The Quality of Randomized Controlled Trials in High-impact Rheumatology Journals, 1998–2018

Michael S. Putman, Ashley Harrison Ragle, and Eric M. Ruderman

ABSTRACT. ***Objective.*** Well-designed randomized controlled trials (RCT) mitigate bias and confounding, but previous evaluations of rheumatology trials found high rates of methodological flaws. Outside of rheumatoid arthritis, no studies in the modern era have assessed the quality of rheumatology RCT over time or regarding industry funding.

Methods. We identified all RCT published in 3 high-impact rheumatology journals from 1998, 2008, and 2018. Quality metrics derived from a modified Jadad scale were analyzed by year of publication and by funding source.

Results. Ninety-six publications met inclusion criteria; 82 of these described the primary analysis of an RCT. Over time (1998–2008–2018), trials were less likely to adequately report dropouts and withdrawals (100% vs 82% vs 60%; $p < 0.01$) or include an active comparator (44% vs 12% vs 13%; $p = 0.01$). Later trials were more likely to evaluate biologic therapy (11% vs 38% vs 83%; $p < 0.01$) and report adequate randomization procedures (39% vs 29% vs 60%; $p = 0.04$). Seventy-nine percent of trials received industry funding. Industry-funded trials were more likely to report double-blinding (86% vs 53%; $p < 0.01$), patient-reported outcome measures (77% vs 41%; $p < 0.01$), and intention-to-treat analyses (86% vs 65%; $p = 0.04$).

Conclusion. Industry-funded trials comprise the majority of RCT published in high-impact rheumatology journals and more frequently report metrics associated with RCT quality. RCT assessing active comparators and nonbiologic therapies have become less common in high-impact rheumatology journals. (First Release July 1 2020; J Rheumatol 2020;47:1446–9; doi:10.3899/jrheum.191306)

Key Indexing Terms:
RANDOMIZED CONTROLLED TRIALS QUALITY RHEUMATOLOGY

Randomized controlled trials (RCT) are the gold standard for assessing the efficacy of pharmaceutical interventions[1]. Randomization and blinding help mitigate issues of bias and confounding, which limit the validity of observational data[2]. However, inadequate randomization, poor allocation concealment, and flawed analysis plans may threaten the validity of RCT[3,4]. Moreover, industry funding has been suggested to bias trials toward a positive result[5]. Because RCT influence US Food and Drug Administration (FDA) approvals and clinical practice guidelines, their quality is of academic interest.

Previous work assessing rheumatology RCT uncovered low rates of adherence to quality metrics. A 2002 report[6] found that fewer than one-third of trials from 1997 to 1998 used adequate randomization, appropriate allocation concealment, or intention-to-treat analysis. A 2003 publication on systemic lupus erythematosus (SLE) trials reported similar results[7]. Two recent studies on RCT in rheumatoid arthritis (RA) did not observe a bias toward positive outcomes among industry studies and both found higher rates of blinding[8,9]. One also found fewer than half of studies adequately describing randomization procedures or allocation concealment[9].

We describe RCT from 3 high-impact rheumatology journals from the years 1998, 2008, and 2018. All disease states were included. Similar to the methodology employed by Hill, *et al*, we scored trials using a modified Jadad scale, which assessed the adequacy of important RCT quality metrics[6,10]. We hypothesized that quality would improve over time and that industry funding would be associated with a higher rate of statistically significant primary outcome measures.

MATERIALS AND METHODS

Journal selection. The top 3 rheumatology journals that publish general interest primary research were identified using the h5-index, which ranks journals according to the volume of highly cited articles over the previous 5 years. A search for "rheumatology" was performed on scholar.google.com/citations?view_op=top_venues[11]. *Annals of the Rheumatic Diseases*, *Arthritis and Rheumatology*, and *Rheumatology* were included. *Nature Reviews: Rheumatology* had a high h5-index but was excluded because it does not publish RCT.

Article selection. Articles from 1998, 2008, and 2018 were included if they described an RCT. The following study types were excluded:

From the Division of Rheumatology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA.

M.S. Putman, MD; A. Harrison Ragle, MD; E.M. Ruderman, MD, Division of Rheumatology, Northwestern University Feinberg School of Medicine.

Address correspondence to Dr. M.S. Putman, 3-731 W. Briar Place, Chicago, Illinois 60657, USA. E-mail: msputman@gmail.com

Accepted for publication March 11, 2020.

nonpharmaceutical clinical interventions, open-label extensions, biosimilar equivalence studies, pharmacodynamic/genetic studies, biomarker studies, and dose-finding studies. Author MSP identified potential articles by title, and eligibility was confirmed by review of the abstract and manuscript.

Article review. MSP extracted title, date of publication, and total number of patients included. "Industry funding" was identified and defined as any financial support. MSP also identified whether the article reported a primary or secondary analysis of an RCT, a biologic or targeted disease-modifying antirheumatic drug, an active comparator trial, or a primary outcome with $p < 0.05$. MSP and AHR conducted a blinded co-review using a modified Jadad scale (Table 1), which retained the elements used by Hill, *et al*[6] (adequate blinding, adequate randomization, allocation concealment, adequate description of withdrawals and dropouts), and incorporated additional elements from the original instrument (identification of a primary outcome, power calculation)[10] as well as additional variables (patient-reported outcome included, sensitivity analysis, adjustment for multiple hypothesis testing). The variables and definitions were extracted in a yes/no fashion.

Analysis. Discrepancies between reviewers were adjudicated by conference between MSP and AHR. Pooled interoperator reliability calculated using the kappa statistic (0.66) indicated substantial agreement[12]. A novel quality scale (normalized from 0 to 10) was constructed from the modified Jadad scale by assigning 1 point for each criterion a study met. Associations between categorical variables were calculated using chi-square testing. The continuous quality scale was compared to year and to journal with an ANOVA model, and to industry funding using the independent samples t-test. All analyses were performed using SPSS version 24.

RESULTS

Out of 3338 articles, 96 were selected for analysis (39 reports in *Annals of the Rheumatic Diseases*, 18 in *Rheumatology*, and 39 in *Arthritis and Rheumatology*). The numbers of publications in each year were 19 (1998), 41 (2008), and 36 (2018). Common conditions included RA (33%), seronegative spondyloarthritis and psoriatic arthritis (14%), osteoarthritis (11%), SLE (9%), vasculitis or myositis (6%), systemic sclerosis (5%), Sjögren syndrome (4%), osteoporosis (4%), and pediatric diseases (3%).

Eighty-two studies were primary analyses of RCT, and 14 were secondary analyses. Among primary analyses (Table 2), articles frequently identified a primary outcome (93%), used an intention-to-treat analysis (82%), and reported a double-blind design (79%). Quality metrics least frequently reported were adjustment for multiple hypothesis testing (16%), performance of a sensitivity analysis (22%), and appropriate reporting of allocation concealment (37%).

Over time (1998–2008–2018), articles reporting primary analyses were less likely to adequately report withdrawals and dropouts (100% vs 82% vs 60%, respectively; $p < 0.01$), report a p value < 0.05 for the primary outcome measure (80% vs 52% vs 37%; $p = 0.02$), or compare 2 active therapies (44% vs 12% vs 13%; $p = 0.01$). This decrease in comparative efficacy research was attributable to changes among industry trials (43% vs 8% vs 11%; $p = 0.01$). Articles were more likely to evaluate biologic therapy over time (11% vs 38% vs 83%; $p < 0.01$) and to report adequate randomization procedures (39% vs 29% vs 60%; $p = 0.04$). The majority of trials received industry funding (79%); changes over time were not statistically significant (78% vs 74% vs 87%; $p = 0.43$). Industry-sponsored trials were more likely to evaluate a biologic therapy (57% vs 18%; $p < 0.01$) and to report double-blinding (86% vs 53%; $p < 0.01$), patient-reported

Table 1. Modified Jadad scale for evaluation of methodologic quality (modified from Jadad, *et al*[10] and Hill, *et al*[6]).

Quality Metric	Outcome
Adequate blinding	Double-blinding reported
Adequate randomization	Any unbiased process for randomization described (e.g., random number table, computer generation, adaptive randomization, etc.)
Allocation concealment	Methods to avoid unblinding described (e.g., central allocation, third party distribution, numbered or coded bottles, sealed envelopes)
Adequate description of withdrawals and dropouts	Participants who did not complete the observation period or who were not included in the analysis are clearly described
Identification of a primary outcome	Primary outcome explicitly stated
Patient-reported outcome included	Validated patient-reported outcomes (e.g., HAQ-DI) or some type of patient-reported visual analog scale included
Power calculation	Study power and alpha error explicitly stated prior to conducting the trial (post-hoc power calculations not sufficient)
Sensitivity analysis	Sensitivity analysis performed that included some alteration of the underlying assumptions of the trial (per-protocol analysis not sufficient)
Adjustment for multiple hypothesis testing	Process used for retaining error rate over multiple comparisons (e.g., Bonferroni or sequential hypothesis testing)
Intention-to-treat (ITT) analysis	ITT or modified ITT performed (e.g., all participants who were randomized and received at least 1 treatment were included in analysis)

Table 2. Characteristics of primary randomized controlled trials (RCT) in the rheumatologic literature and the proportion that met RCT quality metrics (n = 82). Data are percentages.

Characteristics		Year of Publication				Industry Funding		
Variable	Total	1998	2008	2018	p	No	Yes	p
Comparative efficacy research*	20	44	12	13	0.01	29	17	0.25
Evaluated biologic therapy*	49	11	38	83	< 0.01	18	57	< 0.01
Reported p < 0.05 for primary outcome*	51	80	52	37	0.02	60	49	0.45
Adequate blinding	79	67	79	87	0.25	53	86	< 0.01
Adequate randomization	43	39	29	60	0.04	53	40	0.34
Allocation concealment	37	28	35	43	0.55	41	35	0.66
Adequate description of withdrawals	78	100	82	60	< 0.01	88	75	0.25
Identification of a primary outcome	93	83	91	100	0.09	88	94	0.43
Patient-reported outcome included	70	67	77	63	0.50	41	77	< 0.01
Power calculation	72	67	74	73	0.85	59	75	0.18
Sensitivity analysis	22	6	27	27	0.16	24	22	0.86
Adjusted for multiple hypothesis testing	16	11	12	23	0.37	12	17	0.60
Intention-to-treat analysis	82	78	79	87	0.67	65	86	0.04

* Variables not included in the modified Jadad scale.

outcome measures (77% vs 41%; p < 0.01), or intention-to-treat analyses (86% vs 65%; p = 0.04).

The average score on the quality scale was 6.56. There were no significant differences in the overall quality among journals (6.18 for *Annals of Rheumatic Diseases*, 5.19 for *Rheumatology*, and 6.52 for *Arthritis and Rheumatology*; p = 0.14). The quality score increased numerically by year (6.05, 6.50, 6.93), but this was not significant (p = 0.39). Industry funding was associated with a higher score (6.75 vs 5.82 for non-industry; p < 0.01). Secondary analyses of RCT had a lower quality score than primary analyses (3.65 vs 6.56; p < 0.01). There were no significant differences in quality between studies that reported a p value < 0.05 (6.50 vs 6.80; p = 0.51), studies on biologic therapies vs non-biologic therapies (6.18 vs 6.09; p = 0.85), or on studies with more or less than 200 patients (6.67 vs 5.92; p = 0.16).

DISCUSSION

Over time, high-impact rheumatology journals published more RCT that assessed biologic therapy and performed adequate randomization. Adequate description of withdrawals and dropouts, statistically significant primary endpoints, and comparative effectiveness research became less frequent. Industry trials were more likely to use double-blinding, patient-reported outcome measures, and intention-to-treat analysis, and had a significantly higher overall quality score than non-industry trials.

These results extend previous work to a broad range of disease states and the modern era[6,9]. Both overall quality and randomization have improved over time. Contrasting with these positive trends, fewer trials adequately described dropouts and withdrawals. This makes it difficult to assess the risk of attrition bias[13]. Surprisingly, fewer trials met their primary endpoint over time. This may be attributable to requirements that trials be listed on ClinicalTrials.gov, which may have encouraged the publication of negative studies[14,15].

This investigation confirms a high rate of industry sponsorship of RCT. As with RA trials, industry sponsorship was associated with higher rates of adequate blinding and intention-to-treat analysis[9]. Industry-sponsored trials also included patient-reported outcome measures at a higher rate[16]. Similar to studies in RA, we did not observe a significant association between industry funding and positive primary outcome measures[8]. This association has been observed in other fields of medicine[17], and our study may have been underpowered to detect a meaningful difference.

The decline in publication of comparative efficacy research — from nearly half of trials in 1998 to one in 8 by 2018 — appears to be driven by industry funding. This may be appropriate early in drug development, but clinicians need trials comparing the relative effectiveness of therapies to guide clinical decision making. Such trials have started to emerge in RA[18,19], but remain lacking for many disease states. Moreover, in many conditions, off-patent therapies without industry sponsorship have never been compared[20]. If industry funding and incentives drive our research agenda, valuable questions may remain unanswered.

This report has a number of limitations. The exclusion of high-impact general medicine journals and lower-impact rheumatology journals may bias our results. Although we assessed almost 100 articles, this study may be underpowered to detect significant differences over time or between funding sources. Finally, the reporting of quality metrics may not reflect actual practice[21], and other important quality metrics, such as discrepancies between FDA analysis and published analysis[22], were not included.

Trends toward improvement were observed in multiple specific metrics over time. Studies that received industry funding had higher rates of individual quality metrics and

higher quality overall. The decrease of comparative efficacy research and its effect on the practice of rheumatology deserves further investigation.

REFERENCES

1. Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the gold standard — lessons from the history of RCTs. N Engl J Med 2016;374:2175-81.
2. Kyriacou DN, Lewis RJ. Confounding by indication in clinical research. JAMA 2016;316:1818-9.
3. Manson JE, Shufelt CL, Robins JM. The potential for postrandomization confounding in randomized clinical trials. JAMA 2016;315:2273-4.
4. Ioannidis JPA. The importance of predefined rules and prespecified statistical analyses: do not abandon significance. JAMA 2019;321:2067-8.
5. Sismondo S. How pharmaceutical industry funding affects trial outcomes: causal structures and responses. Soc Sci Med 2008;66:1909-14.
6. Hill CL, LaValley MP, Felson DT. Secular changes in the quality of published randomized clinical trials in rheumatology. Arthritis Rheum 2002;46:779-84.
7. Karassa FB, Tatsioni A, Ioannidis JPA. Design, quality, and bias in randomized controlled trials of systemic lupus erythematosus. J Rheumatol 2003;30:979-84.
8. Cohen N, Lavie RG, Manor Y, Mimouni M, Furst DE, Amarilyo G. Questioning a publication bias between industry-funded and non-industry-funded randomized controlled trials on biological and small molecule therapy for rheumatoid arthritis. Semin Arthritis Rheum 2020;50:7-11
9. Khan NA, Lombeida JI, Singh M, Spencer HJ, Torralba KD. Association of industry funding with the outcome and quality of randomized controlled trials of drug therapy for rheumatoid arthritis. Arthritis Rheum 2012;64:2059-67.
10. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials 1996; 17:1-12.
11. Braun T, Glänzel W, Schubert A. A Hirsch-type index for journals. Scientometrics 2006;69:169-73.
12. Viera AJ, Garrett JM. Understanding interobserver agreement: the kappa statistic. Fam Med 2005;37:360-3.
13. Bell ML, Kenward MG, Fairclough DL, Horton NJ. Differential dropout and bias in randomised controlled trials: when it matters and when it may not. BMJ 2013;346:e8668.
14. Ross JS, Tse T, Zarin DA, Xu H, Zhou L, Krumholz HM. Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. BMJ 2012;344:d7292.
15. Zarin DA, Tse T, Williams RJ, Carr S. Trial reporting in ClinicalTrials.gov — the final rule. N Engl J Med 2016; 375:1998-2004.
16. Orbai AM, Bingham CO. Patient reported outcomes in rheumatoid arthritis clinical trials. Curr Rheumatol Rep 2015;17:501.
17. Bhandari M, Busse JW, Jackowski D, Montori VM, Schünemann H, Sprague S, et al. Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials. CMAJ 2004;170:477-80.
18. Gabay C, Emery P, van Vollenhoven R, Dikranian A, Alten R, Pavelka K, et al. Tocilizumab monotherapy versus adalimumab monotherapy for treatment of rheumatoid arthritis (ADACTA): a randomised, double-blind, controlled phase 4 trial. Lancet 2013;381:1541-50.
19. Fleischmann R, Pangan AL, Song IH, Mysler E, Bessette L, Peterfy C, et al. Upadacitinib versus placebo or adalimumab in patients with rheumatoid arthritis and an inadequate response to methotrexate: results of a phase 3, double-blind, randomized controlled trial. Arthritis Rheum 2019;71:1788-800.
20. Aggarwal R, Oddis CV. Therapeutic advances in myositis. Curr Opin Rheumatol 2012;24:635-41.
21. Hill CL, LaValley MP, Felson DT. Discrepancy between published report and actual conduct of randomized clinical trials. J Clin Epidemiol 2002;55:783-6.
22. Amarilyo G, Furst DE, Woo JM, Li W, Bliddal H, Christensen R, et al. Agreements and discrepancies between FDA reports and journal papers on biologic agents approved for rheumatoid arthritis: a meta-research project. PLoS ONE 2016;11:e0147556.