**RMD Open**

Rheumatic & Musculoskeletal Diseases

ORIGINAL RESEARCH

# Missing data in randomised controlled trials of rheumatoid arthritis drug therapy are substantial and handled inappropriately

Nasim A Khan,[1] Karina D Torralba,[2] Fawad Aslam ⓘ [3]

► Additional supplemental material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/rmdopen-2021-001708).

For 'Presented at statement' see end of article.

For numbered affiliations see end of article.

**Correspondence to**
Dr Fawad Aslam;
fawadaslam2@gmail.com

## ABSTRACT

**Objectives**  To analyse the amount, reporting and handling of missing data, approach to intention-to-treat (ITT) principle application and sensitivity analysis utilisation in randomised clinical trials (RCTs) of rheumatoid arthritis (RA). To assess the trend in such reporting 10 years apart (2006 and 2016).

**Methods**  Parallel group drug therapy RA RCTs with a clinical primary endpoint.

**Results**  176 studies enrolling a median of 160 (IQR 62–339) patients were eligible. In terms of actual analysis: 81 (46%) RCTs conducted ITT, 42 (23.9%) conducted modified ITT while 53 (30.1%) conducted non-ITT analysis. Only 58 of 97 (59.8%) RCTs reporting an ITT analysis actually performed it. The median (IQR) numbers of participants completing the trial and included in analysis for primary outcome were 86% (74%–91%) and 100% (97.1%–100%), respectively. 53 (32.7%) and 65 (40.1%) RCTs had >20% and 10%–20% missing primary outcome data, respectively. Missing data handling was unreported by 58 of 171 (33.9%) RCTs. When reported, vast majority used simple imputation methods. No significant trend towards improved reporting was seen between 2006 and 2016. Sensitivity analysis numerically improved from 2006 to 2016 (14.7% vs 21.4%).

**Conclusions**  There is significant discrepancy in the reported and the actual performed analysis in RA drug therapy RCTs. Nearly one-third of RCTs had >20% missing data. The reporting and methods of missing data handling remain inadequate with high usage of non-preferred simple imputation methods. Sensitivity analysis utilisation was low. No trend towards better missing data reporting and handling was seen.

## Key messages

**What is already known about this subject?**
► Missing data are common in randomised controlled trials (RCTs).
► Handling of missing data has implication for the results of RCTs.
► Little is known about missing data reporting and its handling in RCTs of rheumatoid arthritis (RA) treatment.

**What does this study add?**
► Missing data are a significant issue in RCTs in RA.
► More importantly, missing data handling relies heavily on error-prone methods of simple imputation with no trend of improvement.
► Sensitivity analysis usage remains very low.

**How might this impact on clinical practice or future developments?**
► Increased awareness and requirement of using the preferred imputation methods is needed.
► Journals should require reporting of these methods as well as sensitivity analysis.

## INTRODUCTION

Randomised controlled trial (RCT) is the gold standard for comparing treatments and determining efficacy. Randomization reduces bias by balancing distribution of known and unknown confounding factors.[1] The intention-to-treat (ITT) principle preserves randomisation by evaluating all participants as original randomly assigned, regardless of deviations from randomised treatment and actual interventions.[1–3]

Most trials have missing data due to protocol deviation, non-adherence, trial non-completion and other issues.[3] The ideal analysis, a true ITT analysis, has no missing data.[2] Per-protocol (PP) or complete-case (CC) analysis include only subjects who adhere to the assigned treatment and complete the study. These approaches generally introduce bias. Missing data undermine the RCT by introducing confounding, unbalancing baseline characteristics and compromising the internal validity.[4 5] Moreover, compliant subjects do better than non-compliant ones, irrespective of active treatment or placebo status.[6] Thus, restricting analysis to those who completed a trial as PP can lead to misleading

estimates of treatment efficacy.[7] Non-ITT methods may simply reflect confounding from differences of demographics, prognostic features or subject characteristics.[5 8] The ITT approach reflects the real-world settings where subjects drop out, switch treatments or are non-compliant. Although ITT may underestimate the true difference, via dilution, it still represents the best unbiased estimate.[9 10] However, it must be noted that ITT analysis only matters if there are missing data, otherwise ITT and CC are similar.

The term ITT is often used incorrectly[11 12] as RCTs with post-randomisation exclusions are still described as using 'ITT analysis'.[13] The term modified ITT (mITT) reflects analyses where participants are excluded for different reason(s) post-randomisation.[14] mITT analysis usage has increased.[15] Consequently, the Consolidated Standards of Reporting Trials (CONSORT) statement requires a clear description of subject randomisation and analysis rather than simply stating the method.[1]

Missing data may seriously compromise inferences ascertained from RCTs,[16 17] limit the ability to draw conclusions[18] or lead to incorrect inferences about drug safety.[19 20] Missing data can be handled by imputation. Simple imputation (SI) methods include last observation carried forward (LOCF), baseline observation carried forward, non-response imputation (NRI) or poor response imputation. SI assumes lack of temporal change, underestimates true data variability and biases treatment effect estimates.[16 17 21] Multiple imputation, maximum likelihood-based methods such as expectation–maximisation algorithm, or equation-based methods such as full information maximum likelihood or mixed model regression for longitudinal data are generally preferred. These methods incorporate auxiliary information about the missing data and report SE and p values.[16 17 21] Any attempt at replacing missing data makes unverifiable assumptions about the distribution of the unobserved or missing data.[22 23] To test these assumption(s), sensitivity analyses with a different assumption(s) should be performed.[16 17 21 24]

Studies across disciplines have reported inappropriate use of the ITT principle, incorrect handling of missing data, utilisation of weak imputation methods and lack of sensitivity analyses.[13 24–27] Studies in rheumatoid arthritis (RA) exploring this have either been restricted to top medical journals or have also included subjects without RA.[25 27] Our study examines RA publications in all medical journals and gives temporal trends in missing data reporting and handling, before and after the publication of 2010 CONSORT statement.[1] Our goal was to identify areas of improvement and create awareness about such shortcomings.

## METHODS
### Literature search and RCT selection
MEDLINE and Cochrane Central Register of Controlled Trials databases were searched using terms "rheumatoid arthritis" and "arthritis, rheumatoid" with limits of "Clinical Trials" & years "2002–3", "2006–7", "2010–11" and "2016". Inclusion criteria were original report of a single RCT in a population with RA, parallel design, drug therapy intervention and a clinical primary endpoint. Non-English language, abstract only, non-clinical endpoint, open-label extensions, population without RA, phase I studies and non-drug intervention studies were excluded. Studies with safety as the primary outcome were also excluded as safety analyses typically focus on the subgroup of trial participants who received the intervention, although such approach may be misleading by ignoring the reasons trial participants stop treatment.[28]

### Trial selection and data acquisition
Two reviewers independently assessed each eligible RCT. Differences were resolved by consensus. Funding source, experimental agent used, study phase, study duration, study outcome, sample size calculation, participant flow description, overall subject dropout, self-reported data analysis method, actual data analysis method used (determined by study reviewers), missing data proportion and missing data handling methods were assessed.[29] An adequate description of study participant flow was defined as a clear description of the number of subjects who were randomised, number completing and not completing the study with reasons, and number analysed for the primary outcome. We assessed whether participant progress through the phases of parallel RCT (enrolment, allocation, follow-up and analysed) was presented in a flow diagram, as recommended.[1 30] Definitions are given in box 1.

---

**Box 1  Definitions for different methods of missing data handling**

**Definitions**

**ITT analysis**
► All subjects analysed after randomisation. No exclusions.

**Modified ITT analysis excluded the following[15]:**
► Subjects who did not receive any study intervention (treatment-related exclusion).
► Subjects without a baseline assessment (baseline assessment exclusion).
► Subjects without a post-baseline assessment (post-baseline exclusion).
► Subjects found non-eligible after randomisation (target condition-related exclusion).
► Exclusion of study centre(s) due to study conduct issues (centre-related exclusion).

**Non-ITT analysis included the following:**
► Performed per-protocol analysis.
► Performed a case completer analysis.
► Included exclusions not specified in the modified ITT analysis definition.
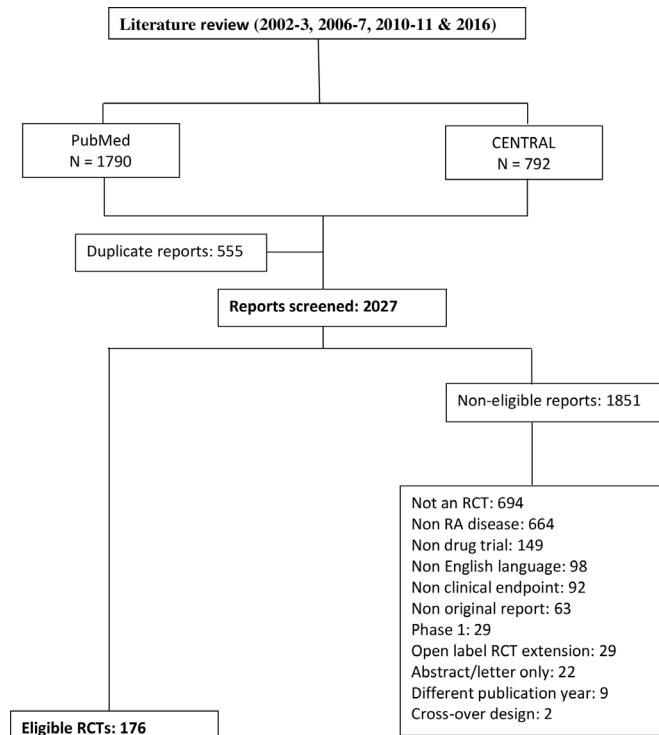
ITT, intention-to-treat.

---

Figure 1  Study selection flow diagram. RA, rheumatoid arthritis; RCT, randomised controlled trial.



**Figure 2**  Discrepancy between self-reported analysis type and actual analysis performed for handling missing data. ITT, intention-to-treat; mITT, modified ITT; N, number.

## Additional data for 2006 and 2016 trials

Additional data collected for the 2006 and 2016 studies included reporting of sample size inflation to account for anticipated dropout, mechanism of missing data, missing data in each study arm, and baseline characteristics comparison of study subjects with complete and missing data, and performance of sensitivity analysis. The 2006 and 2016 RCTs were compared for data analysis, use of ITT principle and handling of missing data.

## Statistical analysis

Categorical data are described as the number (per cent), and the continuous variables are described as the median (IQR). Comparison of RCT characteristics among trials using different analysis approaches and temporal changes from 2006 to 2016 was assessed using Pearson's $X^2$ tests, Fisher's exact test or likelihood ratio tests (for categorical variables) and Mann-Whitney U tests or Kruskal-Wallis tests (for continuous variables). SPSS V.21 was used for data analysis.

## RESULTS
### Study characteristics

The 2027 reports published in 7 years (2002–2003, 2006–2007, 2010–2011, 2016) were screened to identify 176 eligible RCTs (figure 1). These RCTs enrolled a median of 160 (IQR 62–339) patients. In terms of actual analysis: 81 (46%) RCTs conducted ITT analysis, 42 (23.9%) conducted mITT analysis, while 53 (30.1%) conducted non-ITT analysis (27 RCTs conduced CC analysis, while analysis approach for 26 RCTs was unclear). Discordance
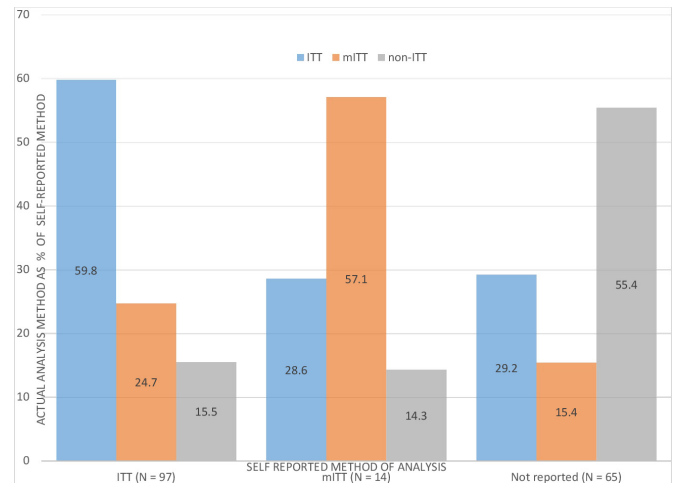
was present between the reported and the actual analysis performed (figure 2). Only 58 of 97 (59.8%) RCTs reporting an ITT analysis actually performed it. Of 42 (23.9%) studies which performed mITT, only 8 (19.0%) described their analysis as mITT. RCT participant flow diagram was unreported or inadequate in 66.5% RCTs: mostly resulted from non-reporting of participants analysed for the primary outcome. A median of 86% participants (IQR 74%–91%) remained enrolled until trial completion, although 14 RCT reports did not specify number of patients completing the trial. Fifty-three (32.7%) and 65 (40.1%) of RCTs had >20% and 10%–20% missing primary outcome data, respectively. A median of 100% (IQR 97.1%–100%) patients were included in the efficacy analysis of primary trial outcome; 13 RCT reports did not specify number of patients analysed for the primary outcome. Most common missing data handling method was LOCF (69 (39.2%)), followed by NRI (49 (27.8%)) and imputing missing value by interpolation (21 (11.9%)). Only three study RCTs used preferred methods to deal with missing data for primary outcome analysis: two used linear mixed model regression,[31 32] while one used expectation–maximisation algorithm to impute missing data.[33] In addition, two RCTs used equation-based methods for sensitivity analyses.

Missing data handling was not reported by 58 of 171 (33.9%) RCTs; 5 RCTs had no missing data. Twelve RCTs using ITT analysis did not specify missing data handling method.

### RCT characteristics associated with analysis type

Table 1 shows the characteristics of RCTs according to the analysis performed. Industry-funded RCTs were more likely to perform ITT or mITT than non-profit-funded RCTs. Trials reporting ITT or mITT analysis were more likely to enrol higher number of participants, be multi-centre, describe sample size calculation, adequately report participant flow, describe missing data handling methods and include higher proportion of participants

**Table 1** Characteristics of the study RCTs according to the actual analysis type

| Characteristics | ITT (N=81) | mITT (N=42) | Non-ITT (N=53) | P value |
|---|---|---|---|---|
| Funding source | | | | <0.001 |
| Industry, full or partial | 59 (72.8) | 31 (73.8) | 14 (26.4) | |
| Non-profit or unspecified | 22 (27.2) | 11 (26.2) | 39 (73.6) | |
| Experimental agent | | | | 0.001 |
| Traditional DMARD | 13 (16.0) | 5 (11.9) | 12 (22.6) | |
| Biological DMARD | 30 (37) | 27 (64.3) | 16 (30.2) | |
| Small molecule | 8 (9.9) | 2 (4.8) | 0 (0.0) | |
| Others | 30 (37.0) | 8 (19.0) | 25 (47.2) | |
| Study phase | | | | 0.086 |
| Phase 2 | 20 (24.7) | 8 (19.0) | 5 (9.4) | |
| Non-phase 2/unspecified | 59 (75.3) | 34 (81.0) | 48 (90.6) | |
| Study centres, multiple | 63 (77.8) | 42 (100) | 19 (35.8) | <0.001 |
| Study duration, months | 6 (3–12) | 9 (3–12) | 6 (3–12) | 0.481 |
| Pre-study sample size calculation reported* | 42 (68.9) | 29 (85.3) | 24 (50) | 0.003 |
| Participant flow diagram | | | | 0.001 |
| Adequate | 24 (29.6) | 24 (57.1) | 11 (20.8) | |
| Inadequate | 26 (32.1) | 9 (21.4) | 12 (22.6) | |
| Not reported | 31 (38.3) | 9 (21.4) | 30 (56.6) | |
| Adequate follow-up description | 63 (77.8) | 40 (95.2) | 33 (62.3) | <0.001 |
| Number of patients enrolled | 165 (70–339) | 317 (160–549) | 70 (41–160) | <0.001 |
| Percentage of patients completing RCT† | 85.8 (70.4–89.3) | 82.9 (76–88.9) | 89 (72.8–92.3) | 0.232 |
| Percentage of missing data† | | | | 0.341 |
| 0–5 | 8 (10.5) | 3 (7.1) | 6 (13.6) | |
| >5–10 | 10 (13.2) | 5 (11.9) | 12 (27.3) | |
| >10–20 | 32 (42.1) | 20 (47.5) | 13 (29.5) | |
| >20 | 26 (34.2) | 14 (33.3) | 13 (29.5) | |
| Percentage of patients analysed for the primary outcome‡ | 100 (100–100) | 98.7 (97.1–99.6) | 90.8 (86.3–97.7) | <0.001 |
| Missing data handling method described§ | 64 (84.2) | 38 (90.5) | 11 (20.8) | <0.001 |
| Efficacy, positive¶ | 54 (69.2) | 35 (87.5) | 29 (63.0) | 0.032 |

Values represent number (%) for categorical variables and median (25th–75th percentile) for the numerical variables.
*N=143 (33 phase 2 RCTs excluded as they have different sample size calculation considerations).
†N=162 (five RCTs performing ITT and nine RCTs performing non-ITT analysis did not clearly report number of patients completing the trial).
‡N=163 (13 RCTs performing non-ITT analysis did not clearly report number of patients analysed for the primary outcome).
§N=169 (all enrolled patients reported to complete seven RCTs with no missing data).
¶Positive efficacy defined as statistically significant outcome for the primary outcome favouring the experimental intervention for superiority trials or meeting the threshold for equivalence in the non-inferiority trials. N=164, 12 RCTs were strategy trials with no intervention designated as experimental.
DMARD, disease-modifying anti-rheumatic drug; ITT, intention-to-treat; mITT, modified ITT; N, number; RCT, randomised controlled trial.

in the primary outcome analysis. However, there was no difference in proportion of participants completing the trials among the three groups. Among the mITT trials, the top two reasons for post-randomisation exclusions were no study treatment received (n=27) and no post-baseline assessment (n=17). Among non-ITT trials, the top two exclusions were study protocol deviation (n=22) and outcome absence at study completion (n=7).

## Temporal trends in missing data reporting from 2006 to 2016

These studies, 10 years apart, were compared to assess trends in missing data reporting and handling. A total of 76 studies (34 published in 2006 and 42 published in 2016) were eligible. Baseline characteristics are given in online supplemental table 1 and were similar except for more RCTs in 2016 having a placebo arm compared with

2006. There were four RCTs of small molecules in 2016 and none in 2006.

Table 2 provides a detailed description of data reporting in 2006 and 2016 RCTs. Of the 65 non-phase 2 RCTs, only 43 (66.2%) adequately reported sample size calculation. Thirteen (20.0%) RCTs inflated sample size by median of 20% (IQR 12.5%–25.5%) anticipating loss to follow-up. RCTs reporting sample size inflation numerically increased from 2006 (4 (12.9%)) to 2016 (9 (26.5%), p=0.167). Phase 2 trials were excluded as they have different sample size considerations.[34] A significant increase in adequate participant flow description occurred between 2006 and 2016.

Self-reporting of ITT analysis declined from 2006 to 2016, while the actual analysis type was unchanged (table 2). Notably, 49 (68.1%) RCTs had >10% and 18 (25%) had >20% missing primary outcome data on completion. Four (5.5%) RCTs lacked information about subjects completing the RCT. The 2016 RCTs showed a trend toward higher percentage of participants completing the trials (table 2). Nearly one-third of RCTs did not report missing outcome data handling method: unimproved from 2006 to 2016. LOCF and NRI were the most common approaches to handle missing data, although a trend in decline of LOCF use was noted. No 2006 study and two 2016 studies used the preferred missing data handling methods.

Ten of 14 (17.5%) RCTs using sensitivity analysis reported results. Alternate imputation was the most used (seven RCTs) method for sensitivity testing. Sensitivity analysis usage improved from 2006 to 2016 (14.7% vs 21.4%). Only one trial each reported mechanism of missing data[35] and baseline characteristics of trial completers and non-completers.[36]

Online supplemental table 2 gives the differential attrition between study arms. In 17 of 68 (25%) RCTs which reported number of patients completing each study arm, the comparator arm had >10% loss of patients compared with experimental arm, while three had >10% loss in the experimental arm compared with the comparator arm.

## DISCUSSION

Our review of RCTs of RA drug therapy showed that missing primary outcome data are a major issue. Reassuringly, however, almost all patients were included in the final primary outcome analysis (mostly by using SI methods for missing data). RCTs in 2016 reported the highest number of sensitivity analyses (20.5%). The 2006 and 2016 comparisons did not show any significant improvement in the parameters of missing data reporting, handling and sensitivity analysis.

Missing data handling methods were not reported by 38.0% RCTs. This is, however, improved from a previous report where 76.5% did not do so.[25] When reported, non-preferred SI methods (LOCF and NRI) were used. Limited missing data handling reporting and preference for SI methods, when 67.0% of RCTs have >10% missing

data, are concerning. To determine how much of missing data is problematic, some authors have reported a '5 and 20' rule where <5% missing data lead to minimal bias, while >20% pose a serious validity issue.[37 38]

Table 3 summarises missing data reporting in four studies, including our study. The 2005 study looked at RCT data published in high-impact journals pertaining to RA, osteoarthritis and osteoporosis.[25] The 2014 paper is of RCTs published in top four medical journals irrespective of medical condition.[24] The 2016 study looking at phase 3 RA RCTs published in top rheumatology and medical journals is also included.[27] The latter report's RCTs may have some overlap with our study.

Misrepresentation of ITT analysis was common with only 59.8% of self-reported ITT being true ITT analysis. Overall, 46.0% actually performed an ITT analysis, which is slightly higher as compared with some other reports of 37%–40%[24 26] and improved from older RA studies at 7.4%.[25] mITT analysis excludes subjects which violate the ITT principles. However, the number of patients actually analysed for the primary outcome was high: 100% for ITT (as expected) and 98.8% for mITT. Thus, the absolute patient exclusions in mITT RCTs are small. Exclusion of such subjects, especially with small numbers, does not always violate the ITT principle.[9 15 27 39 40] Recently, the treatment effect estimation in RA trials was shown to be similar between mITT and ITT analysis.[39]

For the 2006 and 2016 studies, only 20% adjusted the sample size for dropouts despite a quarter having >20% missing outcome data. Sample size inflation accounts for dropout data loss but it does not remedy missing data and any differences between the observed and missed data.[16] Differential attrition can be a source of bias in RCTs.[37] In our analysis, >10% attrition in the placebo or comparator arm versus the experimental arm was seen in 25% of trials which may artificially increase treatment effect in favour of experimental intervention if using SI. The use of SI in differential attrition biases results.[41] A sensitivity analysis is required to assess the impact of such possible bias, but was performed in <20% trials. Impact of differential attrition in RA is unexplored. A small study of musculoskeletal RCTs found differential attrition rates ranging from 1% to 14%[42] without impacting the results. The authors suggested reporting of baseline characteristics of both the analysed and randomised subjects.

Only one study reported a missing data mechanism.[35] Another review reported this number at 7.8% RCTs.[27] Three missing data mechanisms exist: missing completely at random (MCAR), missing at random (MAR) and missing not at random (MNAR) and are discussed elsewhere.[43 44] It is important to obtain data on study completers and the non-completers. It is possible to distinguish between MCAR and MAR by studying the available data. If the completers and the non-completers differ (for example, if older participants have more missing data than younger participants, the data are not MCAR but possibly MAR). It is not possible to distinguish between MAR and MNAR by studying the available data.[24] Only one RCT in our

**Table 2** Comparison of data reporting including missing data and its handling in 2006 and 2016 RCTs

| Characteristics | All (N=76) | Study year 2006 (N=34) | 2016 (N=42) | P value |
|---|---|---|---|---|
| Total patients | | | | 0.913 |
| Median (IQR) | 162 (74–326) | 163 (53–367) | 159 (79–311) | |
| Range | 21–1404 | 21–1093 | 43–1404 | |
| Patient per cent completing study* | | | | 0.089 |
| Median (IQR) | 86.4 (79.5–91.6) | 84.5 (72.8–91.8) | 87.9 (81.7–91.6) | |
| Range | 38.5–100 | 40.9–95.8 | 38-5–100 | |
| Patient per cent analysed for primary outcome† | | | | 0.962 |
| Median (IQR) | 99.7 (97.1–100) | 99.5 (97.3–100) | 99.7 (97–100) | |
| Range | 38.5–100 | 85.8–100 | 38.5–100 | |
| Patient in experiment arm(s) | | | | 0.81 |
| Median (IQR) | 100 (40–217) | 99 (29–252) | 103 (40–177) | |
| Range | 11–1286 | 11–751 | 19–1286 | |
| Patient per cent completing study in experimental arm(s)‡ | | | | 0.405 |
| Median (IQR) | 87.7 (81.3–93.1) | 86.4 (81.7–93) | 88.6 (80.7–94.4) | |
| Range | 37.5–100 | 45.2–100 | 37.5–100 | |
| Patient per cent analysed for primary outcome in experimental arm(s)§ | | | | 0.721 |
| Median (IQR) | 100 (99.6–100) | 100 (99.3–100) | 100 (99.6–100) | |
| Range | 37.5–100 | 88.6–100 | 37.5–100 | |
| Patient in comparator arm¶ | | | | 0.983 |
| Median (IQR) | 56 (30–132) | 68 (25–141) | 53 (31–121) | |
| Range | 10–532 | 10–531 | 15–532 | |
| Patient per cent completing study in comparator arm** | | | | 0.1 |
| Median (IQR) | 84.4 (73–91.5) | 81.9 (64.6–90.6) | 84.9 (78.1–93.5) | |
| Range | 23.9–100 | 30–93.7 | 23.9–100 | |
| Patient per cent analysed for primary outcome in comparator arm | | | | 0.507 |
| Median (IQR) | 100 (97.2–100) | 100 (96.1–100) | 100 (97.6–100) | |
| Range | 39.5–100 | 84.5–100 | 39.5–100 | |
| Adequate sample size calculation description†† | 43 (66.2) | 19 (61.3) | 24 (70.6) | 0.429 |
| Sample size inflation anticipating follow-up loss†† | 13 (20.0) | 4 (12.9) | 9 (26.5) | 0.167 |
| Adequate follow-up description | 64 (84.2) | 28 (82.4) | 36 (85.7) | 0.689 |
| Flow diagram for patient follow-up | | | | 0.016 |
| Adequate | 31 (40.8) | 17 (50.0) | 14 (33.3) | |
| Inadequate | 18 (23.7) | 3 (8.8) | 15 (35.7) | |
| Not reported | 27 (35.5) | 14 (41.2) | 13 (31.0) | |
| Amount of missing outcome data* | | | | 0.042 |
| <5% | 10 (13.9) | 1 (3.1) | 9 (22.5) | |
| 5.1%–10% | 13 (18.1) | 8 (25.0) | 5 (12.5) | |
| 10.1%–20% | 31 (43.1) | 13 (40.6) | 18 (45.0) | |
| >20% | 18 (25) | 10 (31.3) | 8 (20.0) | |

Continued

**Table 2** Continued

| Characteristics | All (N=76) | Study year 2006 (N=34) | 2016 (N=42) | P value |
|---|---|---|---|---|
| Missing data handling method given‡‡ | | | | 0.75 |
| N (%) | 48 (65.8) | 23 (67.6) | 25 (64.1) | |
| Self-reported analysis term used | | | | 0.05 |
| ITT | 33 (43.4) | 20 (58.8) | 13 (43.4) | |
| Modified ITT | 6 (7.9) | 2 (5.9) | 4 (9.5) | |
| None | 37 (48.7) | 12 (35.3) | 25 (59.5) | |
| Actual analysis performed | | | | 0.266 |
| ITT | 29 (38.2) | 12 (35.3) | 17 (40.5) | |
| Modified ITT | 20 (26.3) | 11 (32.4) | 9 (21.4) | |
| Completer/inadequate | 15 (19.7) | 4 (11.8) | 11 (26.2) | |
| Unclear | 12 (15.8) | 7 (20.6) | 5 (11.9) | |
| Top methods to impute missing data | | | | |
| LOCF, N (%) | 29 (38.2) | 17 (50.0) | 12 (28.6) | 0.056 |
| NRI, N (%) | 22 (28.9) | 10 (29.4) | 12 (28.6) | 0.936 |
| Simple imputation, N (%) | 11 (14.5) | 6 (17.6) | 9 (21.4) | 0.68 |
| Performed sensitivity analysis | | | | 0.452 |
| N (%) | 14 (18.4) | 5 (14.7) | 9 (21.4) | |

*N=72, number of patients completing the trial unclear for four RCTs.
†N=73, number of patients included in primary outcome analysis unclear for three RCTs.
‡N=69, number of patients in experimental arm(s) completing the trial unclear for seven RCTs.
§N=70, number of patients in experimental arm(s) included in primary outcome analysis unclear for six RCTs.
¶N=75, one RCT had three arms with different doses of experimental intervention, no comparator group.
**N=68, number of patients in comparator arm completing the trial unclear for seven RCTs and one RCT did not have a comparator group.
††N=65, excluded 11 phase 2 RCTs as they have different sample size calculation considerations.
‡‡N=73, all enrolled patients reported to complete three RCTs with no missing data.
ITT, intention-to-treat; LOCF, last observation carried forward; N, number; NRI, non-response imputation; RCTs, randomised controlled trials.

**Table 3** Trend of salient characteristics over time in RA studies and comparison with a study from top four medical journals

| Characteristic | Time period of included studies/(reference) (publication year) | | | | |
|---|---|---|---|---|---|
| | 1994–2003[25] (2005) | 2006 (Current) | 2016 (Current) | 2013*[24] (2014) | 2008–2013*[27] (2016) |
| Overall ITT analysis performed, % | 7.4 | 35.3 | 38.1 | 40.0 | NA† |
| Overall PP analysis performed, % | 59.3 | 11.7 | 28.5 | NA | NA |
| Preferred missing data handling methods used, % | 1.2 | 2.7 | 4.9 | 27.3 | 1.9 |
| Dropout in each study arm given, % | 69.0 | 82.3 | 85.7 | 97.0 | NA |
| Missing data handling given, % | 23.5 | 61.1 | 48.8 | 36.0 | 94.1 |
| LOCF used to handle missing data‡, % | 52.6 | 41.7 | 22.7 | 12.0 | 56.8 |
| Missing mechanism given, % | NA | 2.9 | 0.0 | NA | 7.8 |
| Sensitivity analysis performed, % | 11.1 | 13.9 | 20.5 | 37.0 | 27.0 |
| Comparison of completers/non-completers, % | 16.7 | 0.0 | 2.3 | 11.6 | NA |

*Studies involving top medical journals.
†86% stated that ITT was used but no information on those that actually performed ITT.
‡LOCF was the most common method used across all years for imputation except 2016 (Ibrahim *et al*) where NRI was most frequent (74.5%) but NRI and LOCF were used simultaneously as well; in the 1994–2003 and 2013 studies, complete case analysis (59.3% and 45.0%, respectively) was the most common method to handle missing data.
ITT, intention-to-treat; LOCF, last observation carried forward; NA, not available; PP, per-protocol; RA, rheumatoid arthritis.

analysis reported data on completers and dropouts.[36] In one study, this was 11.6%[24] and 16.7% in older RA trials.[25] Generally, MAR is a reasonable starting point.[16 21] In our study, LOCF was the most common method (53.7%) to handle missing data in the 95 RCTs that used ITT or mITT. This is different from a prior report of RA trials where NRI was the most common (75%) imputation method.[27] The US Food and Drug Administration as well as the European Medicines Agency used to recommend LOCF method, assuming that it could provide conservative effect estimates. This may have contributed to its widespread use. However, it generally introduces bias even if data are missing completely at random.[45] In an RCT with longitudinal data, it is generally recommended to use a mixed model analysis excluding any systematic treatment effect at baseline. This is called 'constrained analysis'.[46] It is also recommended to adjust for baseline value of the outcome variable.[47] In particular, this is a good way to handle missing data in an RCT with longitudinal data.[48] None of our study RCTs used this approach.

A sensitivity analysis to test a range of plausible mechanisms responsible for the missing data is generally recommended.[7] A sensitivity analysis should be performed with a different mechanism of missingness.[44] Agreement of different sensitivity analyses is reassuring. Sensitivity analysis was reported by 17.5% RCTs in our review. These numbers are similar (21%–27%) for recent studies reviewing sensitivity analysis in RA trials[26 27] but lower than numbers (37% and 27%) from top medical journals.[24 27]

Our study has limitations. The RCTs in our study were time and language restricted. Results may not be applicable to non-RA trials. Some of the deficiencies may result from poor reporting (not mentioning relevant details) and not necessarily poor execution.[49] On the other hand, we have looked at a large number of trials over several years. We have compared the trend in reporting at two time points separated by a decade. We have not restricted our analysis to high-impact journals increasing representation.

Major gaps exist between statistical analysis recommendations and actual reporting. Many papers and recommendations have addressed these deficiencies from a statistical perspective.[16 23 44 50 51] These issues are seemingly better handled in papers published in high-ranking journals signifying greater attention to missing data handling during peer review might improve such reporting.

In conclusion, our study has found significant shortcomings in the reporting and handling of missing data in RA trials: there is wide variability in reporting and performance of ITT analysis, SI methods remain popular despite shortcomings, sensitivity analysis usage is low, there is inadequate explicit reporting of participants with missing outcomes and methods used to handle them, baseline comparisons of completers versus dropouts are deficient. Lack of a significant trend in improved reporting 10 years apart remains a major concern.

**Author affiliations**
¹Rheumatology Section, CHI St Vincent Little Rock Diagnostic Clinic, Little Rock, Arkansas, USA
²Division of Rheumatology, Loma Linda University, Loma Linda, California, USA
³Divison of Rheumatology, Department of Internal Medicine, Mayo Clinic, Scottsdale, Arizona, USA

**ORCID iD**
Fawad Aslam http://orcid.org/0000-0001-6016-6123

## REFERENCES

1 Moher D, Hopewell S, Schulz KF, *et al*. Consort 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c869.
2 White IR, Horton NJ, Carpenter J, *et al*. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ* 2011;342:d40.
3 Liao JM, Stack CB. Annals understanding clinical research: implications of missing data due to dropout. *Ann Intern Med* 2017;166:596–8.
4 Marston L, Sedgwick P. Randomised controlled trials: missing data. *BMJ* 2014;349:g4656.
5 Sedgwick P. Intention to treat analysis versus per protocol analysis of trial data. *BMJ* 2015;350:h681.
6 Horwitz RI, Viscoli CM, Berkman L, *et al*. Treatment adherence and risk of death after a myocardial infarction. *Lancet* 1990;336:542–5.
7 DeMets DL, Cook T. Challenges of Non-Intention-to-Treat analyses. *JAMA* 2019;321:145–6.
8 Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 1979;301:1410–2.
9 Kunz R, Guyatt G. Which patients to include in the analysis? *Transfusion* 2006;46:881–4.
10 Gupta SK. Intention-To-Treat concept: a review. *Perspect Clin Res* 2011;2:109–12.

11 Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999;319:670–4.

12 Gravel J, Opatrny L, Shapiro S. The intention-to-treat approach in randomized controlled trials: are authors saying what they do and doing what they say? *Clinical Trials* 2007;4:350–6.

13 Herman A, Botser IB, Tenenbaum S, *et al*. Intention-To-Treat analysis and accounting for missing data in orthopaedic randomized clinical trials. *J Bone Joint Surg Am* 2009;91:2137–43.

14 Montedori A, Bonacini MI, Casazza G, *et al*. Modified versus standard intention-to-treat reporting: are there differences in methodological quality, sponsorship, and findings in randomized trials? A cross-sectional study. *Trials* 2011;12:58.

15 Abraha I, Montedori A. Modified intention to treat reporting in randomised controlled trials: systematic review. *BMJ* 2010;340:c2697.

16 Little RJ, D'Agostino R, Cohen ML, *et al*. The prevention and treatment of missing data in clinical trials. *N Engl J Med* 2012;367:1355–60.

17 O'Neill RT, Temple R. The prevention and treatment of missing data in clinical trials: an FDA perspective on the importance of dealing with it. *Clin Pharmacol Ther* 2012;91:550–4.

18 Ware JH. Interpreting incomplete data in studies of diet and weight loss. *N Engl J Med* 2003;348:2136–7.

19 Lagakos SW. Time-to-event analyses for long-term treatments--the APPROVe trial. *N Engl J Med* 2006;355:113–7.

20 Molnar FJ, Man-Son-Hing M, Hutton B, *et al*. Have last-observation-carried-forward analyses caused us to favour more toxic dementia therapies over less toxic alternatives? A systematic review. *Open Med* 2009;3:e31–50.

21 Thabane L, Mbuagbaw L, Zhang S, *et al*. A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *BMC Med Res Methodol* 2013;13:92.

22 Altman DG. Missing outcomes in randomized trials: addressing the dilemma. *Open Med* 2009;3:e51–3.

23 Cro S, Morris TP, Kenward MG, *et al*. Sensitivity analysis for clinical trials with missing continuous outcome data using controlled multiple imputation: a practical guide. *Stat Med* 2020;39:2815–42.

24 Bell ML, Fiero M, Horton NJ, *et al*. Handling missing data in RCTs; a review of the top medical journals. *BMC Med Res Methodol* 2014;14:118.

25 Baron G, Boutron I, Giraudeau B, *et al*. Violation of the intent-to-treat principle and rate of missing data in superiority trials assessing structural outcomes in rheumatic diseases. *Arthritis Rheum* 2005;52:1858–65.

26 Joseph R, Sim J, Ogollah R, *et al*. A systematic review finds variable use of the intention-to-treat principle in musculoskeletal randomized controlled trials with missing data. *J Clin Epidemiol* 2015;68:15–24.

27 Ibrahim F, Tom BDM, Scott DL, *et al*. A systematic review of randomised controlled trials in rheumatoid arthritis: the reporting and handling of missing data in composite outcomes. *Trials* 2016;17:272.

28 Evans SR, Bigelow R, Chuang-Stein C, *et al*. Presenting risks and benefits: helping the data monitoring Committee do its job. *Ann Intern Med* 2020;172:119–25.

29 Khan NA, Lombeida JI, Singh M, *et al*. 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.

30 Egger M, Jüni P, Bartlett C, *et al*. Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001;285:1996–9.

31 Goekoop-Ruiterman YPM, de Vries-Bouwstra JK, Allaart CF, *et al*. Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2007;146:406–15.

32 Fautrel B, Pham T, Alfaiate T, *et al*. Step-Down strategy of spacing TNF-blocker injections for established rheumatoid arthritis in remission: results of the multicentre non-inferiority randomised open-label controlled trial (STRASS: spacing of TNF-blocker injections in rheumatoid arthritis study). *Ann Rheum Dis* 2016;75:59–67.

33 Wilkinson TJ, Lemmey AB, Jones JG, *et al*. Can creatine supplementation improve body composition and objective physical function in rheumatoid arthritis patients? A randomized controlled trial. *Arthritis Care Res* 2016;68:729–37.

34 Stallard N. Optimal sample sizes for phase II clinical trials and pilot studies. *Stat Med* 2012;31:1031–42.

35 Kremer JM, Genant HK, Moreland LW, *et al*. Effects of abatacept in patients with methotrexate-resistant active rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2006;144:865–76.

36 Rajaei E, Mowla K, Ghorbani A, *et al*. The effect of omega-3 fatty acids in patients with active rheumatoid arthritis receiving DMARDs therapy: double-blind randomized controlled trial. *Glob J Health Sci* 2015;8:18–25.

37 Schulz KF, Grimes DA. Sample size slippages in randomised trials: exclusions and the lost and wayward. *Lancet* 2002;359:781–5.

38 Straus SE, Glasziou P, Richardson SW. *Prognosis. Evidence-based medicine : how to practice and teach ebm*. Edinburgh: Elsevier Health Sciences, 2019: 190.

39 Dossing A, Tarp S, Furst DE, *et al*. Modified intention-to-treat analysis did not bias trial results. *J Clin Epidemiol* 2016;72:66–74.

40 Fergusson D, Aaron SD, Guyatt G, *et al*. Post-Randomisation exclusions: the intention to treat principle and excluding patients from analysis. *BMJ* 2002;325:652–4.

41 Bell ML, Kenward MG, Fairclough DL, *et al*. Differential dropout and bias in randomised controlled trials: when it matters and when it may not. *BMJ* 2013;346:e8668.

42 Hewitt CE, Kumaravel B, Dumville JC, *et al*. Assessing the impact of attrition in randomized controlled trials. *J Clin Epidemiol* 2010;63:1264–70.

43 Dziura JD, Post LA, Zhao Q, *et al*. Strategies for dealing with missing data in clinical trials: from design to analysis. *Yale J Biol Med* 2013;86:343–58.

44 National Research Council (US) Panel on Handling Missing Data in Clinical Trials. The prevention and treatment of missing data in clinical trials, 2010. Available: https://www.ncbi.nlm.nih.gov/books/NBK209904/ 10.17226/12955 [Accessed Mar 2021].

45 Lydersen S. Last observation carried forward. *Tidsskr Nor Laegeforen* 2019;139. doi:10.4045/tidsskr.19.0061. [Epub ahead of print: 28 May 2019].

46 Coffman CJ, Edelman D, Woolson RF. To condition or not condition? Analysing 'change' in longitudinal randomised controlled trials. *BMJ Open* 2016;6:e013096.

47 J T, L B, T H, *et al*. Different ways to estimate treatment effects in randomised controlled trials. *Contemp Clin Trials Commun* 2018;10:80–5.

48 Twisk JW, Rijnhart JJ, Hoekstra T, *et al*. Intention-To-Treat analysis when only a baseline value is available. *Contemp Clin Trials Commun* 2020;20:100684.

49 Soares HP, Daniels S, Kumar A, *et al*. Bad reporting does not mean bad methods for randomised trials: observational study of randomised controlled trials performed by the radiation therapy Oncology Group. *BMJ* 2004;328:22–4.

50 Altman DG, Schulz KF, Moher D, *et al*. The revised consort statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134:663–94.

51 Moher D, Hopewell S, Schulz KF, *et al*. Consort 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. *Int J Surg* 2012;10:28–55.