A Review of Methods for Propensity Score Matched Subgroup Analysis and their Application in Peer Reviewed Research Studies

Shirley V Wang[1], Mengdong He[1], Yinzhu Jin[1], Jessica M Franklin[1], Richie Wyss[1], HoJin Shin[2], Yong Ma[3], Stephine Keeton[3], Bruce Fireman[4], Sara Karami[5], Jacqueline M Major[5], Sebastian Schneeweiss[1], Joshua J Gagne[1]

[1]Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA; [2]Harvard T.H. Chan School of Public Health, Boston, MA; [3]Office of Biostatistics, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD; [4]Kaiser Permanente Northern California, Oakland, CA; [5]Office of Pharmacovigilance and Epidemiology, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, MD

BACKGROUND

- The United States Food and Drug Administration's (FDA) Sentinel Program is a national system for medical product monitoring involving a national distributed network of healthcare databases and a suite of routine querying tools
- The FDA is often interested in examining treatment effect in pre-specified subgroups.
- Trade offs of different methods using propensity score (PS) matching for subgroup analyses is not well understood

OBJECTIVE

- To summarize methods papers comparing the performance of different methods to conduct PS matched subgroup analyses.
- To describe how often different methods for PS matching in subgroup analyses are implemented in practice.

METHODS

- We conducted separate searches for our systematic literature review of methodology and applied research papers
- For methodology papers, we searched PubMed for:
  - "subgroup", "effect modification", "moderation analysis*", or "treatment heterogeneity" co-occurred with "propensity score" in titles or abstracts without restricting publication date.
- For applied research papers, we searched PubMed for:
  - "propensity score match*" and "subgroup analysis*" in any field without restriction by publication date; or published in Pharmacoepidemiology and Drug Safety (PDS) between July 1, 2015 and June 30, 2016 with "propensity score match*" in any field.

RESULTS

Methodology papers

- We identified 5 relevant methods papers after reviewing abstracts of 606 papers from the PubMed search (Table 1)
- These reported small improvements in covariate balance and bias with use of a subgroup-specific PS instead of an overall PS when the true PS varied by subgroup.
- Methods papers only compared strategies that involved re-matching on PS within subgroups.

Table 1. Methodology papers comparing performance of different ways to use propensity score matching for subgroup analyses

Paper	Data	# Subgroups	Performance	PS methods used in subgroups
Green KM, et al. 2014	Empirical	1	Balance (Abs. Std. Diff)	Full matching
Girman CJ, et al. 2014	Empirical	1	Balance (Mean Abs. Std. Diff)	1:1 greedy matching Decile adjustment
Rassen JA, et al. 2012[1]	Empirical, Simulation	3, 1	Balance (Mahalanobis distance) Bias (difference in estimates)	1:1 nearest neighbor matching Decile adjustment Direct adjustment for PS
Kreif N, et al. 2012[2]	Empirical, Simulation	1, 1	Balance (weighted Std. Diff) Bias (RMSE)	1:1 Nearest neighbor matching Genetic matching IPTW
Radice R, et al. 2012	Empirical, Simulation	1, 1	Balance (weighted Std. Diff) Bias (RMSE)	1:1 Nearest neighbor matching Genetic matching, IPTW

[1] Results from simulations directly adjusting for PS, no simulation results available for matching on PS
[2] Results from studies of cost-effectiveness, continuous outcome - cost and quality adjusted life years (QALY)
Std. Diff = standardized difference; Abs = absolute; IPTW = inverse probability of treatment weighting; PS = propensity score; RMSE = root mean squared error

Applied papers

- 83 of 129 papers from the PubMed search met inclusion criteria after abstract review (cohort study, PS matching for main and subgroup analysis, English PDF available)
- Applied papers frequently used PS for subgroup analysis in ways not evaluated in methods papers (Table 2)
  - 33% used PS to match in the overall cohort, then split the 1:1 matched cohort into subgroups without further adjustment
  - 25% provided insufficient detail to clearly determine how PS matching was implemented for subgroup analysis

Table 2. Applied research papers using propensity score matching methods for subgroup analysis

Strategy	N	Proportion
Use overall PS to match for main analysis and do matched analysis in subgroups	14	0.17
Use subgroup specific PS for subgroup analysis	6	0.07
Use overall PS to match within subgroups, aggregate for main analysis	1	0.01
Use overall PS to match for main analysis, within matched cohort, split into subgroups and do analysis without further matching	27	0.33
Use overall PS to match for main analysis, within matched cohort, split into subgroups and use multivariable or other adjustment in subgroups	14	0.17
Unclear (lack sufficient details in publication)	21	0.25
Total	83	1.00

N = number of applied research papers by strategy type

CONCLUSIONS

- The performance of several alternative methods for using a PS to match for subgroup analyses have been evaluated in the methods literature, however the evaluated methods do not include evaluation of the methods for PS matched subgroup analyses most commonly used in applied studies.
- Further evaluation is needed to understand the relative performance of strategies for PS matching in subgroup analyses, particularly within settings with low exposure, infrequent outcomes and multiple subgroups of interest.

CONFLICTS OF INTERST

This work was supported by the United States Food and Drug Administration (FDA) (Grant HHSF223200910006I).Dr. Shirley V Wang was Principal Investigator (PI) of a grant from Novartis Pharmaceuticals Corporation, and a consultant to Aetion, Inc. for unrelated work. Dr. Joshua J Gagne was PI of a grant from Novartis Pharmaceuticals Corporation and a consultant to Aetion, Inc. and to Optum, Inc. for unrelated work. Dr. Sebastian Schneeweiss is PI of the Harvard-Brigham Drug Safety and Risk Management Research Center funded by FDA and investigator-initiated grants from Novartis, and Boehringer Ingelheim. His work is partially funded by grants/contracts from PCORI, FDA, and NHLBI. He is also a consultant to WHISCON, LLC and to Aetion, Inc., a software manufacturer of which he also owns equity.