Using Racially Biased Proxy as Outcomes for Prediction Models – Are We (Re)producing Health Inequalities?

THE CHOICE INSTITUTE
School of Pharmacy

Sara Khor[1], Anirban Basu[1], Erin E Hahn[2], Eric C Haupt[2], Lindsay Joe L Lyons[2], Connor P Henry[1], Veena Shankaran[3], Patrick J Heagerty[1], Aasthaa Bansal[1]

[1]University of Washington, Seattle, WA, USA, [2]Southern California Permanente Medical Group, Pasadena, CA, USA, [3]Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Proxy Outcomes may be Racially Biased

- Cancer recurrence is an important outcome for clinicians and patients, but not consistently captured in structured fields of electronic health record or claims data
- We often rely on proxies for cancer recurrence that are based on healthcare utilization patterns and diagnoses codes
- Healthcare utilization and timing of diagnoses often varies systematically across racial subgroups
- Proxy outcomes for recurrence that are based on utilization and diagnosis patterns can be racially biased

RESEACH GOALS

To examine the racial bias in using a utilization-based proxy for colorectal cancer (CRC) recurrence

Methods

- **Data:** Kaiser Permanente Southern California, patients diagnosed with Stage 1-3 CRC and underwent resection between 2008-2013
- **Sample:** Random sample, stratified by race/ethnicity and recurrence status (n=276)
- **Outcome:** 5-year recurrence status post cancer resection, accounting for censoring
- **Proxy for recurrence**: based on utilization and timing of health services (i.e.., surgery, chemo, imaging, and hospice referral) associated with metastatic disease
- **Analysis:** Compare chart review recurrence status (gold standard) vs. proxy recurrence status based on utilization patterns
  - Bootstrap confidence intervals
  - Qualitatively summarize patterns

KEY TAKEAWAYS

- Recurrence rates calculated using proxies based on utilization and diagnosis patterns may be **disproportionately inflated for Hispanic and Black/African American patients**
  - Significantly higher false positives among Hispanics

IMPLICATIONS

- Using a racially biased proxy outcome to estimate disease burden for population health management or for outcomes research may **produce inaccurate estimates for racial/ethnic minority groups**
- Building risk prediction models using an outcome label that is racially biased may **overestimate health needs for Hispanic/Black patients and lead to over-testing, higher anxiety, and higher healthcare costs for these individuals**

Funding Acknowledgement: This project was supported by National Cancer Institute (1R37CA218413-01A1) and the Health Technology Fund.

For more Information, please contact khors@uw.edu

Key Findings

Table 1. Performance of the proxy recurrence outcome

	Overall (n=276)	Non-Hispanic White (n =85)	Black/African American (n=85)	Hispanic (n=84)
True recurrence rate	40%	33%	47%	**42%**
Utilization-based recurrence rate	43%	32%	51%	**50%**
PPV	90%	96%	91%	81%*
NPV	98%	97%	98%	98%

*PPV is 15% lower among Hispanics compared to NHWs (95% 0.7-29%)

Figure 1. Comparing cumulative incidences of recurrence

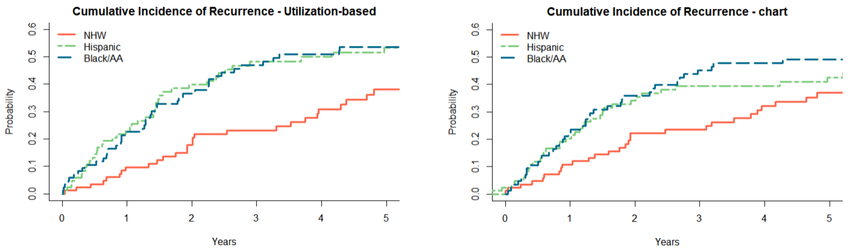
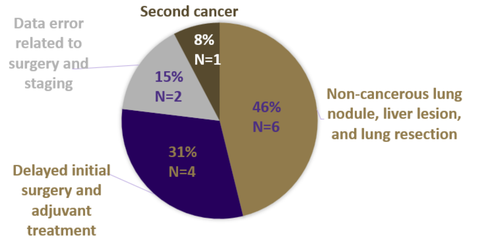


Figure 2. Reasons for False Positives (n=13)



Limitations

- Small sample size limited our ability to examine patterns of false positives among racial/ethnic minorities