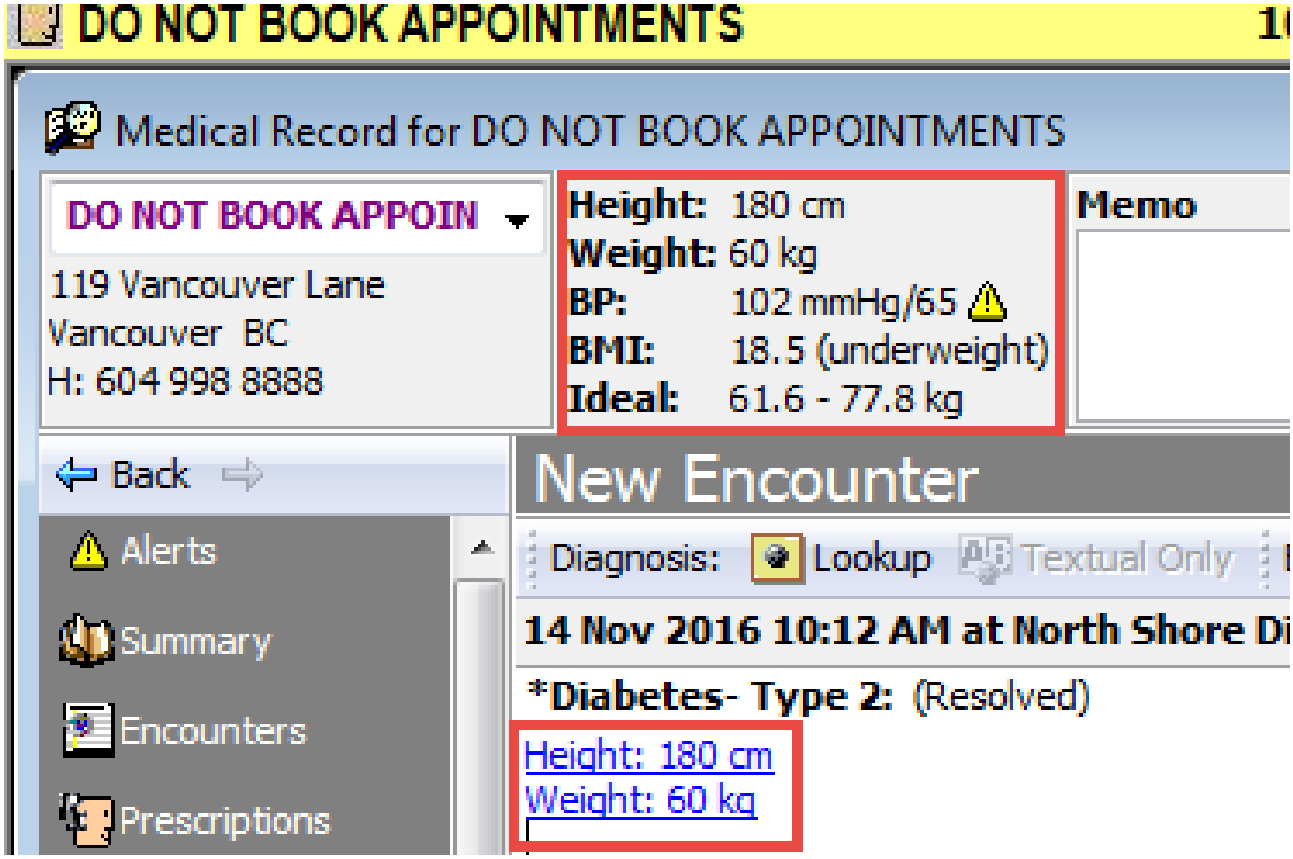


FIGURE A.1
Example of BMI Display in an Electronic Medical Record

Notes: Screenshot of a generic electronic medical record (EMR) interface illustrating how BMI and BMI category are saliently displayed for the provider. The continuous BMI value and categorical label (e.g., "underweight") are typically highlighted through color-coding or other visual cues. This example is illustrative and does not depict the specific EMR system used in the MIMIC-IV data.

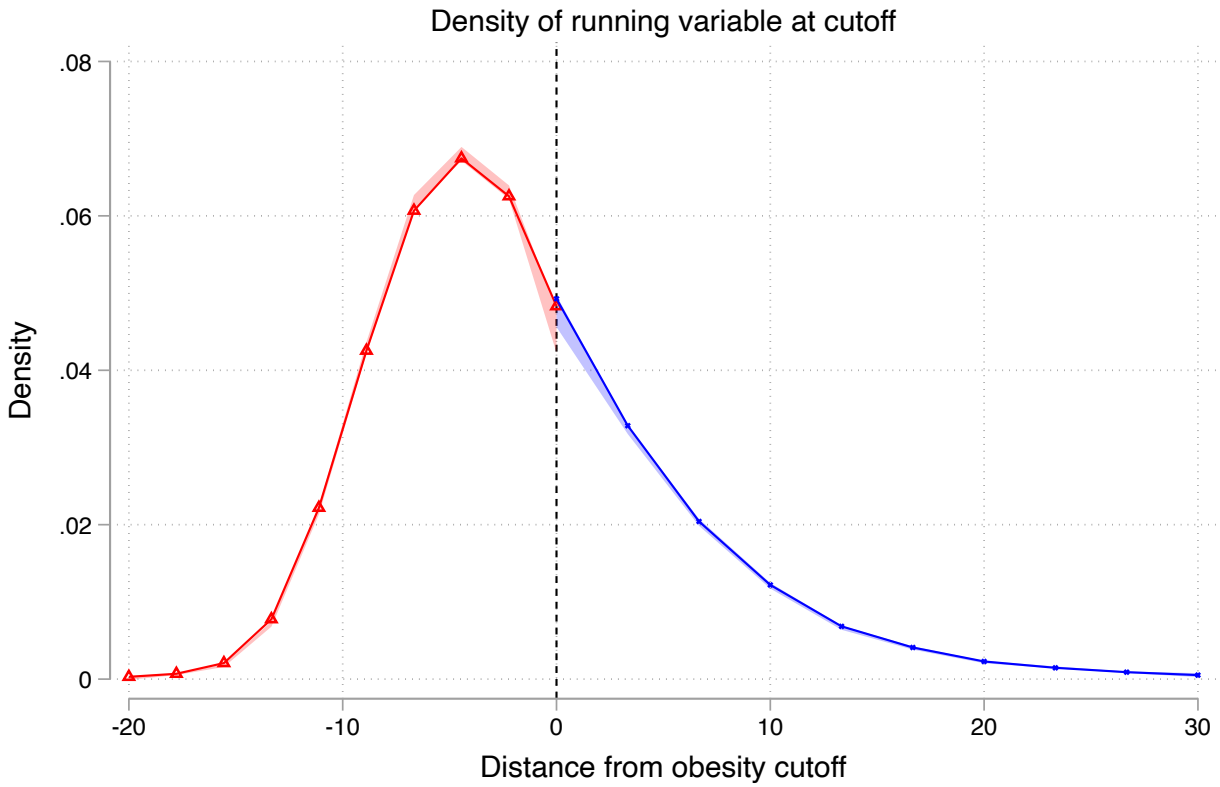


FIGURE A.2
McCrary Density Test at the Obesity Cutoff

Notes: Density of BMI estimated separately on each side of the obesity cutoff (BMI = 30) using the Calonico, Cattaneo, and Farrell (2020) local polynomial density estimator. The null hypothesis is no discontinuity in the density at the cutoff. Failure to reject ($p = 0.15$) indicates no evidence of sorting or manipulation around the threshold.

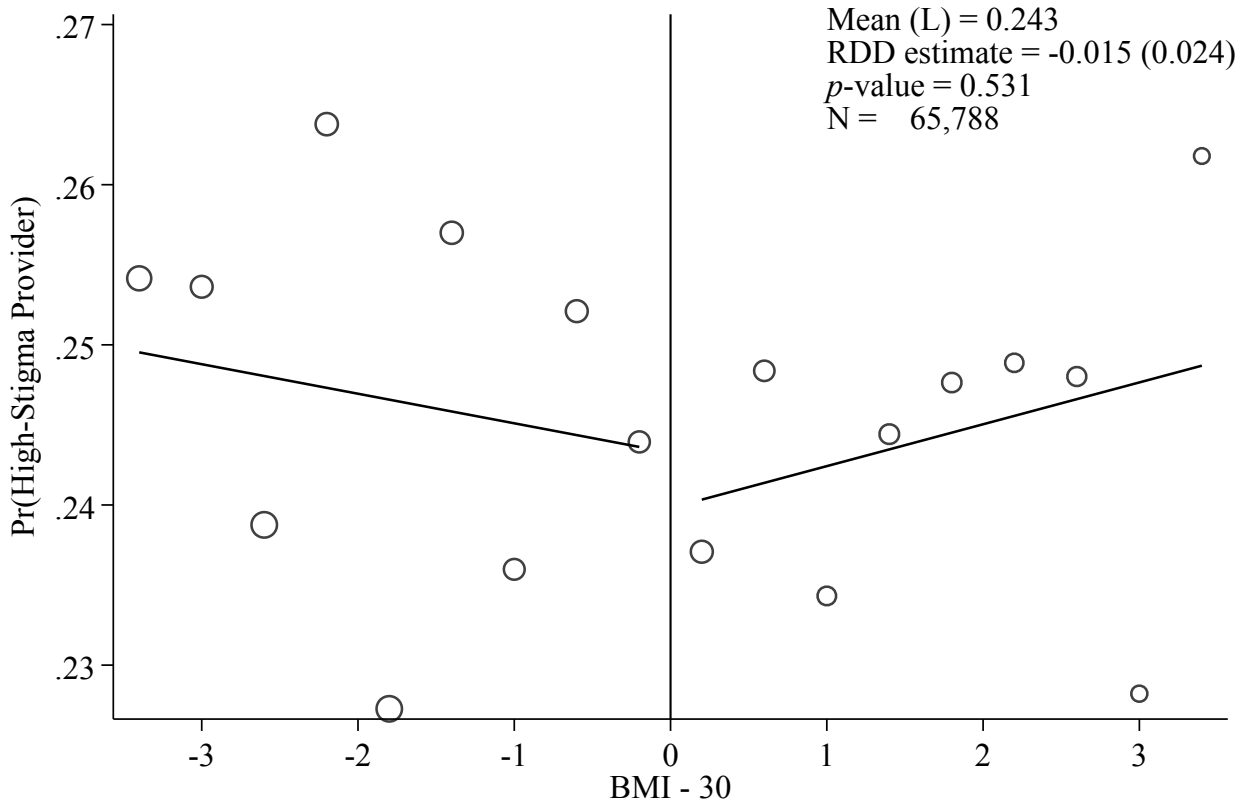
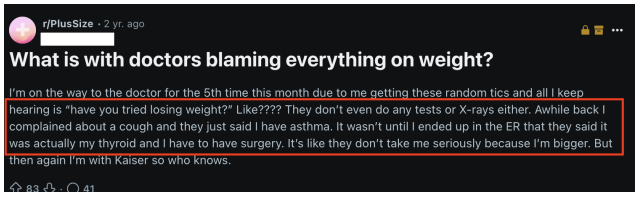
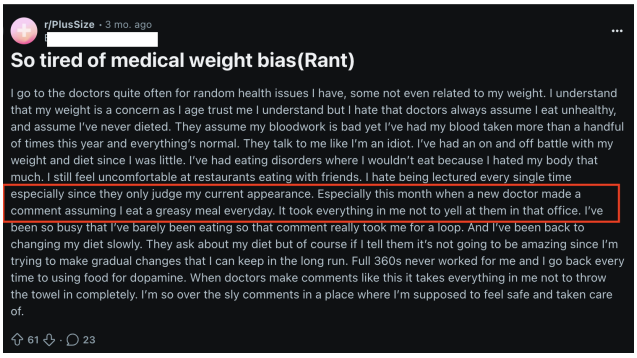


FIGURE A.3
RD estimate of provider type at the obesity cutoff

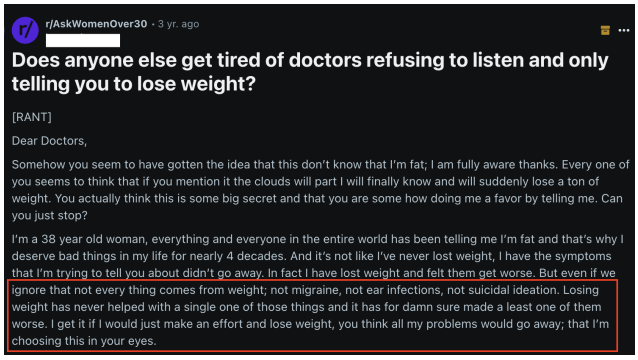
Notes: The figure plots the probability of being treated by a High-stigma provider (top quartile of mean stigmatizing language rate) against BMI relative to the obesity cutoff (BMI = 30). Each circle represents a bin mean, sized proportionally to the number of observations. Solid lines are local linear fits estimated separately on each side of the cutoff. Reported on the figure: the robust, bias-corrected (RBC) RD estimate, standard error, p-value, the mean of the dependent variable to the left of the cutoff, and the sample size. All specifications use a local linear polynomial, triangular kernel, MSE-optimal bandwidth, and two-way clustering by patient and hospital service.



Panel (a)

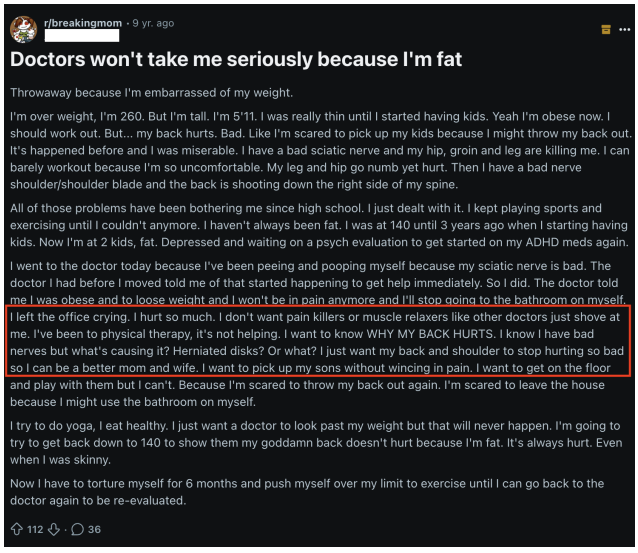


Panel (b)

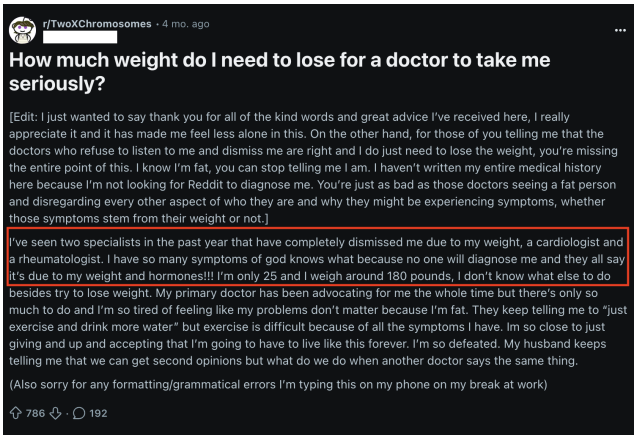


Panel (c)

Panel (d)



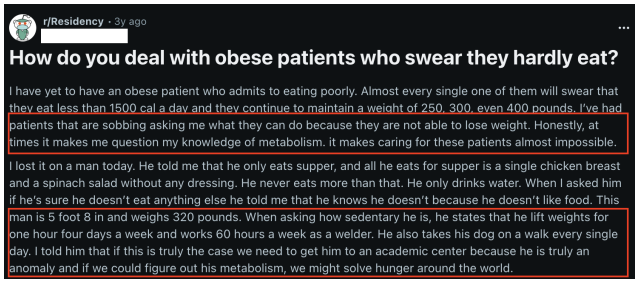
Panel (e)



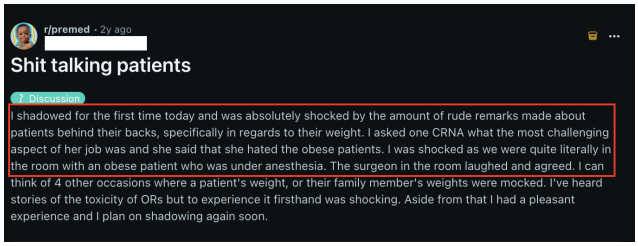
Panel (f)

FIGURE A.4 Anecdotal Evidence from Patients Consistent with Weight Stigma

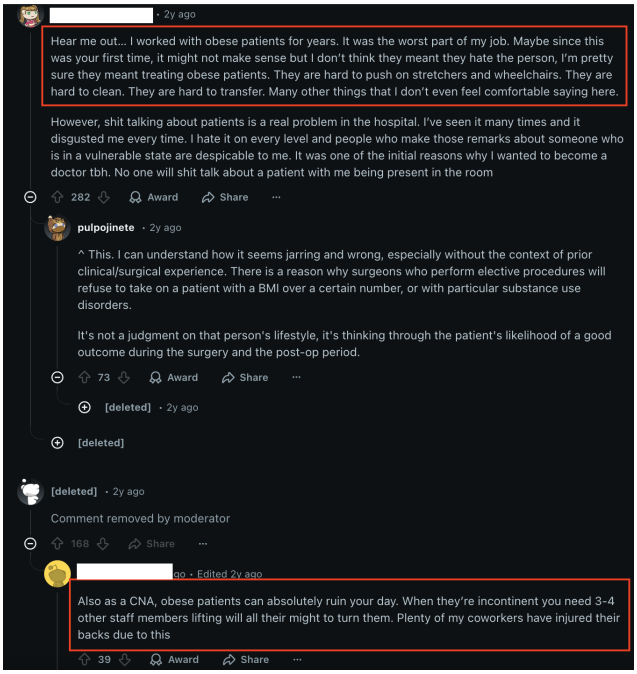
Notes: Screenshots of posts from Reddit subreddits in which patients describe experiences consistent with weight stigma in healthcare settings, including symptom dismissal, withheld diagnostic testing, and moralizing comments from providers. Posts have been cropped and identifying information redacted.



Panel (a)



Panel (b)



Panel (c)

FIGURE A.5 Anecdotal Evidence from Providers Consistent with Weight Stigma

Notes: Screenshots of posts from Reddit subreddits in which providers discuss caring for overweight and obese patients. Posts include language consistent with weight stigma, including biased descriptions of patients and moral judgments about lifestyle. Posts have been cropped and identifying information redacted.

Outcome	**Pred. Mortality**	**Obs. Mortality**		
	(1)	(2)	(3)	(4)
RD estimate (RBC)	0.000 (0.001)	0.001 (0.002)	0.002 (0.002)	0.006** (0.002)
P-value (RBC)	0.747	0.825	0.433	0.017
Mean, left of threshold	0.011	0.013	0.009	0.010
Cutoff	30	24.80	34.50	30
Bandwidth (L/R)	4.216	2.739	5.321	2.684
Observations left	105,253	54,322	137,512	107,909
Observations right	57,056	112,408	29,218	57,430
Effective obs left	42,195	27,959	36,172	25,330
Effective obs right	27,368	31,175	17,080	18,758

TABLE A.1 RD placebo and donut checks

Notes: Each column reports the robust, bias-corrected (RBC) RD estimate, standard error (in parentheses), and RBC p-value using the Calonico, Cattaneo, and Titiunik (2014, 2017) procedure. All specifications use a triangular kernel, local linear polynomial, and patient-clustered standard errors. Col (1): outcome is predicted mortality from a logit of observed mortality on baseline covariates (age, sex, race, insurance, marital status, Elixhauser comorbidities and mortality index); cutoff at BMI = 30. Cols (2)–(3): outcome is observed in-hospital mortality; cutoffs are placebo thresholds at the within-group median BMI below and above 30, respectively. Col (4): donut-hole RD dropping observations within ±0.1 of the obesity cutoff. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

Outcome	RD estimate	p-value	Mean (left)	Bandwidth h	N (L/R)	Eff. N (L/R)
Female	0.025* (0.014)	0.073	0.47	2.659	108,342 / 58,388	25,504 / 19,673
Age at admission	0.149 (0.420)	0.722	61.86	3.906	108,342 / 58,388	39,481 / 26,523
Black	-0.005 (0.009)	0.578	0.15	3.501	108,342 / 58,388	35,368 / 24,539
Medicaid	0.002 (0.009)	0.812	0.17	4.565	107,255 / 58,032	46,302 / 29,325
Medicare	0.003 (0.011)	0.817	0.49	4.870	107,255 / 58,032	49,233 / 30,581
Private insurance	-0.007 (0.011)	0.533	0.32	4.672	107,255 / 58,032	47,457 / 29,731
Married	0.005 (0.012)	0.698	0.50	4.003	106,252 / 57,411	39,829 / 26,612
Hispanic	-0.003 (0.006)	0.620	0.06	4.844	108,342 / 58,388	49,575 / 30,688
Num of Elixhauser comorbidities	0.064 (0.047)	0.173	3.09	5.731	108,323 / 58,365	60,125 / 34,389
Elixhauser mortality index	0.195 (0.214)	0.363	7.46	3.622	108,323 / 58,365	36,641 / 25,035

TABLE A.2 Covariate balance at the obesity cutoff

Notes: Each row reports the robust, bias-corrected (RBC) RD estimate, standard error, and p-value for a pre-determined covariate using the Calonico, Cattaneo, and Titiunik (2014, 2017) procedure at the obesity cutoff (BMI = 30). Mean (left) is the sample mean just below the cutoff. N (L/R) is the total number of observations on each side of the cutoff; Eff. N (L/R) is the effective number of observations within the MSE-optimal bandwidth on each side; h is the MSE-optimal bandwidth. All specifications use a local linear polynomial, triangular kernel, and patient-clustered standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

	Dep Var: In-Hospital Mortality		
	(1)	(2)	(3)
RD estimate (RBC)	-0.002 (0.002)	-0.000 (0.003)	-0.004 (0.003)
P-value (RBC)	0.372	0.889	0.276
Mean, left of threshold	0.013	0.009	0.009
Cutoff BMI	25.0	35.0	40.0
Cutoff label	Overweight	Obese II	Obese III
Bandwidth (L/R)	2.691	5.158	4.860
Observations left	55,922	139,589	154,916
Observations right	110,808	27,141	11,814
Effective obs left	27,344	32,302	14,622
Effective obs right	30,629	15,647	6,609

TABLE A.3 RD estimates of in-hospital mortality at alternative BMI cutoffs

Notes: Each column reports the robust, bias-corrected (RBC) RD estimate, standard error, and p-value using the Calonico, Cattaneo, and Titiunik (2014, 2017) procedure at an alternative BMI cutoff where no "obese" label is assigned: (1) BMI = 25 (overweight), (2) BMI = 35 (obese class II), and (3) BMI = 40 (obese class III). The dependent variable is in-hospital mortality. Mean (left) is the sample mean to the left of the cutoff. All specifications use a local linear polynomial, triangular kernel, MSE-optimal bandwidth, and patient-clustered standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

Panel A: Diagnostic	
Order Type	Order Subtypes
blood bank	blood tests
cardiology	ecg, echo, echo - surface, echo - tee (transesophageal), ett (non-imaging), holter monitor, stress echo
lab	blood gases/whole blood, csf, drug monitoring, Unspecified, urine
neurology	eeg, emg, evoked potential, tcd
ob	fetal monitoring
radiology	angio, cross-sectional interventional radiology, ct scan, general xray, interventional neuro, interventional radiology, mri, noninvasive vascular, nuclear med, ultrasound, Unspecified
respiratory	pulmonary function test
Panel B: Non-Diagnostic	
Order Type	Order Subtypes
adt orders	admit, discharge, postop, transfer
blood bank	cryoprecipitate product order, derivative order, frozen plasma product order, platelet product order, red cell product order
cardiology	cardiac cath, lvad settings
consults	acute care surgery, addiction, anesthesia (chronic pain service), brace and dme orders, cardiology (ep), cardiology (general), cardiology (structural), cardiology (vascular medicine), colorectal surgery, dermatology, diabetes consult, discharge followup appointment, endocrinology, ercp/advanced endoscopy, ethics, gastroenterology, gerontology, gerontology (gifts), hand surgery, hematology-oncology, hepatology, hepatopancreaticobiliary surgery, infectious disease, interventional pulmonology, interventional radiology, lactation, medical procedure service, medicine, medicine procedure service, minimally invasive surgery, nephrology, neurology, neuroscience-icu, neurosurgery, nursing, ob/gyn, occupational therapy, oral maxillofacial surgery (omfs), orthopedics, otolaryngology (ent), palliative care, palliative care/ethics support, pheresis, physical therapy, plastic surgery, podiatry, psychiatry, pulmonary, rheumatology, sleep, social work, speech/swallowing, spiritual care/chaplaincy, surgical oncology, thoracic surgery, transplant surgery, trauma, urology, vascular, wound/ostomy
critical care	cardiac monitoring, hemodynamic monitoring, intra-cranial monitoring, suspected vap pathway
general care	activity, code status, covid-19 precautions de-escalation, dialysis, neuraxial procedure ¡b¿(anesthesia only)¡/b¿, nicu discharge order, other, precautions, restraints, telemetry, telemetry: cardiac, cont spo2, therapeutic devices/dvt/prophylaxis, tubes/drains, use of patient-provided equipment, vitals/monitoring, wound care
hemodialysis	hemodialysis
iv therapy	3cg or radiology tip confirmation ok to use line, access, flushes, iv access, iv access request, iv drip, iv fluids, nicu and non-routine nursery diet order, radiology ok to use line, radiology tip confirmation ok to use line, Unspecified
medications	medication, pain medication/bowel regimen, Unspecified
nutrition	calorie count, diet order, nicu and non-routine nutrition order, npo/diet for procedure, nutrition consult, po fluid restriction, tubefeeding order
ob	cmfm
respiratory	bipap, cpap for osa, extubate, mechanical ventilation, oxygen therapy, respiratory therapy consult, special procedures
tpn	central tpn, ppn, premixed central standard, premixed peripheral standard

TABLE A.5 List of diagnostic and non-diagnostic orders

Notes: This table lists all order type and subtype combinations from the MIMIC-IV Provider Order Entry (POE) table, classified as diagnostic or non-diagnostic. Panel A shows all diagnostic order type–subtype combinations, and Panel B shows all remaining order type–subtype combinations.

	Low-stigma			High-stigma			Diff-in-D		
Characteristics	Estimate	SE(RBC)	P-value	Estimate	SE(RBC)	P-value	Estimate	SE(RBC)	P-value
Female	0.026	0.017	0.122	0.044*	0.023	0.059	0.018	0.026	0.496
Age at admission	-0.064	0.488	0.896	0.389	0.913	0.670	0.453	0.977	0.643
Black	-0.003	0.008	0.728	-0.011	0.012	0.341	-0.009	0.014	0.548
Medicaid	0.014	0.008	0.106	-0.008	0.014	0.575	-0.021*	0.012	0.067
Medicare	0.005	0.013	0.710	-0.006	0.032	0.854	-0.011	0.036	0.765
Private insurance	-0.017**	0.008	0.041	0.007	0.021	0.736	0.024	0.022	0.288
Married	0.003	0.019	0.860	-0.009	0.017	0.605	-0.012	0.030	0.686
Hispanic	0.001	0.004	0.858	-0.003	0.006	0.663	-0.003	0.007	0.622
Num of Elixhauser comorbidities	0.078	0.094	0.406	0.021	0.111	0.849	-0.057	0.191	0.766
Elixhauser mortality index	0.170	0.288	0.555	0.445	0.412	0.280	0.275	0.647	0.671
Predicted Mortality	0.001	0.001	0.232	-0.001	0.001	0.165	-0.002	0.001	0.134

TABLE A.6 Diff-in-Disc covariate balance by provider stigma group

Notes: Each row reports the robust, bias-corrected (RBC) RD estimate, standard error, and p-value for a pre-determined covariate using the Calonico, Cattaneo, and Titiunik (2014, 2017) procedure at the obesity cutoff (BMI = 30). The first three columns report RD estimates among Low-stigma providers (quartiles 1–3); the next three among High-stigma providers (quartile 4); the final three report the Diff-in-Disc estimate (High − Low). All specifications use a local linear polynomial, triangular kernel, and two-way clustering by patient and hospital service. Predicted mortality is generated from a logit of observed mortality on baseline covariates. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

	Dep Var: Any Stigma in Notes				
	(1)	(2)	(3)	(4)	(5)
RD estimate					
Low-stigma providers	-0.036*	-0.032	-0.038*	-0.036*	-0.023*
	(0.019)	(0.027)	(0.020)	(0.019)	(0.014)
High-stigma providers	0.044	0.043*	0.044	0.044	0.019**
	(0.029)	(0.022)	(0.041)	(0.028)	(0.009)
Diff-in-Disc estimate	**0.0800****	**0.0749****	**0.0823***	**0.0800****	**0.0418****
	(0.034)	**(0.038)**	**(0.047)**	**(0.033)**	**(0.017)**
Polynomial	Linear	Quadratic	Linear	Linear	Linear
Covariates	No	No	Yes	No	No
Clustering	Pat., Svc.	Pat., Svc.	Pat., Svc.	Prov., Svc.	Pat., Svc.
Method	RD	RD	RD	RD	OLS
Bandwidth	MSE-opt.	MSE-opt.	MSE-opt.	MSE-opt.	±1.5
Total N	65,788	65,788	65,788	65,788	20,018

TABLE A.7 Robustness of the Diff-in-Disc effect on stigmatizing language

Notes: Each column reports the robust, bias-corrected (RBC) RD estimate and standard error (in parentheses) for Low- and High-stigma providers separately, as well as the Diff-in-Disc estimate (High − Low), using variations of the baseline specification. The dependent variable is an indicator for any stigmatizing language in the discharge summary. Column (1) is the baseline: local linear polynomial, triangular kernel, MSE-optimal bandwidth, and two-way clustering by patient and hospital service. Column (2) uses a local quadratic polynomial. Column (3) adds covariates (age, race, sex, private insurance) to the local linear specification. Column (4) clusters by provider and hospital service instead of patient and hospital service. Column (5) estimates the RD parametrically via OLS using a linear interaction model on a ±1.5 BMI window around the cutoff. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

	Dep Var: In-hospital mortality				
	(1)	(2)	(3)	(4)	(5)
RD estimate					
Low-stigma providers	0.003	0.003	0.003	0.003	-0.000
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
High-stigma providers	0.013***	0.014***	0.013***	0.013***	0.008***
	(0.003)	(0.004)	(0.003)	(0.002)	(0.001)
Diff-in-Disc estimate	**0.0104***	**0.0113***	**0.0106***	**0.0104****	**0.0083*****
	(0.006)	**(0.007)**	**(0.006)**	**(0.005)**	**(0.002)**
Polynomial	Linear	Quadratic	Linear	Linear	Linear
Covariates	No	No	Yes	No	No
Clustering	Pat., Svc.	Pat., Svc.	Pat., Svc.	Prov., Svc.	Pat., Svc.
Method	RD	RD	RD	RD	OLS
Bandwidth	MSE-opt.	MSE-opt.	MSE-opt.	MSE-opt.	±3.0
Total N	166,730	166,730	166,730	166,730	51,630

TABLE A.8 Robustness of the Diff-in-Disc effect on in-hospital mortality

Notes: Each column reports the robust, bias-corrected (RBC) RD estimate (except Column 5) and standard error (in parentheses) for Low- and High-stigma providers separately, as well as the Diff-in-Disc estimate (High − Low), using variations of the baseline specification. The dependent variable is in-hospital mortality. Column (1) is the baseline: local linear polynomial, triangular kernel, MSE-optimal bandwidth, and two-way clustering by patient and hospital service. Column (2) uses a local quadratic polynomial. Column (3) adds covariates (age, race, sex, private insurance) to the local linear specification. Column (4) clusters by provider and hospital service instead of patient and hospital service. Column (5) estimates the RD parametrically via OLS using a linear interaction model on a ±3 BMI window around the cutoff. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

	Leave-One-Out		Obese-Only Mean		Excl. Obese ICD	
	(1)	(2)	(3)	(4)	(5)	(6)
	Stigma	Mortality	Stigma	Mortality	Stigma	Mortality
RD estimate						
Low-stigma providers	-0.035*	0.003	-0.065***	0.004	-0.040*	0.004
	(0.019)	(0.003)	(0.019)	(0.002)	(0.023)	(0.003)
High-stigma providers	0.039	0.013***	0.152***	0.012***	0.041***	0.011***
	(0.029)	(0.003)	(0.021)	(0.003)	(0.012)	(0.003)
Diff-in-Disc estimate	**0.0743****	**0.0106****	**0.2169*****	**0.0081***	**0.0812****	**0.0072**
	(0.034)	**(0.005)**	**(0.019)**	**(0.004)**	**(0.029)**	**(0.005)**
Total N	65,788	166,730	65,788	166,730	59,676	145,499

TABLE A.9 Robustness of Diff-in-Disc estimates to alternative provider classification methods

Notes: Each column reports the robust, bias-corrected (RBC) RD estimate, standard error (in parentheses), and Diff-in-Disc estimate for Low- and High-stigma providers using three alternative methods of classifying provider type. Columns (1)–(2) use a leave-one-out measure, where each patient's provider stigma rate is calculated across all other patients of that provider, excluding the focal patient. Columns (3)–(4) classify providers by their mean stigma rate among obese patients only (BMI $\geq$ 30). Columns (5)–(6) exclude patients with an obese ICD diagnosis code from both the provider classification and the estimation sample. In odd-numbered columns the dependent variable is an indicator for any stigmatizing language in the discharge summary; in even-numbered columns it is in-hospital mortality. The Diff-in-Disc estimate is the difference in the RD effect between High- and Low-stigma providers. All specifications use a local linear polynomial, triangular kernel, MSE-optimal bandwidth, and two-way clustering by patient and hospital service. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

	estimate	SE (RBC)	p-value	total N	Eff. N (L)	Eff. N (R)	h
Outcome: Any Stigma in Notes							
RD estimate by height quartile							
Height Q1 (lowest)	0.023	0.028	0.410	17,038	3,150	2,875	1.759
Height Q2	-0.031	0.052	0.548	15,657	2,809	2,403	1.681
Height Q3	0.001	0.040	0.971	15,635	2,231	1,942	1.362
Height Q4 (highest)	-0.065	0.046	0.162	15,888	2,306	2,039	1.302
Diff-in-Disc (Q4 − Q1)	**-0.0877***	**0.046**	**0.058**	**65,788**			

TABLE A.10 Heterogeneous RD effects on stigmatizing language by height quartile

Notes: Each row reports the robust, bias-corrected (RBC) RD estimate, standard error, and p-value using the Calonico, Cattaneo, and Titiunik (2014, 2017) procedure at the obesity cutoff (BMI = 30). The outcome is an indicator for any stigmatizing language in the discharge summary. The sample is split into quartiles of patient height (Q1 = shortest, Q4 = tallest). The Diff-in-Disc estimate is the difference in the RD effect between the highest and lowest height quartiles (i.e., Q4 − Q1). All specifications use a local linear polynomial, triangular kernel, and two-way clustering by patient and hospital service. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

	Dep. var.: $\Delta Stigma_{t-(t-1)}$
	(1)
$\Delta BMI_{t-(t-1)}$	0.030**
	(0.011)
p-value	0.008
Observations	18,322
Patient FE	Yes
Provider FE	Yes
Clustered SE	patient, provider

TABLE A.11 First-difference: changes in stigmatizing language on changes in BMI

Notes: The table reports a first-difference regression of $\Delta Stigma_{t-(t-1)}$ on $\Delta BMI_{t-(t-1)}$, where each observation is a within-patient change across consecutive hospital encounters. The dependent variable is the change in an indicator for any stigmatizing language in the discharge summary; the independent variable is the change in BMI. Standard errors (in parentheses) are two-way clustered by patient and provider. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

Outcome	estimate	SE (RBC)	p-value	total N	Eff. N (L)	Eff. N (R)	h	mean (sample)
Word count	-28.092**	10.523	0.008	133,318	34,771	21,896	4.272	1674.57
Character count	-186.768**	68.711	0.007	133,318	33,585	21,599	4.197	11096.16
Average word length	-0.002	0.005	0.644	133,318	34,357	21,818	4.208	5.39

TABLE A.12 Effect of the obesity cutoff on discharge-note length

Notes: Each row reports the robust, bias-corrected (RBC) RD estimate, standard error, and p-value for a text-based measure of discharge-summary length using the Calonico, Cattaneo, and Titiunik (2014, 2017) procedure at the obesity cutoff (BMI = 30). Outcomes are word count, character count, and average word length. Mean (sample) is the sample average of the outcome. Unlike the main specifications, the sample includes all patients with a discharge summary and a recorded BMI (not restricted to the overweight and obese range). All specifications use a local linear polynomial, triangular kernel, and two-way clustering by patient and hospital service. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.

	Dep Var: In-Hospital Mortality				
	(1)	(2)	(3)	(4)	(5)
RD estimate (RBC)	-0.0030 (0.0061)	-0.0051 (0.0059)	-0.0144** (0.0054)	-0.0024 (0.0051)	-0.0060 (0.0068)
P-value (RBC)	0.620	0.392	0.008	0.631	0.377
Mean, left of threshold	0.025	0.024	0.019	0.015	0.008
Cutoff information					
BP category	Elevated	Prehypertension	Stage 1	Stage 2	Crisis
Syst / Diast threshold	120/–	120/80	130/80	140/90	180/120
Logic	Syst only	OR	OR	OR	OR
Classification	AHA/ACC	JNC-7	AHA/ACC	AHA/ACC	AHA/ACC
Bandwidth (h) (L/R)	0.621	0.536	0.565	0.675	0.996
Sample					
Observations left	11,336	11,336	15,619	22,860	31,474
Observations right	10,155	20,792	16,509	9,268	654
Effective obs left	5,981	5,643	6,782	7,824	1,699
Effective obs right	5,713	8,111	7,948	5,934	502

TABLE A.13 Placebo RD: In-hospital mortality at hypertension blood pressure cutoffs

Notes: Each column reports the robust, bias-corrected (RBC) RD estimate of the effect of crossing a hypertension threshold on in-hospital mortality. Column (1) estimates the RD at the AHA/ACC "Elevated" boundary (systolic $\geq$120 mmHg), restricting the sample to patients with diastolic blood pressure below 80 mmHg, since the Elevated category is defined solely by systolic pressure conditional on normal diastolic pressure. Columns (2)–(5) use a binding-score running variable: the maximum of the standardized distances of systolic and diastolic blood pressure from their respective cutoffs, implementing OR logic (either dimension binding). Column (2) uses the JNC-7 "Prehypertension" cutoff (systolic $\geq$120 *or* diastolic $\geq$80 mmHg). Columns (3)–(5) use the AHA/ACC 2017 classification: Stage 1 (130/80), Stage 2 (140/90), and Hypertensive Crisis (180/120). Blood pressure is measured as the first recorded value during the hospitalization. All specifications use a local linear polynomial, triangular kernel, MSE-optimal bandwidth, and patient-clustered standard errors. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$.