Reporting of the Trimmed Population in Propensity Score Analyses

Dawn Odom, Xiaolei Zhou, Steven Thomas

RTI Health Solutions, Research Triangle Park, NC, United States

CONFLICT OF INTEREST

None to report.

BACKGROUND

- In nonrandomized studies examining treatment comparisons, propensity scores (PS) are frequently used to account for measured confounding.
- As part of this methodology, the PS distribution is typically reviewed, and, in a practice commonly referred to as "trimming," patients with a PS in the areas of nonoverlap and/or with extreme PS values may be excluded with the aim of improving validity. Likewise, sample reductions can result from use of a PS-matching algorithm.
- Regardless of the PS method used, patients can be excluded from both the study treatment group and the comparator group.
- In studies examining safety events, trimming or excluding patients can lead to an incomplete safety profile of the treated patients and reduce the generalizability of results.
- It has been recommended that information on treated patients who were excluded from the analysis be presented to provide a more complete understanding of the study population.[1]

OBJECTIVE

- By performing a targeted literature review, we aimed to quantify the PS methods used in nonrandomized safety cohort studies in recent years and determine if summary information is provided on the trimmed or excluded patients.

METHODS

- A targeted review was conducted in PubMed using the key term "propensity score" for years 2014 and 2015 in the following six leading pharmacoepidemiology journals:
  - *American Journal of Epidemiology*
  - *Drug Safety*
  - *Epidemiology*
  - *European Journal of Epidemiology*
  - *Journal of Clinical Epidemiology*
  - *Pharmacoepidemiology and Drug Safety*
- Articles were included if they were nonrandomized cohort studies with the principal objective of examining treatment safety and presented results from a PS analysis using the following PS methods: stratification, matching, regression adjustment, or inverse probability weighting. The PS analysis could have served as a primary, secondary, or sensitivity analysis.
- In addition, each study was evaluated to determine if any sample reduction/trimming occurred. If so, details on the sample reduction and summary information on patient characteristics and safety outcomes for the overall population and the excluded population were extracted.
- The review was conducted in January 2016. Two independent reviewers examined each article to determine eligibility and classification. In the event of a conflicting review, a third opinion determined eligibility.

DISCUSSION AND CONCLUSIONS

- When PS trimming is conducted in safety cohort studies, researchers often report information on the overall and analysis (e.g., PS-matched) sample without presenting information on the excluded patients, thus missing an important piece of information when evaluating the study's generalizability and a treatment's safety profile.
- When patients on the study treatment are excluded/trimmed because of difficulty in obtaining relevant comparator patients (particularly with PS matching), safety information, such as the crude incidence rate or prevalence (with 95% confidence interval), can still be obtained on these excluded individuals.
  - This can be especially important for patients with a high probability of receiving the study treatment. A high incidence or prevalence of a safety outcome with a narrow confidence interval may provide important safety signals that a root cause analysis or more targeted safety study is needed.
- However, when PS methods are used as a secondary or sensitivity analysis, it may be less important to report this information if the primary analytic methods did not exclude patients. Further research should be conducted to account for the type of analysis (i.e., primary vs. secondary) when reviewing this literature.
- Despite the limitations of our small targeted review, this qualitative synthesis provides a general sense of information reported in recent pharmacoepidemiology studies involving PS analysis.

CONTACT INFORMATION

Dawn Odom, MS
Global Head Biostatistics

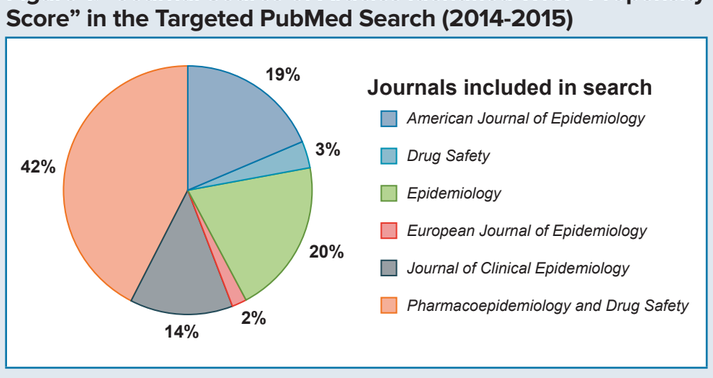
RTI Health Solutions
200 Park Offices Drive
Research Triangle Park, NC 27709

Phone: +1.919.316.3529
E-mail: dodom@rti.org

RESULTS

- A total of 59 articles were identified in the PubMed search. As shown in Figure 1, most (42%) were from *Pharmacoepidemiology and Drug Safety*, followed by *Epidemiology* (20%) and the *American Journal of Epidemiology* (19%), with the remaining studies found in the other targeted journals.
- Of the 59 articles identified in the PubMed search, 18 articles were eligible for qualitative synthesis (Figure 2, Table 1). The majority were excluded as methods articles (24 articles), reviews (4 articles), or editorials (6 articles). Seven articles were excluded because they were not a safety cohort study or did not include results from the PS analyses (Figure 2).

Figure 1. Journal Source of Articles Identified With "Propensity Score" in the Targeted PubMed Search (2014-2015)



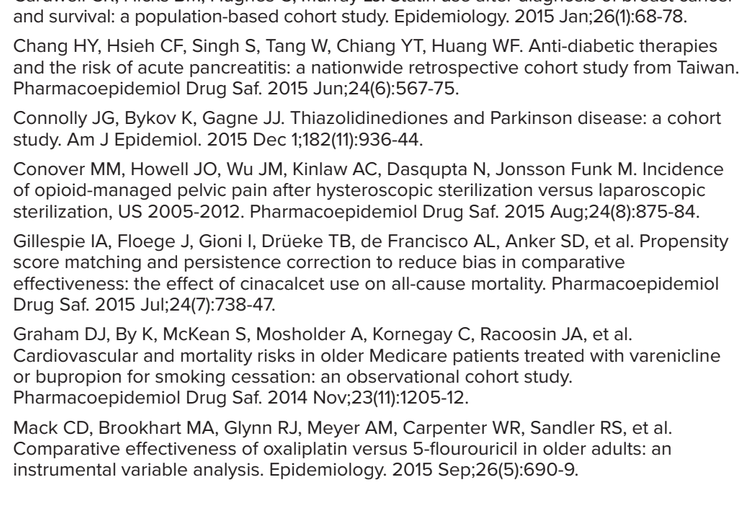
Summary of PS Methods and Sample Size Reduction

- PS matching was used by 13 articles, making it the most frequently used method among the 18 eligible articles, followed by regression adjustment (4 articles) and inverse probability weighting (3 articles); stratification was implemented in 2 of the articles (Table 1). Note that 4 articles applied more than one PS method (e.g., reported a primary and secondary analysis using different PS methods).
- Only a single article performed PS distribution trimming when using PS weighting, regression adjustment, or stratification analysis. The single article trimmed percentiles in a stratified analysis.[12]
- Of the articles that used PS matching, the sample size of the treatment group and the comparator group were commonly reduced as part of the matching algorithm. In some articles, the reduction of sample size for the treatment group was more than 50% (Table 2).
- Additionally, two articles that used PS-matching methods trimmed the overall population prior to implementing the matching algorithm.

Summary of Reported Patient Characteristics and Safety Outcomes

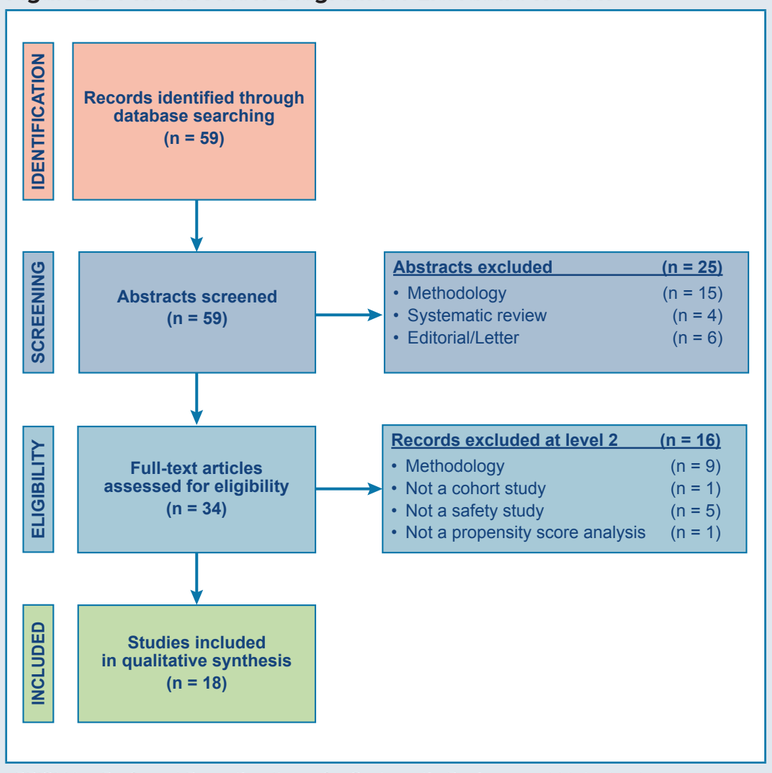
- When PS matching was used, researchers commonly reported patient characteristics on the overall sample and the PS-matched sample; in addition, seven (54%) reported safety comparative analysis using the overall sample (e.g., hazard ratio of the safety event) (Figure 3).
- The one article that performed PS distribution trimming in a stratified analysis reported patient characteristics and safety outcome information (i.e., prevalence) by treatment group for what seemed to be the overall sample.
- Across all methods, no article reported summary statistics on demographics or safety outcomes specifically on the excluded/trimmed patients.

Figure 3. Summary of PS Methods



Note: Each method was counted separately in studies presenting multiple analyses involving different PS methods.

Figure 2. PRISMA Flow Diagram for Literature Review



PRISMA = Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

Table 1. Articles Included in Qualitative Synthesis and PS Method(s) Used

Article	PS Method(s) Used
Antoniou, et al.[2]	Regression adjustment
Badillo R, et al.[3]	Matching
Cardwell CR, et al.[4]	Matching
Chang HY, et al.[5]	Weighting
Connolly JG, et al.[6]	Matching
Conover MM, et al.[7]	Matching, weighting
Gillespie IA, et al.[8]	Matching
Graham DJ, et al.[9]	Matching, stratification
Mack CD, et al.[10]	Matching, weighting
Mehta S, et al.[11]	Regression adjustment
Mines D, et al.[12]	Stratification
Nielsen NM, et al.[13]	Matching
Romanelli RJ, et al.[14]	Matching
Sohn M, et al.[15]	Matching
Swanson SA, et al.[16]	Regression adjustment
Weinhandl ED, et al.[17]	Matching
Wood ME, et al.[18]	Matching, regression adjustment
Zhou EH, et al.[19]	Matching

Table 2. Sample Sizes in the Overall and PS-Matched Sample

Articles That Used PS Matching	Sample Size (Study Drug + Comparator)	
Article	**Overall Sample**	**PS-Matched Sample**
Badillo R, et al.[3]	43,438 (13,626 + 29,812)	12,684 (6,342 + 6,342)
Cardwell CR, et al.[4]	17,880 (4,282 + 13,598)	3,784 (1,892 + 1,892)
Connolly JG, et al.[6]	29,397 (5,230 + 24,167)	10,450 (5,225 + 5,225)
Conover MM, et al.[7]	71,875 (26,927 + 44,948)	51,160 (25,580 + 25,580)
Gillespie IA, et al.[8]	9,101 (1,168 + 7,933)	2,322 (532 + 1,790)
Graham DJ, et al.[9]	88,957 (74,824 + 14,133)	58,350 (46,680+ 11,670)
Mack CD, et al.[10]	3,660 (1,565 + 2,095)	Not available
Nielsen NM, et al.[13]	1,531,832 (16,234 + 1,515,598)	70,200 (16,028 + 54,172)
Romanelli RJ, et al.[14]	16,364 (5,156 + 11,208)	8,470 (4,235 + 4,235)
Sohn M, et al.[15]	403,345 (6,510+396,835)	28,316 (6,236 + 22,080)
Weinhandl ED, et al.[17]	7,791,072 (15,635 + 7,775,437)	46,899 (15,633 + 31,266)
Wood ME, et al.[18]	4,204 (375 + 3,829)	1,095 (365 + 730)
Zhou EH, et al.[19]	58,617 (3,964 + 54,653)	19,820 (3,964 + 15,856)

REFERENCES

1. Sturmer T, Rothman KJ, Avorn J, Glynn RJ. Treatment effects in the presence of unmeasured confounding: dealing with observations in the tails of the propensity score distribution--a simulation study. Am J Epidemiol. 2010 Oct 1;172(7):843-54.
2. Antoniou T, Yao Z, Camacho X, Mamdani MM, Juurlink DN, Gomes T. Safety of valproic acid in patients with chronic obstructive pulmonary disease: a population-based cohort study. Pharmacoepidemiol Drug Saf. 2015 Mar;24(3):256-61.
3. Badillo R, Schmidt R, Mortensen EM, Frei CR, Mansi I. Statin therapy and gastrointestinal hemorrhage: a retrospective cohort study with propensity score-matching. Pharmacoepidemiol Drug Saf. 2015 Aug;24(8):849-57.
4. Cardwell CR, Hicks BM, Hughes C, Murray LJ. Statin use after diagnosis of breast cancer and survival: a population-based cohort study. Epidemiology. 2015 Jan;26(1):68-78.
5. Chang HY, Hsieh CF, Singh S, Tang W, Chiang YT, Huang WF. Anti-diabetic therapies and the risk of acute pancreatitis: a nationwide retrospective cohort study from Taiwan. Pharmacoepidemiol Drug Saf. 2015 Jun;24(6):567-75.
6. Connolly JG, Bykov K, Gagne JJ. Thiazolidinediones and Parkinson disease: a cohort study. Am J Epidemiol. 2015 Dec 1;182(11):936-44.
7. Conover MM, Howell JO, Wu JM, Kinlaw AC, Dasqupta N, Jonsson Funk M. Incidence of opioid-managed pelvic pain after hysteroscopic sterilization versus laparoscopic sterilization, US 2005-2012. Pharmacoepidemiol Drug Saf. 2015 Aug;24(8):875-84.
8. Gillespie IA, Floege J, Gioni I, Drüeke TB, de Francisco AL, Anker SD, et al. Propensity score matching and persistence correction to reduce bias in comparative effectiveness: the effect of cinacalcet use on all-cause mortality. Pharmacoepidemiol Drug Saf. 2015 Jul;24(7):738-47.
9. Graham DJ, By K, McKean S, Mosholder A, Kornegay C, Racoosin JA, et al. Cardiovascular and mortality risks in older Medicare patients treated with varenicline or bupropion for smoking cessation: an observational cohort study. Pharmacoepidemiol Drug Saf. 2014 Nov;23(11):1205-12.
10. Mack CD, Brookhart MA, Glynn RJ, Meyer AM, Carpenter WR, Sandler RS, et al. Comparative effectiveness of oxaliplatin versus 5-flourouricil in older adults: an instrumental variable analysis. Epidemiology. 2015 Sep;26(5):690-9.
11. Mehta S, Pulungan Z, Jones BT, Teigland C. Comparative safety of atypical antipsychotics and the risk of pneumonia in the elderly. Pharmacoepidemiol Drug Saf. 2015 Dec;24(12):1271-80.
12. Mines D, Tennis P, Curkendall SM, Li DK, Peterson C, Andrews EB, et al. Topiramate use in pregnancy and the birth prevalence of oral clefts. Pharmacoepidemiol Drug Saf. 2014 Oct;23(10):1017-25.
13. Nielsen NM, Svanström H, Stenager E, Magyari M, Koch-Henriksen N, Pasternak B, et al. The use of valproic acid and multiple sclerosis. Pharmacoepidemiol Drug Saf. 2015 Mar;24(3):262-8.
14. Romanelli RJ, Jukes T, Segal JB. Compliance after switching from branded to generic statins. Pharmacoepidemiol Drug Saf. 2014 Oct;23(10):1093-100.
15. Sohn M, Talbert J, Blumenschein K, Moga DC. Atypical antipsychotic initiation and the risk of type II diabetes in children and adolescents. Pharmacoepidemiol Drug Saf. 2015 Jun;24(6):583-91.
16. Swanson SA, Hernandez-Diaz S, Palmsten K, Mogun H, Olfson M, Huybrechts KF. Methodological considerations in assessing the effectiveness of antidepressant medication continuation during pregnancy using administrative data. Pharmacoepidemiol Drug Saf. 2015 Sep;24(9):934-42.
17. Weinhandl ED, Gilbertson DT, Collins AJ, Foley RN. Relative safety of peginesatide and epoetin alfa. Pharmacoepidemiol Drug Saf. 2014 Oct;23(10):1003-11.
18. Wood ME, Frazier JA, Nordeng HM, Lapane KL. Prenatal triptan exposure and parent-reported early childhood neurodevelopmental outcomes: an application of propensity score calibration to adjust for unmeasured confounding by migraine severity. Pharmacoepidemiol Drug Saf. 2016 May;25(5):493-502.
19. Zhou EH, Gelperin K, Levenson MS, Rose M, Hsueh YH, Graham DJ. Risk of acute myocardial infarction, stroke, or death in patients initiating olmesartan or other angiotensin receptor blockers - a cohort study using the Clinical Practice Research Datalink. Pharmacoepidemiol Drug Saf. 2014 Apr;23(4):340-7.

Presented at: 32nd International Conference on Pharmacoepidemiology & Therapeutic Risk Management, August 25-28, 2016, Dublin, Ireland