

BACKGROUND

Randomized controlled trials (RCTs) are the gold standard of evidence-based medicine, but surgical RCTs face unique challenges that can compromise methodological quality. In plastic surgery, factors such as variability in surgical technique, difficulty blinding, patient preferences, and subjective outcomes make high-quality trial design especially difficult. These challenges have contributed to lower-quality RCTs and limited their impact on clinical practice. Although the number of RCTs in plastic surgery has increased alongside a broader shift toward evidence-based medicine and reporting standards like CONSORT, the quality of these studies remains uncertain. This review evaluates the methodological and reporting quality of RCTs published in Plastic and Reconstructive Surgery from 2013 to 2022 to identify progress and areas for improvement.

METHODS

We conducted a systematic review of all clinical randomized controlled trials (RCTs) published in Plastic and Reconstructive Surgery from 2013 to 2022. Articles were identified through PubMed/MEDLINE using the terms “randomized,” “clinical,” and “trial,” with inclusion limited to clinical RCTs. Two independent reviewers screened all studies and extracted data using a double-screening approach to minimize bias, with an initial pilot phase to calibrate consistency. For each included study, we recorded key trial characteristics such as topic, intervention type, sample size, authorship, and reporting of primary outcome and analysis set. We also captured indicators of methodological rigor such as epidemiology/statistical support and funding. Methodological quality was assessed using two validated scoring tools: the 5-point Jadad score, which evaluates randomization, blinding, and accountability of trial participants; and a modified 10-point Cochrane Risk-of-Bias (RoB) tool, which assesses additional domains such as allocation concealment, outcome reporting, and investigator expertise.

Jadad Criteria	Points	Yes (%)	No (%)
Was the trial described as randomized?	+1	146 (100)	0
Was the randomization method described and appropriate?	+1	91 (62.3)	55 (37.7)
Was the trial described as double-blinded?	+1	50 (34.2)	96 (65.8)
Was the double-blinding method described and appropriate?	+1	45 (30.8)	101 (69.2)
Was there an adequate description of patient withdrawals and drop-outs?	+1	104 (71.2)	42 (28.8)
Deduct if randomization method was described but inappropriate	-1	3 (2.1)	143 (97.9)
Deduct if blinding method was described but inappropriate	-1	1 (0.7)	145 (99.3)

Table 1: Analysis of Jadad Score

Item	Criteria	Yes (%)	No (%)	Unclear (%)
Randomization sequence generation	Described method with no a priori knowledge of allocation	91 (62.3)	11 (7.53)	44 (30.1)
Allocation concealment	Concealment method stated and allocation hidden until patient assignment	65 (44.5)	8 (5.48)	73 (50.0)
Blinding of surgeons	Considered blinded if explicitly stated. For “single-” or “double-blinded” trials, unmentioned parties were assumed unblinded.	34 (23.3)	58 (39.7)	54 (37.0)
Blinding of assessors		83 (56.8)	28 (19.2)	35 (24.0)
Blinding of patients		63 (43.2)	40 (27.4)	43 (29.4)
Infrequent loss to follow-up	<20% loss of randomized cohort	89 (61.0)	22 (15.1)	35 (23.9)
No selective outcome reporting	All collected/analyzed outcomes reported; no vague reporting (e.g., “NS”)	124 (84.9)	22 (15.1)	0
Objective outcomes, assessed to limit bias	Outcomes appropriate, clearly defined, and assessed objectively (e.g. blinding or multiple assessors)	142 (97.3)	4 (2.7)	0
Adequate sample size	A priori power calculation performed and met	52 (35.6)	94 (64.4)	0
Equal investigator expertise	Same surgeon for both treatments, or comparable expertise described	82 (56.2)	7 (4.79)	57 (39.0)

Table 2: Analysis of RoB Score

RESULTS

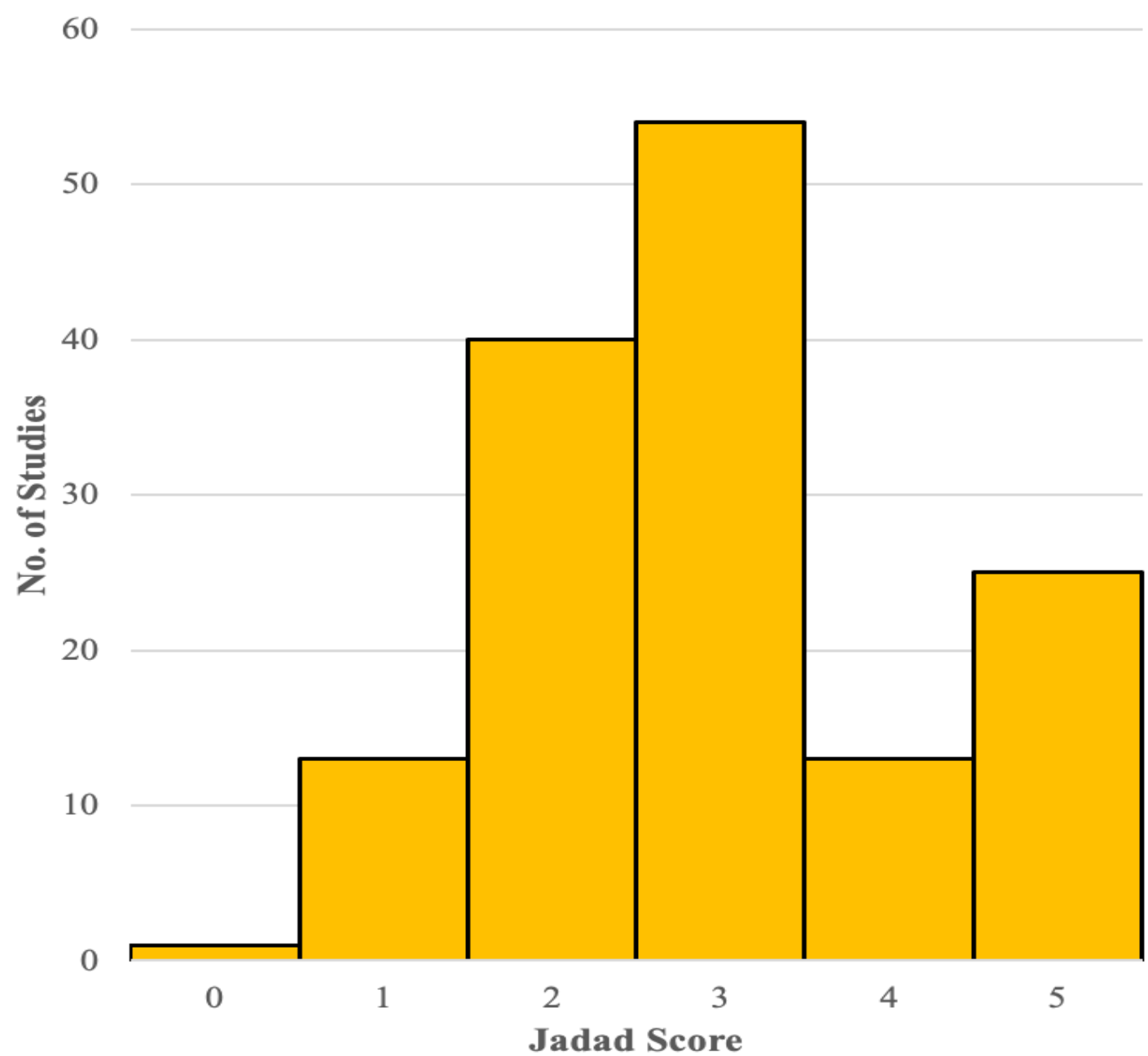


Figure 1: Frequency of Jadad scores

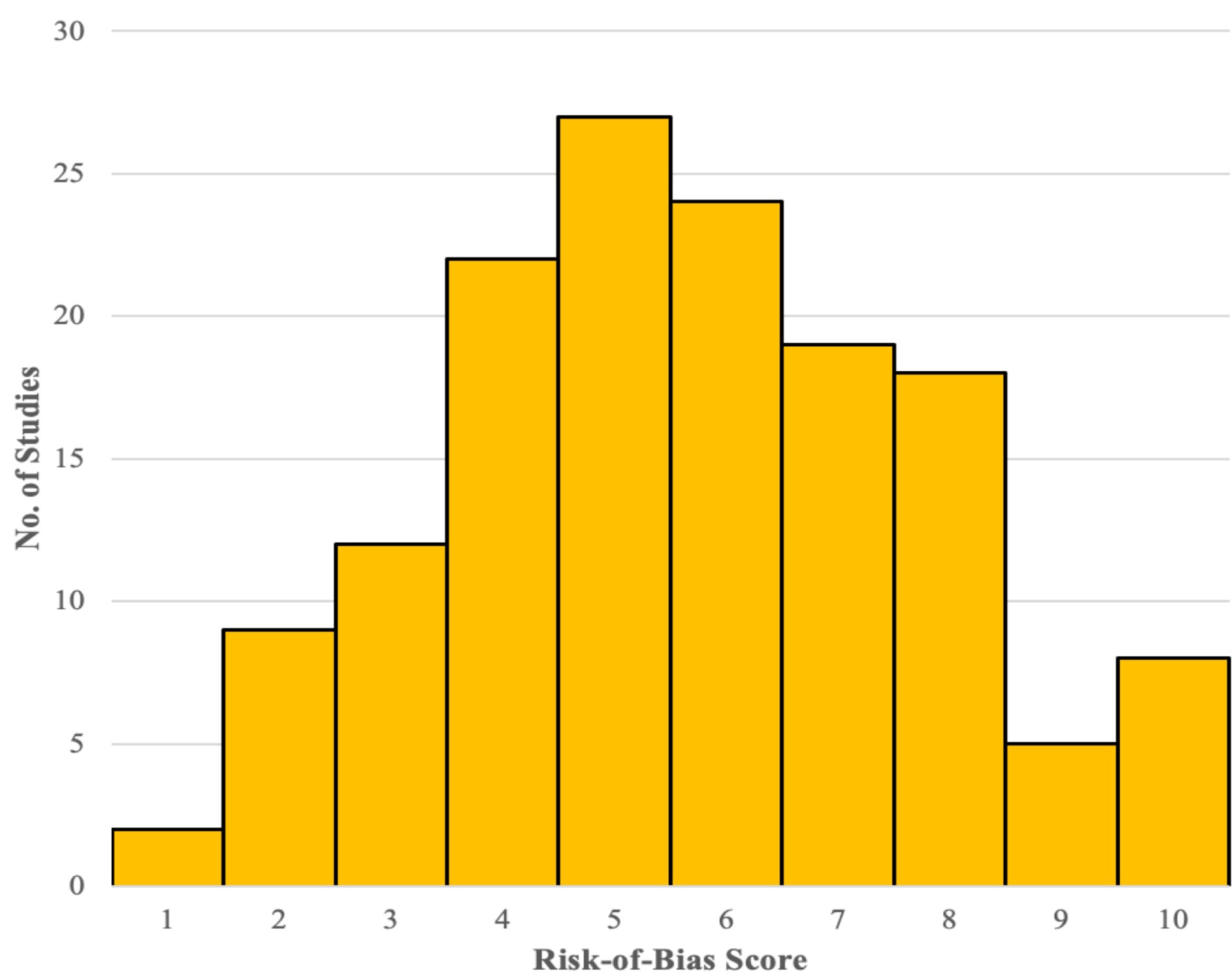


Figure 2: Frequency of RoB scores

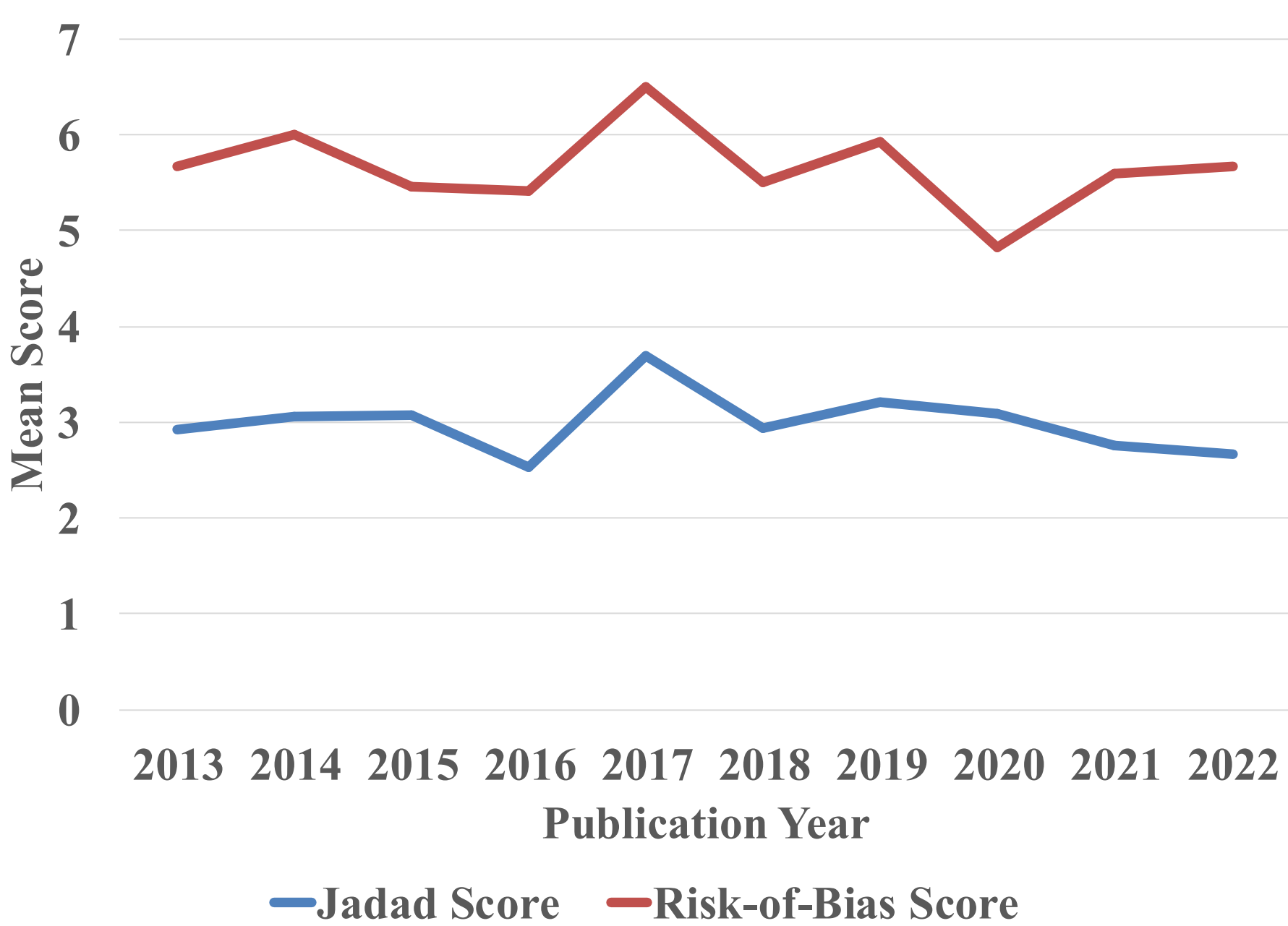


Figure 3: Temporal trend of Jadad and RoB scores

- 227 articles identified; 146 met inclusion criteria (Figure 1).
- Data abstracted by two independent reviewers with high agreement ($\kappa = 0.90$, $p < 0.001$).
- From 2013–2022, an average of 14.6 ± 2.5 RCTs were published annually.
- Mean sample size was 93.5 ± 98.7 patients.
- Surgeons were first authors in 83.6% and senior authors in 85.6% of studies.
- Superiority designs dominated (93.2%); only 6.8% were non-inferiority trials; no equivalence trials.
- Most common topics: cosmetic (26.7%), reconstruction (23.3%), breast (20.5%).
- 73.2% lacked epidemiological affiliation; 65.1% did not cite statistical support; 47.9% of trials reported receiving funding.
- Most common interventions: drugs (38.3%) and medical devices (32.9%).
- Primary outcomes were explicitly stated in 54.1% of trials; 73.3% of trials reported positive results.
- Analysis set usage: 21.2% ITT, 30.8% PP, 4.8% both, and 43.2% unspecified.
- Mean Jadad score was 2.96 ± 1.21 and mean RoB score was 5.65 ± 2.14
- Multiple regression: surgical technique ($\beta = -0.69$, $p = 0.004$) and device trials ($\beta = -0.99$, $p < 0.001$) associated with lower Jadad scores; adequately powered trials associated with higher scores ($\beta = 0.49$, $p = 0.015$).
- Multiple regression showed breast trials associated with higher RoB scores ($\beta = 1.2$, $p = 0.004$); surgical technique ($\beta = -1.1$, $p = 0.009$) and device trials ($\beta = -1.2$, $p = 0.002$) associated with lower RoB scores.
- Subgroup analysis showed breast trials had declining Jadad ($\beta = -0.19$, $p = 0.024$) and RoB scores ($\beta = -0.28$, $p = 0.042$) over time

Jadad Score	Coefficient	Standard Error	t	P-value	95% Confidence Interval	
Primary outcome stated	0.33	0.19	1.7	0.087	-0.049	0.72
Adequately powered sample	0.49	0.20	2.5	0.015	0.098	0.88
Intervention-Surgical Technique	-0.69	0.24	-2.9	0.004	-1.2	-0.22
Intervention-Medical Devices/Equipment	-0.99	0.21	-4.7	<0.001	-1.4	-0.57

Table 3: Multivariable regression analysis for Jadad scores

Risk-of-Bias Score	Coefficient	Standard Error	t	P-value	95% Confidence Interval	
Number of Authors	0.097	0.064	7.5	0.13	-0.03	0.22
Intervention-Surgical Technique	-1.1	0.42	-2.7	0.009	-1.9	0.29
Intervention-Medical Devices/Equipment	-1.2	0.37	-3.2	0.002	-1.9	-0.47
Subject-Breast	1.2	0.4	3.0	0.004	0.40	2.0
Primary outcome stated	0.66	0.35	1.9	0.061	-0.03	1.4

Table 4: Multivariable regression analysis for RoB scores

CONCLUSION

RCT quality in PRS has improved over time but remains moderate, with persistent gaps in reporting and methodology. Emphasizing rigorous design and transparent reporting is essential to strengthen future surgical trials.

Questions?

Email cgordonw@med.umich.edu