Columbia University

reliable Artificial Intelligence for Health and Medicine (reAIM) Lab

Neural Information Processing Systems

Path-specific effects for pulse-oximetry guided decisions in critical care

NeurIPS 2025

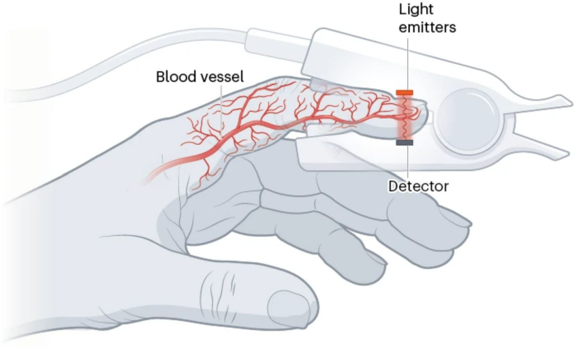
Kevin Zhang, Yonghan Jung, Divyat Mahajan, Karthikeyan Shanmugam, Shalmali Joshi

Evidence of racial bias in pulse oximeter device measurements

nature

NEWS & VIEWS FORUM | 19 October 2022

Skin colour affects the accuracy of medical oxygen sensors



npr

COVID-19 made pulse oximeters ubiquitous. Engineers are fixing their racial bias

Original Investigation | Pulmonary Medicine

Analysis of Discrepancies Between Pulse Oximetry and Arterial Oxygen Saturation Measurements by Race and Ethnicity and Association With Organ Dysfunction and Mortality

An-Kwok Ian Wong, MD, PhD[1,2]; Marie Charpignon, MS[3]; Han Kim, MSE[4] ; et al

Original Investigation

Assessment of Racial and Ethnic Differences in Oxygen Supplementation Among Patients in the Intensive Care Unit

Eric Raphael Gottlieb, MD, MS[1,2,3]; Jennifer Ziegler, MD, MSc[4]; Katharine Morley, MD, MPH[2,5] ; et al

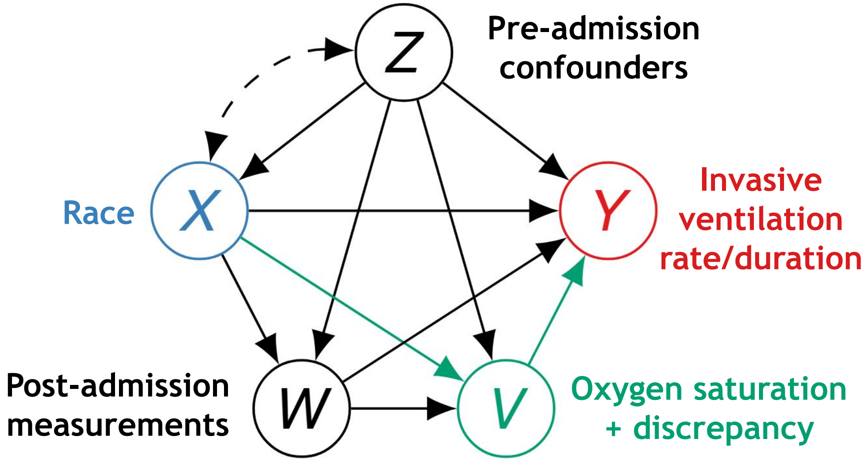
Original Investigation

Racial and Ethnic Discrepancy in Pulse Oximetry and Delayed Identification of Treatment Eligibility Among Patients With COVID-19

Ashraf Fawzy, MD, MPH[1]; Tianshi David Wu, MD, MHS[2,3]; Kunbo Wang, MS[4] ; et al

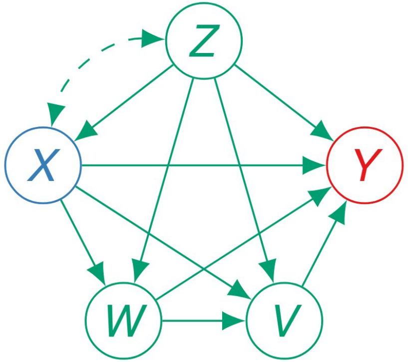
Quantifying *path-specific causal effect* of *pulse-oximeter discrepancies* on clinical decision-making in the ICU

1. Extend existing causal mediation analysis for multiple mediators
2. Develop robust estimators with theoretical guarantees and better performance in small-sample regimes (e.g. ICU data)



Study	Causal Analysis	Path-Spec. Analysis	Multi-Med.	ICU Data	Finite Guar.
Sjoding, 2020	Assoc.	×	×	✓	×
Wong, 2021	Assoc.	×	×	✓	×
Gottlieb, 2022	Part.	Part.	×	✓	×
Miles, 2017	✓	✓	✓	×	×
Tchetgen, 2012	✓	✓	×	×	×
VanderWeele, 2014	✓	✓	×	×	×
Ours	✓	✓	✓	✓	✓

Causal fairness model with two mediators

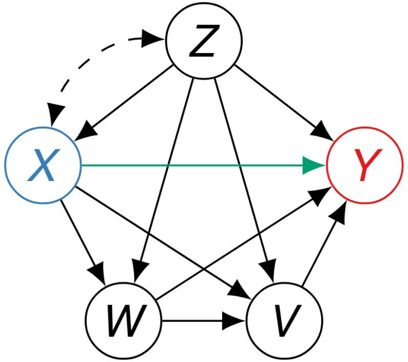


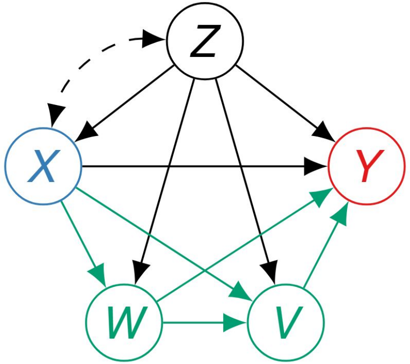
Total Effect (TE)

$$\mathbb{E}[Y_{x_1}] - \mathbb{E}[Y_{x_0}]$$

$$\mathbb{E}[Y_{x_1, W_{x_0}, V_{x_0}}] - \mathbb{E}[Y_{x_0}]$$

Natural Direct Effect (NDE)



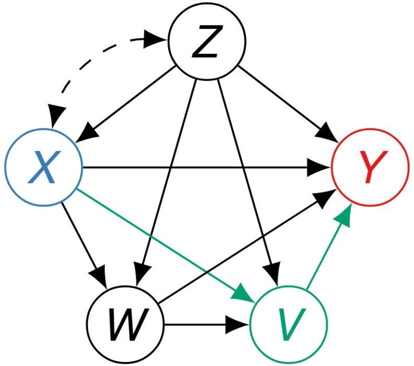


Natural Indirect Effect (NIE)

$$\mathbb{E}[Y_{x_1}] - \mathbb{E}[Y_{x_1, W_{x_0}, V_{x_0}}]$$

$$\mathbb{E}[Y_{x_1}] - \mathbb{E}[Y_{x_1, V_{x_0}, W_{x_1}}]$$

V-specific Direct Effect (VDE)



Pearl, *UAI*, 2001

Pulse oximetry application

VDE doubly robust estimator

The VDE is identifiable using the back-door adjustment and given by

$$\mathbb{E}[Y_{x_1, V_{x_0}, W_{x_1}}] = \sum_{w,v,z} \mathbb{E}[Y \mid x_1, w, v, z] P(v \mid x_0, w, z) P(w \mid x_1, z) P(z). \quad (1)$$

We use the doubly robust estimator,

$$\begin{aligned} \varphi((Y, V, W, X, Z); \boldsymbol{\mu}_0, \boldsymbol{\pi}_0) \triangleq\ & \pi_0^3(V, W, X, Z)\{Y - \mu_0^3(V, W, X, Z)\} \\ & + \pi_0^2(W, X, Z)\{\mu_0^3(V, W, x_1, Z) - \mu_0^2(W, X, Z)\} \\ & + \pi_0^1(X, Z)\{\mu_0^2(W, x_0, Z) - \mu_0^1(X, Z)\} + \mu_0^1(x_1, Z). \end{aligned}$$

$$\mathbb{E}[\varphi((Y, V, W, X, Z); \boldsymbol{\mu}_0, \boldsymbol{\pi}_0)] = \text{Eq. (1)}$$

Nuisance parameters defined following Jung, *NeurIPS*, 2024

Theoretical guarantees of the VDE estimator

Algorithm

1. (***Sample-Splitting***)
   Take any L-fold partition of the data $\mathcal{D} = \cup_{\ell=1}^{L} \mathcal{D}_\ell$.

2. (***Learning by Partitions***)
   Learn $\hat{\mu}_\ell^i, \hat{\pi}_\ell^i$ *using* $\mathcal{D} \setminus \mathcal{D}_\ell$ *and compute* $\hat{\psi}_\ell \triangleq \mathbb{E}_{\mathcal{D}_\ell}[\varphi(Y, V, W, X, Z); \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\pi}}]$.

3. (***Aggregation***)
   Construct $\hat{\psi} \triangleq \frac{1}{L}(\hat{\psi}_1 + \cdots, \hat{\psi}_L)$.

Theorem 2 (informal)

Suppose the second and third moments of φ *exist.*

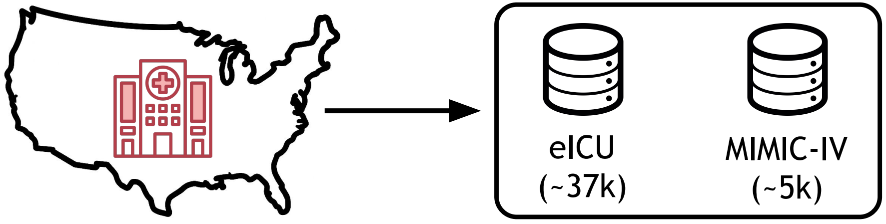
Let $R_1 \triangleq \frac{1}{L}\sum_{\ell=1}^{L}(\mathbb{E}_{\mathcal{D}_\ell}[\hat{\varphi}_\ell] - \mathbb{E}_P[\varphi_0])$. *Then,*

$$\hat{\psi} - \psi_0 = R_1 + \frac{1}{L}\sum_{\ell=1}^{L}\sum_{i=1}^{3}\mathbb{E}[\{\hat{\mu}_\ell^i - \mu_0^i\}\{\pi_0^i - \hat{\pi}_\ell^i\}].$$

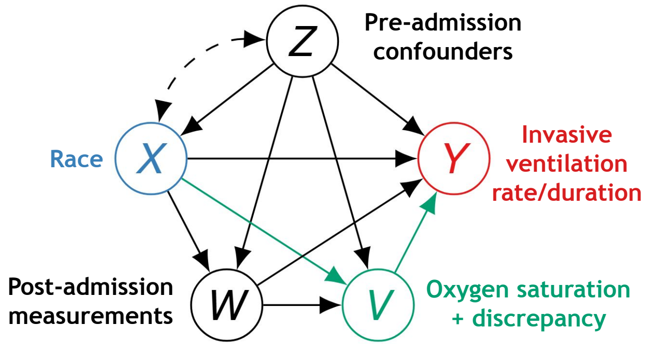
Moreover, $R_1 \xrightarrow{p} 0$ *and* R_1 *is asymptotically normal.*

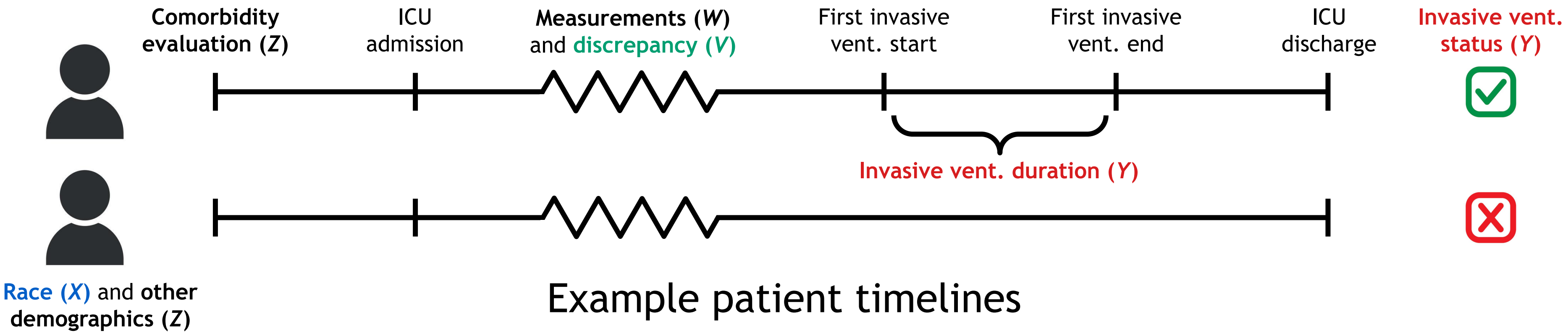
Since $\mathbb{E}[\pi_0^i] = 1$, we use the self-normalized (SN) estimator, $\hat{\pi}_{\text{SN}}^i \leftarrow \hat{\pi}^i / \mathbb{E}_{\mathcal{D}}[\hat{\pi}^i]$, as this improves the estimation stability in practice.

Critical care data on invasive ventilation-related outcomes



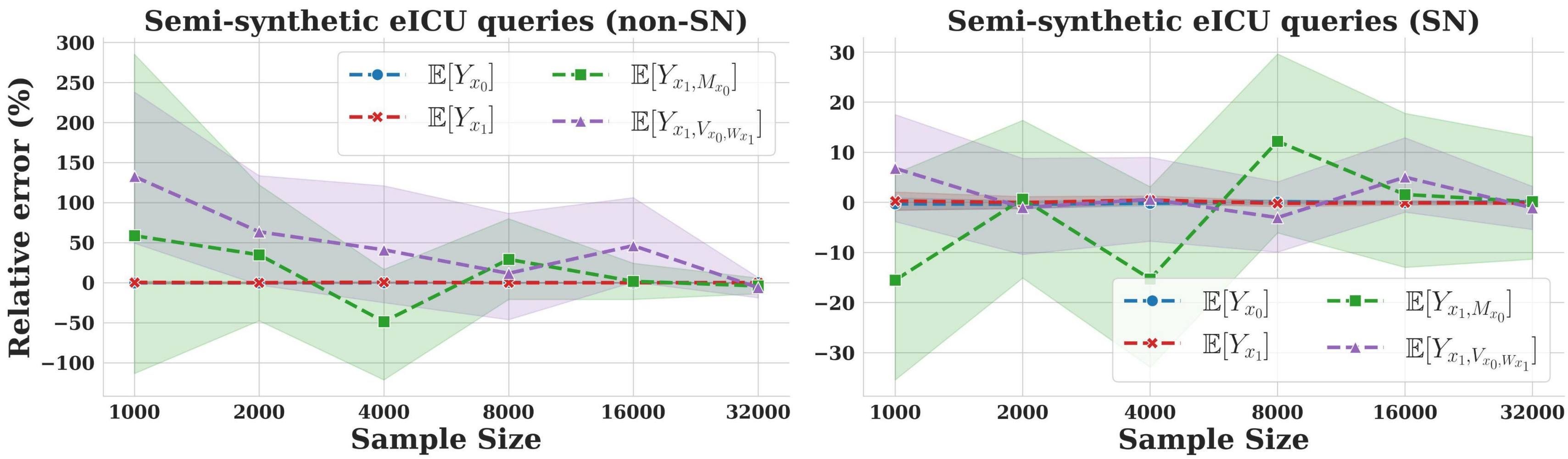
EWA to summarize measurement trajectories





Example patient timelines

Semi-synthetic experiments

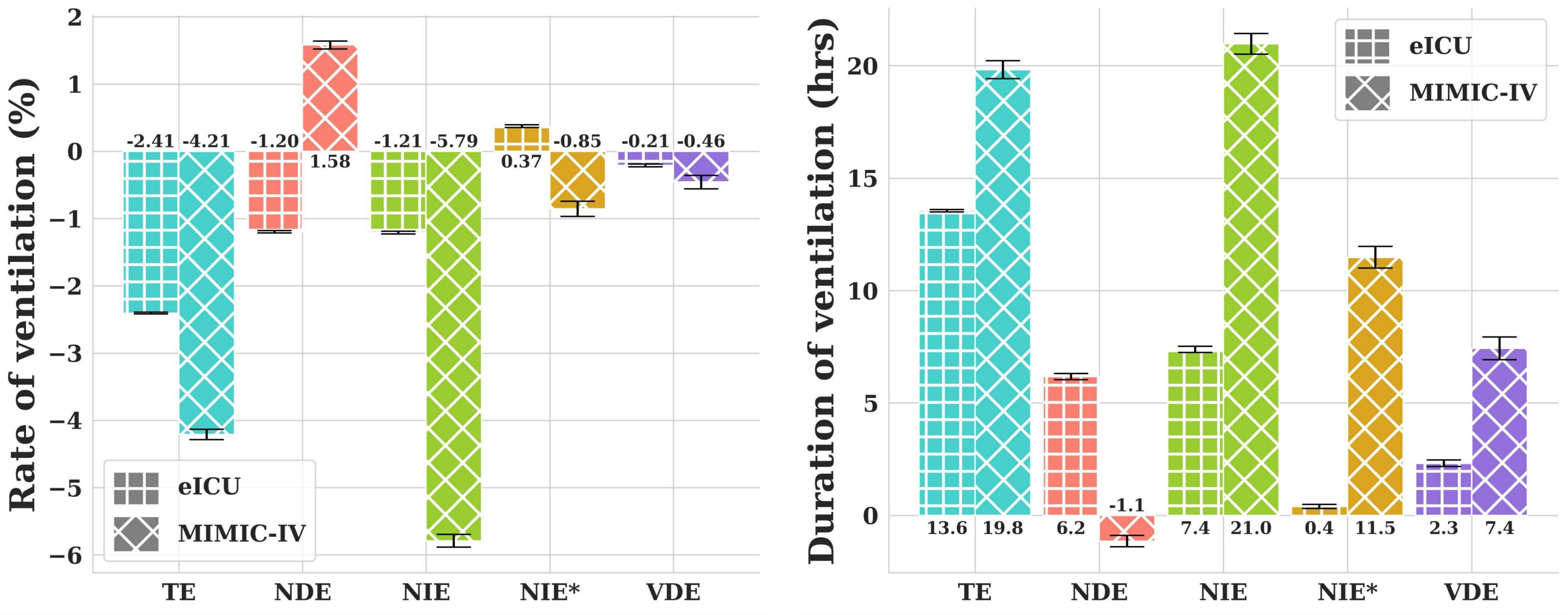


Bands show 95% CIs

The self-normalized estimator (right) has smaller variance compared to the canonical doubly robust estimator (left)

Real-world experiments

Positive effect = more frequent or longer treatment for Black patients vs. White patients



Error bars show 95% CIs

Ventilation rate VDE is negligible (95% CI **eICU [-0.23, -0.19]**, **MIMIC-IV [-0.56, -0.36]**), but more pronounced for ventilation duration (95% CI **eICU [2.2, 2.5]**, **MIMIC-IV [6.9, 7.9]**)

Conclusions and future work

- We examine heterogeneity by race in a clinically actionable healthcare process using path-specific causal analysis with theoretical guarantees

- Bias primarily manifests in ventilation duration, not in treatment initiation

- Future work could model temporality explicitly

In-person poster session @ San Diego

Thu 4 Dec, 11am-2pm PST, Exhibit Hall C/D/E

References

Ashraf Fawzy, Tianshi David Wu, Kunbo Wang, et al. Racial and ethnic discrepancy in pulse oximetry and delayed identification of treatment eligibility among patients with covid-19. *JAMA Internal Medicine*, 182(7):730–738, 2022.

Eric Raphael Gottlieb, Jennifer Ziegler, Katharine Morley, et al. Assessment of racial and ethnic differences in oxygen supplementation among patients in the intensive care unit. *JAMA Internal Medicine*, 182(8):849–858, 2022.

Yonghan Jung, Jin Tian, and Elias Bareinboim. Unified covariate adjustment for causal inference. In *Advances in Neural Information Processing Systems*, volume 37, pages 6448–6499, 2024.

Caleb H Miles, Ilya Shpitser, Phyllis Kanki, et al. On semiparametric estimation of a path-specific effect in the presence of mediator-outcome confounding. *Biometrika*, 107:159–172, 2017.

Judea Pearl. Direct and indirect effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, pages 411-420, 2001.

Michael W Sjoding, Robert P Dickson, Theodore J Iwashyna, et al. Racial bias in pulse oximetry measurement. *The New England Journal of Medicine*, 383 (25):2477–2478, 2020.

Eric Tchetgen Tchetgen and Ilya Shpitser. Semiparametric theory for causal mediation analysis: efficiency bounds, multiple robustness, and sensitivity analysis. *Annals of Statistics*, 40:1816–1845, 2012.

Tyler VanderWeele and Stijn Vansteelandt. Mediation analysis with multiple mediators. *Epidemiologic Methods*, 2(1):95–115, 2014.

An-Kwok Ian Wong, Marie Charpignon, Han Kim, et al. Analysis of discrepancies between pulse oximetry and arterial oxygen saturation measurements by race and ethnicity and association with organ dysfunction and mortality. *JAMA Network Open*, 2021.