



Bias Gaps in Real-World Data-Based Economic Evaluation: Underreporting, Weak Control, and the Need to Upgrade Propensity Score

Yu Xin¹, Yilin Deng², Changjin Wu³, Yuanyi Cai⁴, Jun Hao⁵, ling Zuo¹, Wen Hui¹

Contact: Yu Xin (xinyu1006@scu.edu.cn)
Wen Hui (huiwen@scu.edu.cn)

1 West China Hospital of Sichuan University, Chengdu, China, 2 West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China, 3 College of Public Health, Chongqing Medical University, Chongqing, China, 4 School of Health Management, China Medical University, Shenyang, China, 5 The First Affiliated Hospital of Nanchang University, Nanchang, China

BACKGROUND

- **The Problem:** Traditional trial-based economic evaluation (EE) is often plagued by the efficacy-effectiveness gap, as it fails to reflect real-world resource consumption and population heterogeneity in clinical practice.
- **The Opportunity:** Real-world data (RWD)-based EE holds significant advantages in terms of result generalizability and extrapolation, as it can theoretically directly reflect the economic outcomes of research interventions.
- **The Critical Barrier:** RWD is frequently prone to selection bias and confounding factors, which may compromise the validity and reliability of subsequent economic evaluation results.

REASERCH OBJECTIVE

To assess the methodological robustness of RWD-based EE through two core dimensions: 1) Transparency in bias reporting. 2) Bias control methods in selection bias and confounding.

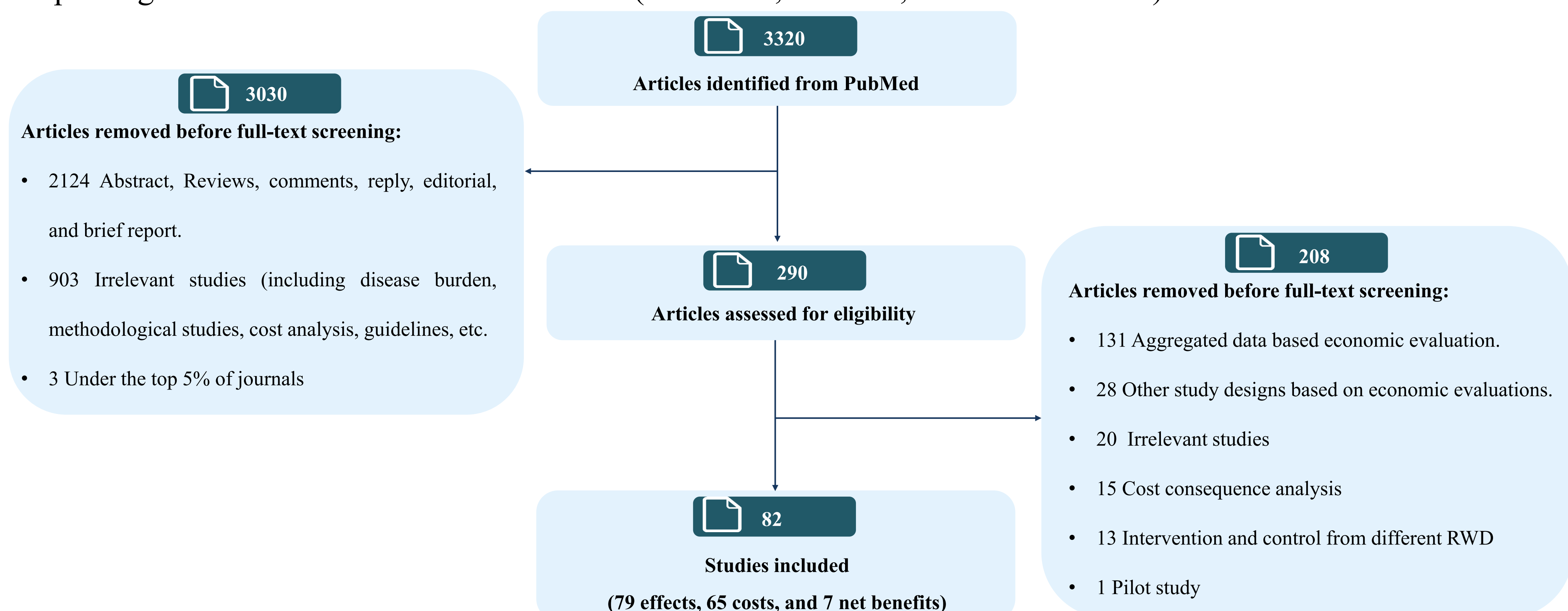
METHODS

- **Search Strategy:** Full EE published between Jan 1, 2010, and December 31, 2024 were included.
- **Quality Filter:** Studies were strictly limited to those published in journals ranked in the top 5% by Journal Impact Factor across 51 Clinical Medicine subcategories.
- **Data Extraction:** Data extraction was conducted by two trained groups of independent reviewers, capturing information at two levels. 1) *Study-Level:* Explicit identification and control of selection bias and confounding were evaluated. 2) *Outcome-Level:* Bias control methods applied to specific outcome measures (effect, cost, net benefit) were assessed.

RESULTS

1. Search and Screening Results

The initial search retrieved 3,320 studies. After title, abstract and full-text article screening, a total of 82 EEs were included, encompassing 151 outcome measures from RWD (79 effects, 65 costs, and 7 net benefits).

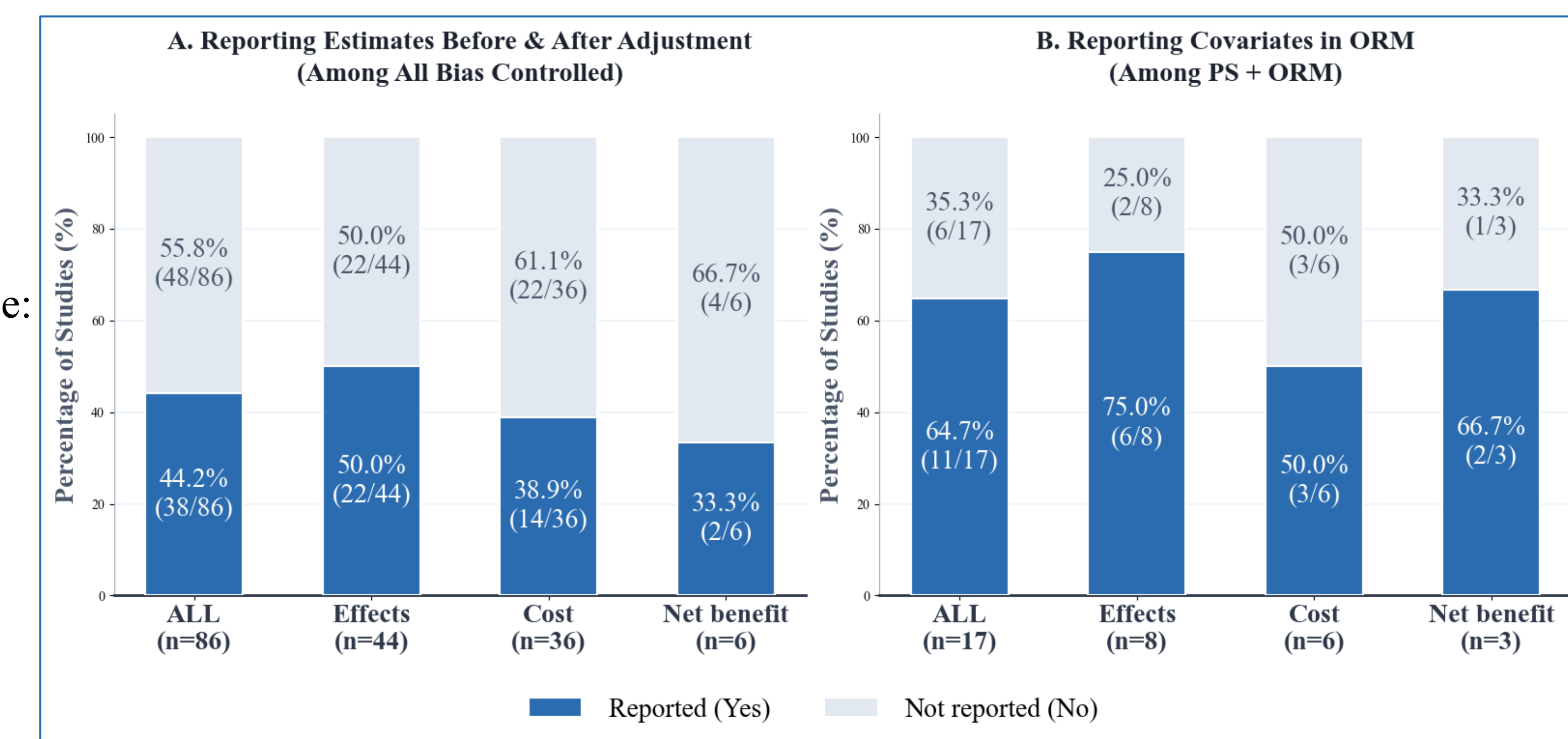
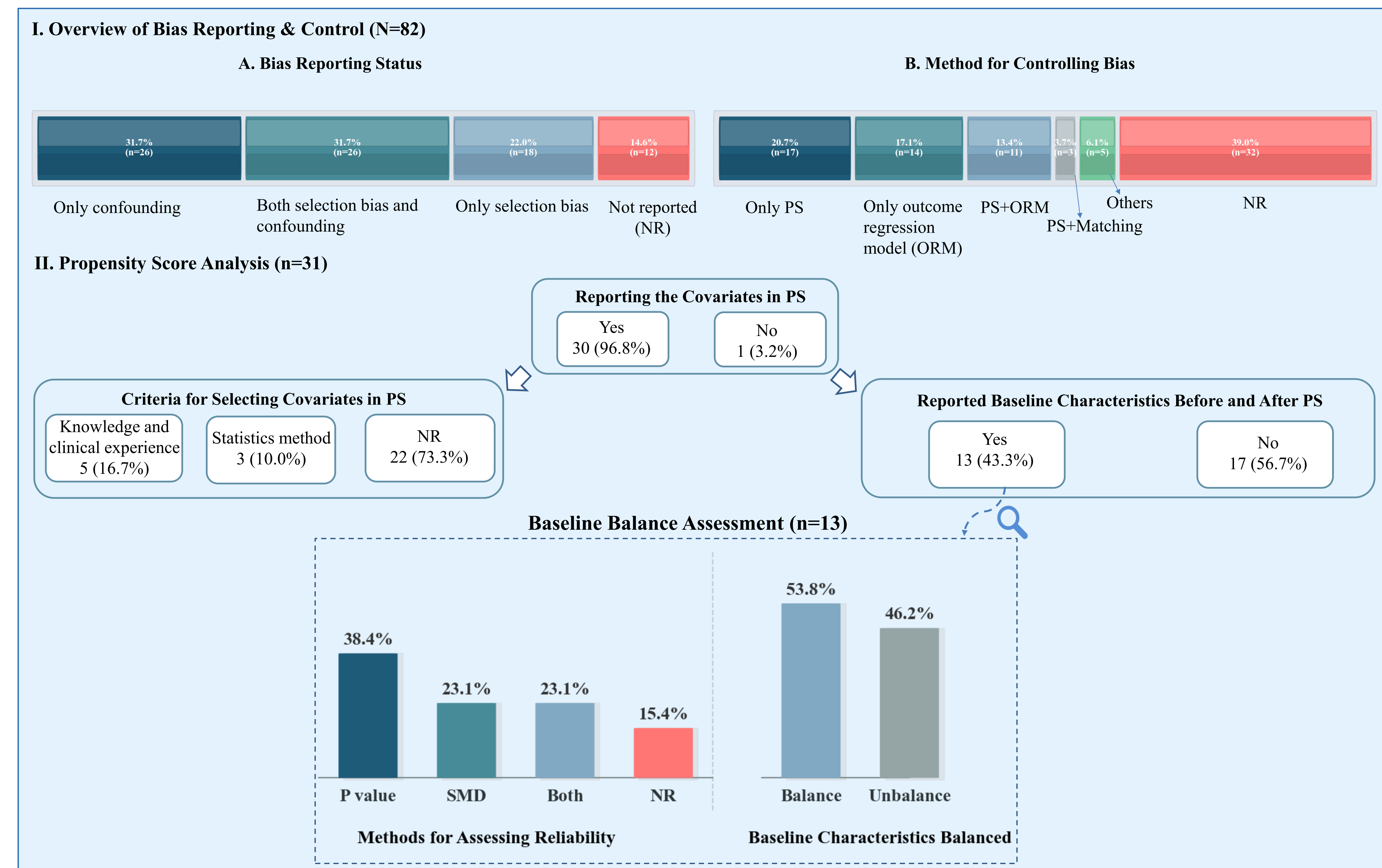


2. Analysis of Bias Control for Study

- **High Prevalence of Weak Bias Control:** 14.6% omitted bias reporting, and 39.0% of studies failed to control bias.
- **The “Black Box” of Covariate Selection:** Among propensity score (PS) studies, a striking 73.3% lacked any rationale for covariate selection.
- **The Illusion of Balance:** Only 43.3% of PS studies assessed baseline balance. Of these, 53.8% achieved adequate balance ($|SMD| < 0.1$; SMD = standardized mean difference).

3. Analysis of Bias Control for Outcome Measures

- **Pervasive Lack of Bias Control:** A substantial 43.1% (65/151) of all outcome measures lacked any statistical bias control.
- **Missing Impact Verification:** The majority (55.8%) of bias-controlled measures failed to report estimates both before and after adjustment. This varied by outcome type: effect (50.0%), cost (61.1%), and net benefit (66.7%).
- **Not reported Covariates in PS + ORM:** Covariate reporting in ORM was inadequate overall (only 64.7% reported), with clear differences by outcome: effect (75.0%), cost (50.0%), and net benefit (66.7%).



CONCLUSIONS

- **Systemic Methodological Deficits:** Current RWD-based EE suffers from inadequate reporting and suboptimal control of selection bias and confounding.
- **The PS Quality Gap:** Although PS is widely adopted, the quality of its implementation and reporting requires significant improvement to enhance the credibility of observational economic evaluations. Most studies omit core methodological details, and nearly half fail to achieve adequate balance.
- **Cost-Specific Double-Robustness Issues:** Cost measures via PS + ORM show the lowest rate of covariate reporting, undermining the transparency of double-robust estimates.