Interpreting diagnostics to assess threats to valid causal inference: *A comparative safety study of cardiovascular outcomes among patients treated for castration-resistant prostate cancer*

PRESENTER: **Mitchell M. Conover**

INTRODUCTION

- Standard heuristic for determining the validity of a causal comparison:
  - Reject any analysis with a covariate that has StdzDiff>0.10 after adjustment / PS-balancing (Austin, 2009)
- Heuristic is unnecessarily conservative for studies that assess/report high-dimensional sets of covariates
- We implemented a diagnostic process for assessing validity of causal estimates which ignores imbalance for obvious non-confounders.

METHODS

- **Context**: new-user cohort study among patients treated for castration-resistant prostate cancer (CRPC)
  - *Target*: Abiraterone acetate (Zytiga) + predniso(lo)ne
  - *Comparator*: Enzalutamide (Xtandi)
  - *Outcomes*: cardiovascular events
- **Data**: Administrative claims databases
  - Optum Clinformatics® Data Mart Date of Death (DOD)
  - IBM MarketScan Medicare Supplemental (MDCR)
- **Analysis**: Large-scale propensity score (PS) matching using regularized regression
- We proposed to generate 600 distinct analyses / effect estimates (Table 1).
  - Identified "reliable" effect estimates *a priori* using covariate-level and analysis-level diagnostics (Figure 1)

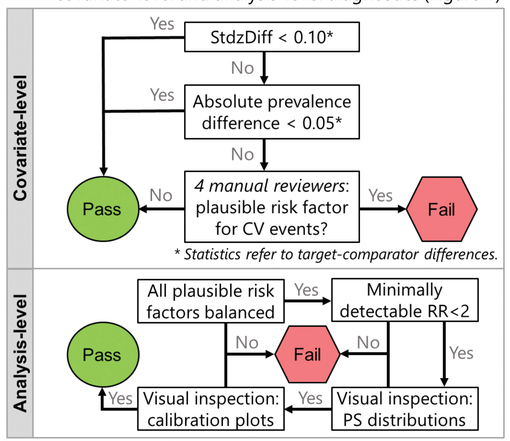

Figure 1. Covariate-level and analysis-level diagnostic procedures

A *priori* diagnostic feasibility assessments should ignore imbalance when covariates are obviously not meaningful confounders.

The widely-used heuristic rejecting any study with a covariate having standardized difference > 0.10 is unnecessarily conservative.

Table 1. Analysis variations that yield 600 unique effect estimates

5 study populations 1. All patients* 2. CVD in prior 180 days 3. No CVD in prior 180 days 4. < 3 CVD occurrences in prior 180 days 5. ≥ 3 CVD occurrences in prior 180 days
6 outcomes-of-interest 1. Heart failure (HF)* 2. Acute myocardial infarction (AMI)* 3. Ischemic stroke* 4. Hemorrhagic stroke 5. Sudden cardiac death 6. Composite (any)
2 databases:1) Optum DOD*, 2) IBM MDCR
10 analysis variations, including: • 2 PS matching strategies 1. 1:1 matching 2. 1:100 variable-ratio matching* • 2 time-at-risk definitions: 1) as-treated*, 2) intent-to-treat • 3 strategies for handling repeat observations: 1) keep all*, 2) keep first, 3) drop all

** Analyses comprised of combinations of the marked parameters are those that passed diagnostic assessment*

RESULTS

- We manually reviewed 473 covariates which were imbalanced in >1 of the 600 proposed analyses based on:
  - Absolute standardized difference >0.1
  - Absolute prevalence difference > 0.05
- 419 (89%) covariates classified as plausible CV risk factor by >1 reviewer
- 9/600 analyses were sufficiently balanced
  - Ignoring imbalance for clear non-confounders
- 4/9 remaining analyses sufficiently powered (MDRR<2)
- 4/4 remaining analyses passed visual inspection of:
  1. Preference score distribution overlap
  2. Negative control calibration plots

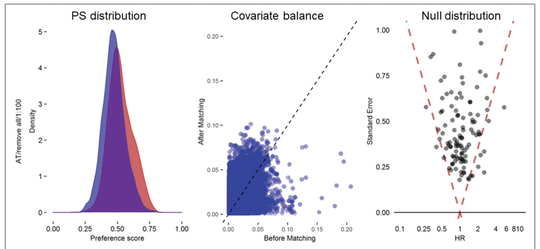

Figure 2. PS distributions, covariate balance, empirical null distribution calibration plots for the (primary) heart failure analysis in Optum DOD

Mitchell M. Conover, James Weaver, Dina Gifkins

janssen PHARMACEUTICAL COMPANIES OF Johnson & Johnson

OHDSI