Interpreting the Effectiveness of Cancer Screening from National Population Statistics

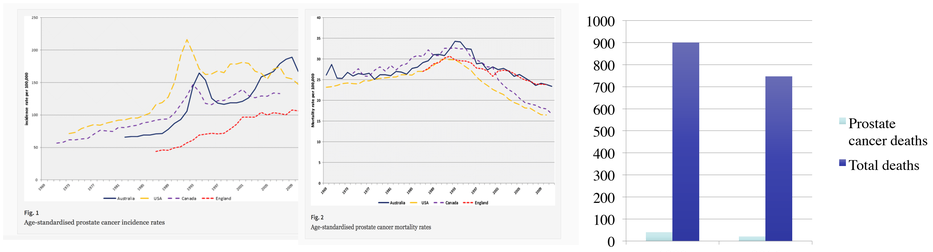
Joaquin Chapa MD[1], Alyson Haslam, PhD, Vinay Prasad, MD MPH[2-4]

1. Department of Medicine **2. Division of Hematology-Oncology, Knight Cancer Institute 3. Department of Public Health and Preventive Medicine 4. Center for Health Care Ethics**

Lessons from Prostate Cancer: Conflicting evidence sources; conflicting guidelines

- RCTs of PSA screening have shown either no benefit or marginal benefit on disease-specific survival, and no benefit an overall survival

- Discordant with this fact: prostate cancer mortality has fallen



Prostate cancer incidence and mortality over time. Source: Feletto, E., Bang, A., Cole-Clark, D. et al. World J Urol (2015)

- The USPSTF does not consider observational data in its guidelines, the Urologic society and American Cancer Society do: observational and modeling data are cited as providing circumstantial causal evidence supporting the efficacy of screening

- Society guidelines simply assume that population-level data is accurate and consistent enough to support efficacy claims of a marginal intervention over time

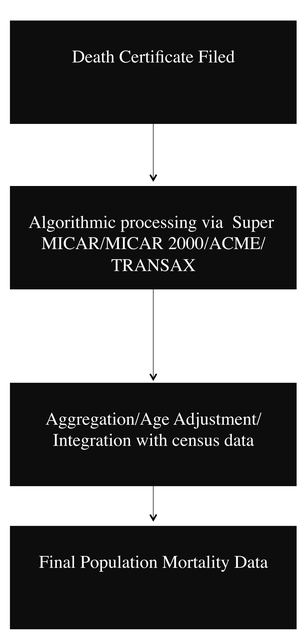
- To investigate the appropriateness of these claims, we investigated the source inputs and processes by which prostate cancer mortality data is constructed

Sausage-making: How Mortality Data is Created

- A patient dies; a provider completes a death certificate

- Natural language on death certificate is converted into ICD-10 codes
- Mortality Medical Data System: Cause of death determined via an algorithm of hierarchical causal events
- *"The injury that initiated the morbid train of events leading to death"*

- Data is aggregated, age adjusted and integrated with US Census data



<u>Differences between RCT and population-level data processes:</u>

- **<u>Population Data</u>: Algorithm determines cause of death without review**

- **<u>PLCO</u>: Up to 3 human reviewers if discordance with death certificate**

- **<u>ESRPC</u>: 3-person review incorporating imaging and clinical data**

Sources of Error, Noise and Bias Across Data Levels

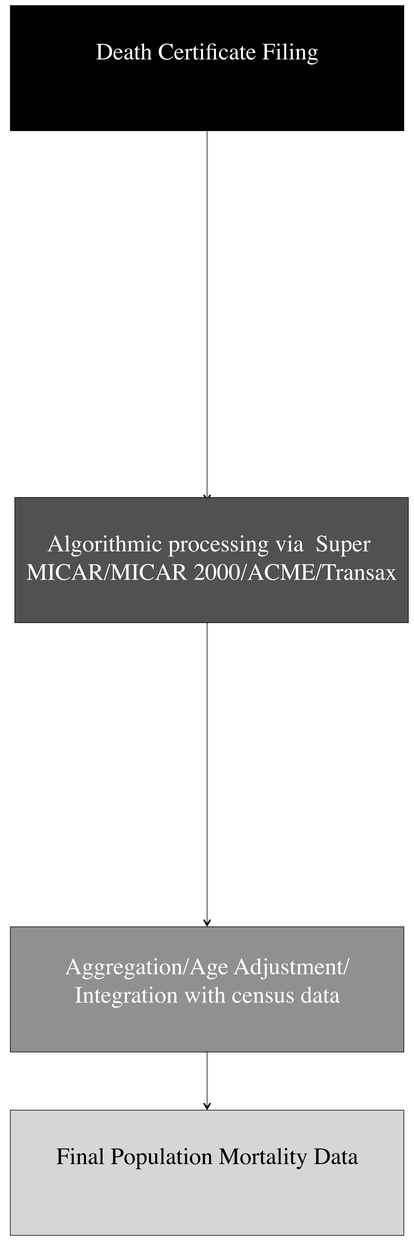
<u>Input level:</u>
- Training is poor: 30-50% of certificate contain major errors and COD determinants can be markedly altered with minimal interventions
- 30-60% of COD diagnoses ultimately are changed on autopsy
- Autopsy rates down>50% over the last 4 decades; less frequently performed in illness.
- Attribution and "sticky-diagnosis" biases have been identified
- Unquantified but other likely contributors: shift work models and less continuity, duty hours restrictions, shift to hospitalist inpatient models

<u>Processing level:</u>

<u>Quantifiable:</u>
- Coding changes: In New Mexico, 50% of the fall in mortality from 1985-1995 attributable to completely artifactual algorithmic changes
- Differing systems: in PLCO 1.4% of deaths were falsely attributed to other causes. Correct attribution leads to 10% increase in prostate cancer-specific mortality
- Improved treatments over time
- Systematic loss of prostate cancers in SEER registries

<u>Unquantifiable:</u>
- Only unreasonable causal mechanisms are disallowed by ACME. Diseases with many manifestations can be accepted as a COD for nearly any condition.
- Only a single disease is allowed to cause prostate cancer: HIV infection
- Causes of death attributable to prostate cancer: cirrhosis, hemopericardium, chronic tubulointerstitial nephritis, etc, etc, etc
- Algorithmic biases: "Senility and Ill-defined condition" rules

<u>Population Level</u>
- The period over which prostate cancer deaths are falling is a period over which the U.S. has undergone demographic shifts and growth of populations with lower prostate cancer mortality
- Competing and changing sources of death



Conclusions

- Population level mortality data is riddled with sources of systematic noise, error and bias that are present at every step involved in constructing this data
- These processes differ substantially from those employed in RCTs of PSA screening in ways that likely make them systematically less accurate
- The processes by which population-level mortality data are constructed have changed over time, with significant effects on the measured mortality rate that are entirely artifactual
- This is a noisy and error prone process that has limited reliability across time and differs markedly from the trial data to which it is often be compared
- It should not be accepted as compelling causal evidence when weighed against RCT data
- These concerns are likely generalizable to other observational mortality trends across time
- Arguments based on population-level data should more closely examine and provide evidence supporting the accuracy and reliability of the underlying processes by which data is collected, aggregated and processed