



# Modeling Mistrust in End-of-Life Care

Willie Boag, Harini Suresh,

Leo Anthony Celi, Peter Szolovits, Marzyeh Ghassemi

[github.com/wboag/eol-mistrust](https://github.com/wboag/eol-mistrust)



## Racial Disparities in End-of-Life Care

Yarnell et al. (2017), Muni et al. (2011), and Lee et al. (2016) observed racial disparities in aggressive end-of-life care.

**We replicate this in both MIMIC and eICU: black patients received (sometimes statistically significantly) longer durations of aggressive EOL treatment.**

Figure 3-1: **Mechanical Ventilation:** CDF of ventilation duration by race, where dotted lines represent the median duration treatment for a population. In multiple datasets, the median black patient receives statistically significant longer ventilation durations than the median white patient.

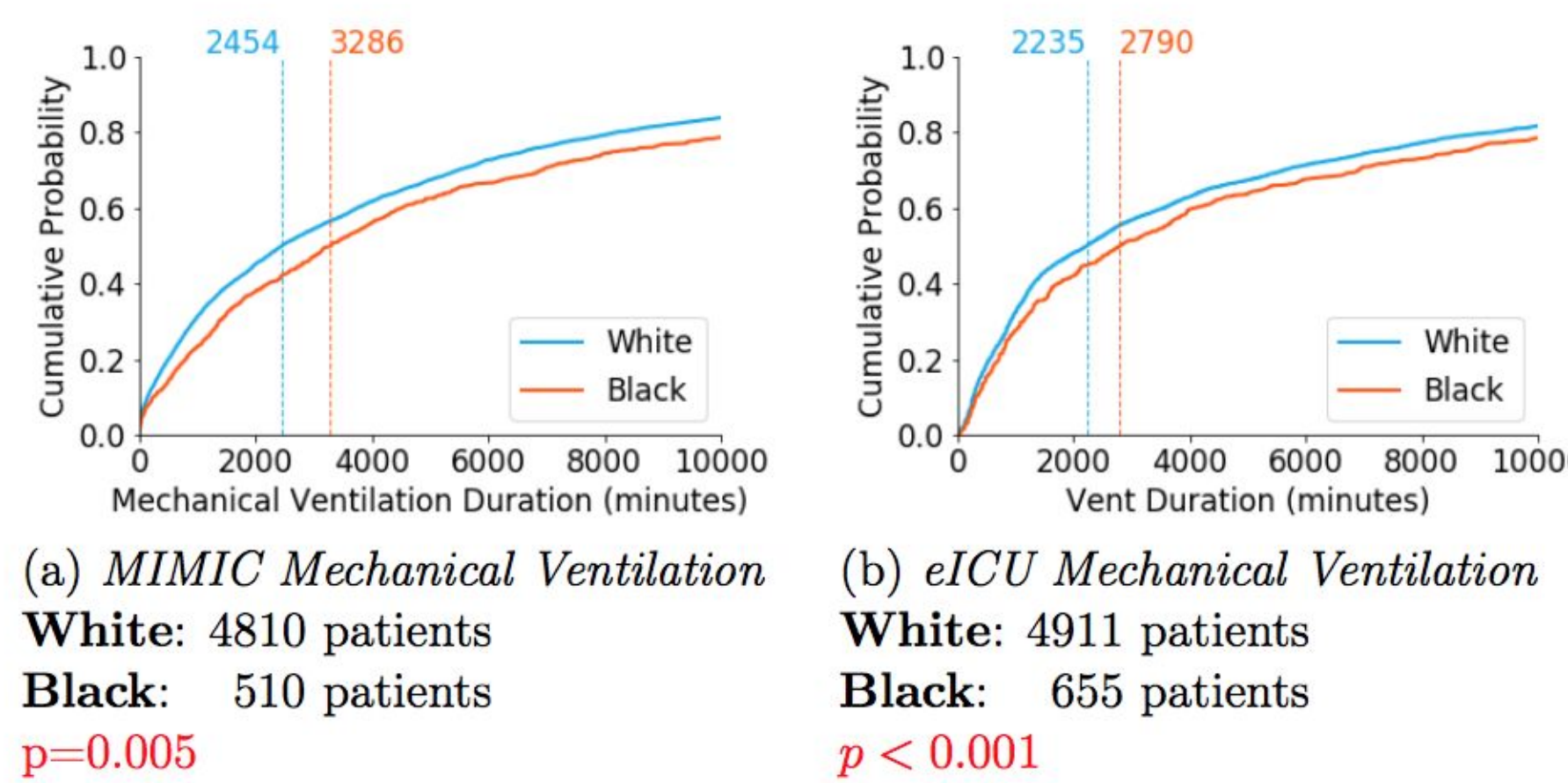
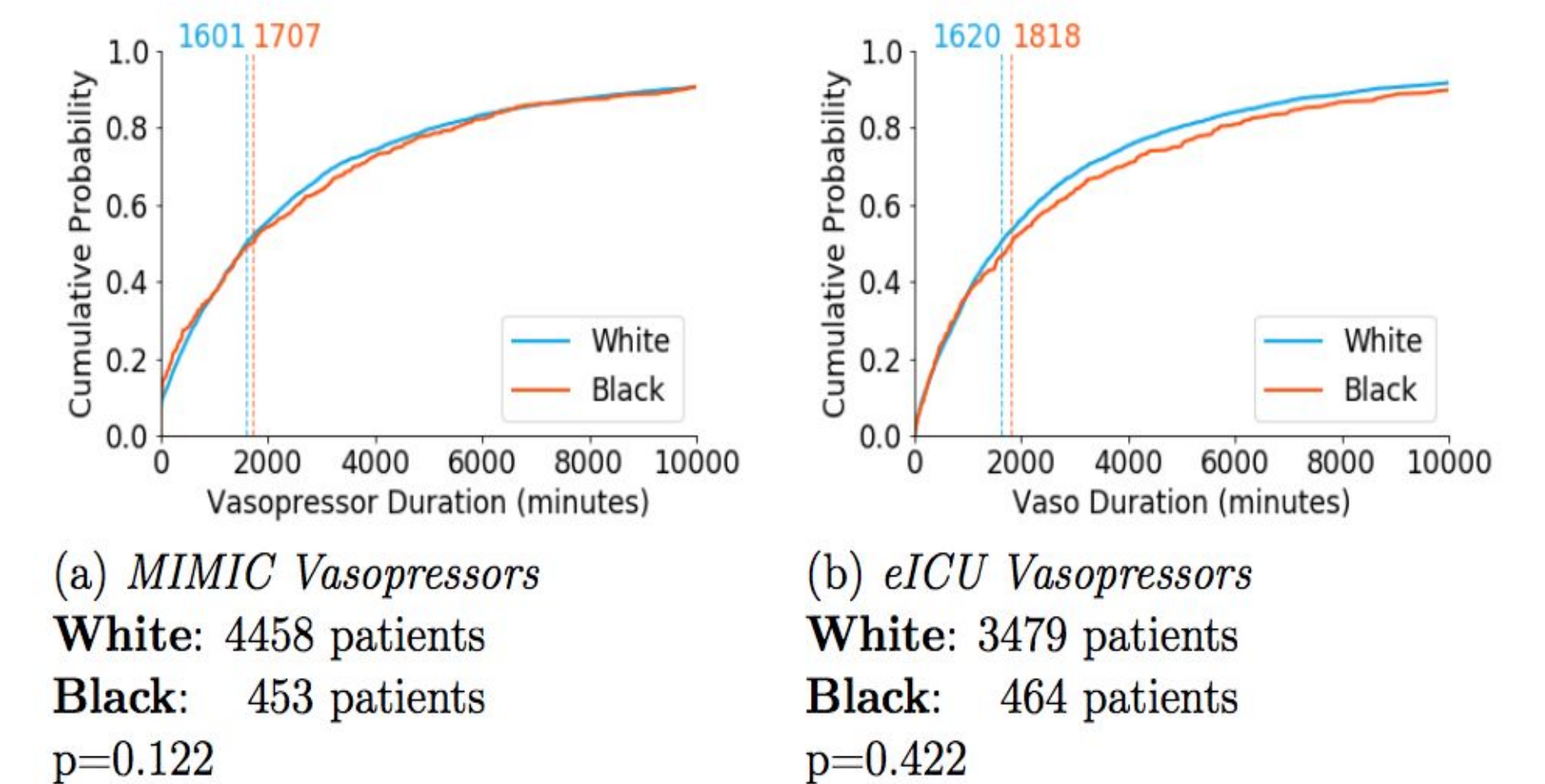


Figure 3-2: **Vasopressors:** In both datasets, the median black patient receives a longer duration of vasopressors than the median white patient. This trend is not statistically significant in either dataset.



## Quantifying Mistrust

Figure 4-1: An example of a nursing note documenting mistrust (in red). Situation-specific identifying information has been blacked out.

# Social: Pt refused to sign ICU consent and expressed wishes to be DNR/DNI, seemingly very frustrated and mistrusting of healthcare system in relation to [redacted]. Also, w/ hx of poor medication compliance and follow-up.

Table 4.1: Coded interpersonal feature types from chartevents.

1:1 sitter present?	baseline pain level (0 to 10)	received bath?	bedside observer
behavioral intent	currently experiencing pain	disease state	consults
education barrier	education learner	education method	family meeting?
education readiness	harm by partner?	education topic	judgement
follows commands?	family communication method	gcs - verbal response	informed?
hair washed?	goal richmond-ras scale	headache?	health care proxy?
pain management	non-violent restraints?	orientation	pain (0 to 10)
pain assess method	understand & agree with plan?	pain level acceptable?	reason for restraint
restraint device	richmond-ras scale (-5 to +4)	rsbi deferred	riker-sas scale
safety measures	violent restraints ordered?	security	security guard
side rails	status and comfort	sitter	skin care?
spiritual support	behavior during application	support systems	stress
verbal response	teaching directed toward	wrist restraints?	social work consult?

Table 4.2: Top-3 most positively and negatively informative chartevent features for tuning the mistrust metric.

feature	weight
state: alert	-1.0156
riker-sas scale: agitated	0.7013
pain: none	-0.5427
richmond-ras scale: 0 alert and calm	-0.3598
education readiness: no	0.2540
pain level: 7-mod to severe	0.2168

We use an L1-regularized binary logistic regression to predict NONCOMPLIANCE (labeled using regex search of notes) with chartevents binary indicators as features.

The classifier's probability serves as the mistrust score.

**The median black patient has a statistically significantly higher mistrust score than the median white patient.**

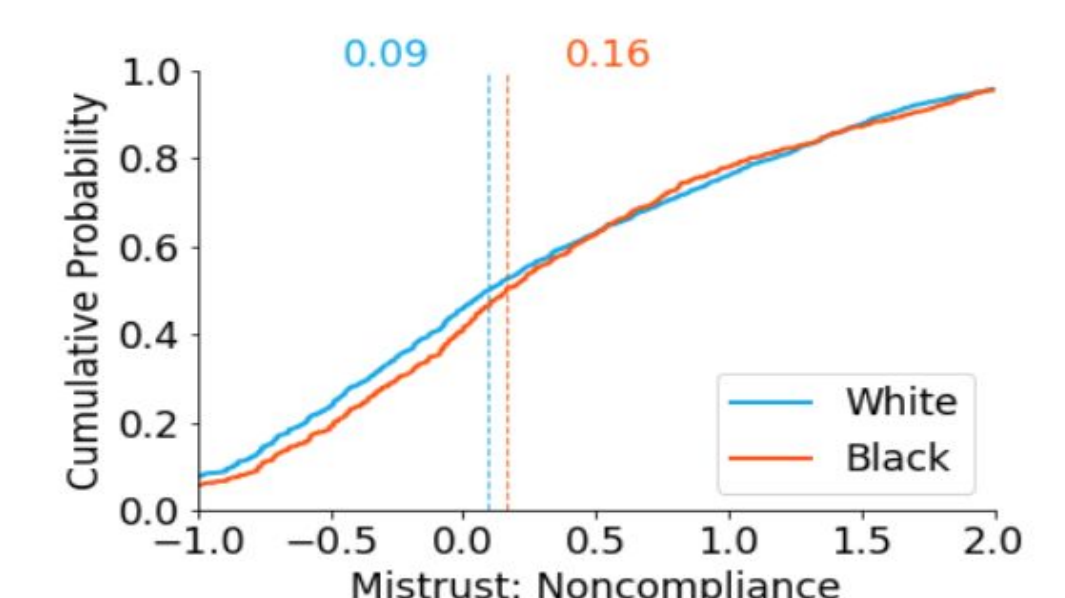


Figure 4-2: Racial disparity in noncompliance-derived mistrust metric. White: 9923 patients, Black: 1202 patients,  $p < 0.001$

## Mistrust-Based Disparities

**Mistrust-based treatment disparities are starker than race-based treatment disparities.**

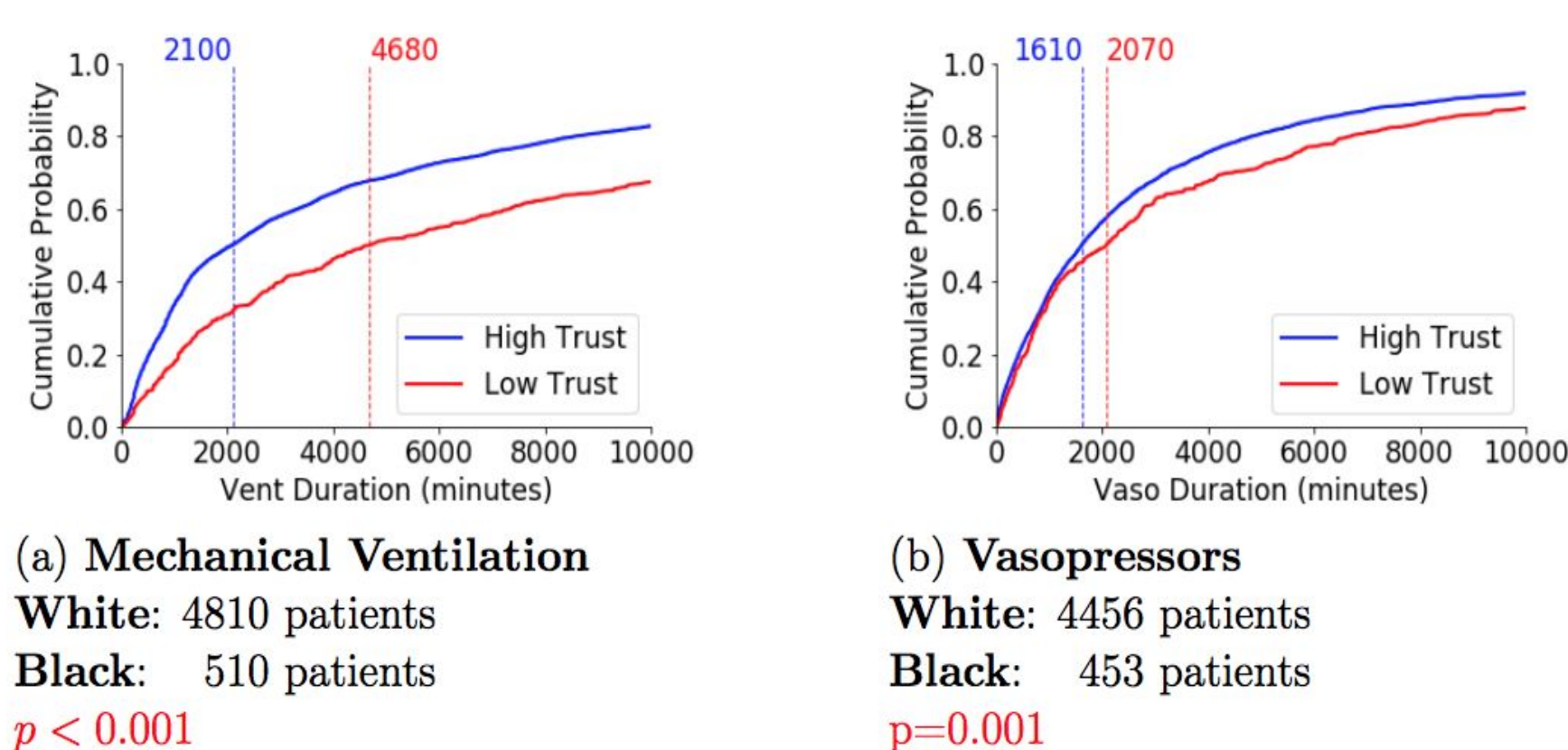


Table 6.2: Effect of race and mistrust features on various binary classification tasks. Performance is measured by AUC and averaged over 100 runs.

Features	Left AMA (n=48,071)	Code Status (n=39,815)	In-Hospital Mortality (n=48,071)
Baseline	0.859 ± .014	0.763 ± .013	0.600 ± .011
Baseline + Race	0.861 ± .014	0.766 ± .014	0.614 ± .011
Baseline + Noncompliant	<b>0.869</b> ± .012	0.767 ± .013	0.614 ± .010
Baseline + Autopsy	0.861 ± .012	<b>0.773</b> ± .011	0.603 ± .012
Baseline + Negative-Sentiment	0.859 ± .013	0.765 ± .014	<b>0.615</b> ± .010
Baseline + ALL	<b>0.873</b> ± .012	<b>0.782</b> ± .012	<b>0.635</b> ± .010

Using various mistrust metrics as features on top of demographics-derived baseline improves performance for clinical tasks.

Table 3. Pairwise Pearson correlations between severity scores and mistrust score.

	OASIS	SAPS II	Mistrust
OASIS	1.0	0.680	0.095
SAPS II	0.680	1.0	0.045
Mistrust	0.095	0.045	1.0

Acuity scores (OASIS and SAPS II) have high correlation with each other, but low correlation with the mistrust metric.

The metric is not just capturing severity of illness.

## Acknowledgements

This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. 1122374. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.