



HAL
open science

Reporting quality of heart failure randomized controlled trials 2000–2020: Temporal trends in adherence to CONSORT criteria

Mohamed B Jalloh, Veronica A Bot, Cristiana Z Borjaille, Lehana Thabane, Guowei Li, Javed Butler, Faiez Zannad, Christopher B Granger, Harriette G.C. van Spall

► To cite this version:

Mohamed B Jalloh, Veronica A Bot, Cristiana Z Borjaille, Lehana Thabane, Guowei Li, et al.. Reporting quality of heart failure randomized controlled trials 2000–2020: Temporal trends in adherence to CONSORT criteria. *European Journal of Heart Failure*, 2024, 10.1002/ejhf.3229 . hal-04570306

HAL Id: hal-04570306

<https://hal.univ-lorraine.fr/hal-04570306>

Submitted on 7 May 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reporting quality of heart failure randomized controlled trials 2000–2020: Temporal trends in adherence to CONSORT criteria

Mohamed B. Jalloh¹, Veronica A. Bot², Cristiana Z. Borjaille¹, Lehana Thabane^{2,3}, Guowei Li⁴, Javed Butler^{5,6}, Faiez Zannad⁷, Christopher B. Granger⁸, and Harriette G.C. Van Spall^{1,2,3,9*}

¹Department of Medicine, McMaster University, Hamilton, ON, Canada; ²Department of Health Research Methods, Evidence, and Impact, McMaster University, Hamilton, ON, Canada; ³Research Institute of St. Joseph's, Hamilton, ON, Canada; ⁴Center for Clinical Epidemiology and Methodology (CCEM), Guangdong Second Provincial General Hospital, Guangzhou, China; ⁵Baylor Scott and White Research Institute, Dallas, TX, USA; ⁶University of Mississippi, Jackson, MS, USA; ⁷Université de Lorraine, INSERM and Centre Hospitalier Régional Universitaire, Nancy, France; ⁸Division of Cardiology, Duke University School of Medicine, Durham, NC, USA; and ⁹Baim Institute for Clinical Research, Boston, Massachusetts, USA

Received 12 November 2023; revised 25 February 2024; accepted 26 March 2024

Aim

Heart failure (HF) is a major cause of morbidity and mortality in older adults. Randomized controlled trials (RCTs) inform HF policy and practice, but the accurate interpretation of results is contingent on clear and transparent reporting. The CONSolidated Standards Of Reporting Trials (CONSORT) statement serves as a guide to RCT reporting. We evaluated the quality of reporting in HF RCTs in high-impact journals by assessing their adherence to CONSORT.

Methods and results

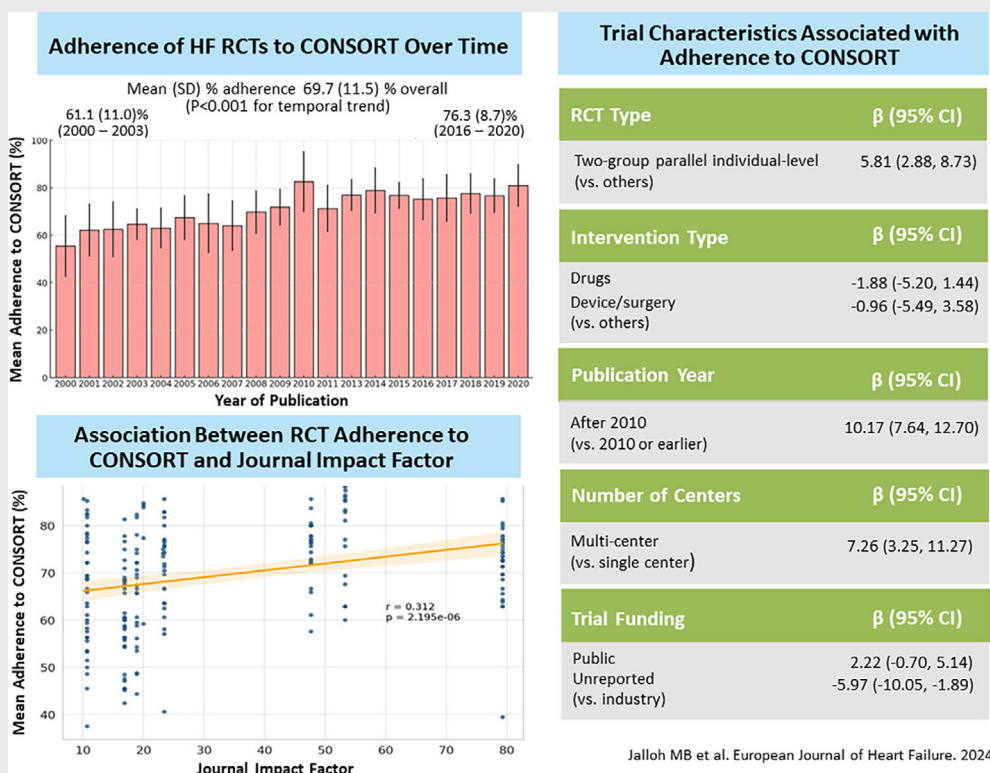
We searched MEDLINE, EMBASE and CINAHL for HF RCTs published in high-impact journals 2000–2020. We assessed the proportion of CONSORT criteria that individual HF RCTs adhered to, and used the Jonckheere–Terpstra test to examine temporal trends in adherence. Multivariable linear regression explored the association between trial characteristics and adherence to CONSORT. Primary analysis assessed adherence to CONSORT 2010 update. A sensitivity analysis assessed adherence to the original (1996) CONSORT criteria. Among 221 RCTs analysed, the mean (standard deviation [SD]) adherence was suboptimal overall (mean [SD] adherence 69.7 [11.5]%) (5513/7913 criteria), with a temporal increase in adherence over the 20-year period ($p < 0.001$). Factors associated with adherence included publication after versus during/before 2010 ($\beta = 10.17$, 95% confidence interval [CI] 7.64–12.70; $p < 0.001$); two-group parallel individual-level randomization versus other (including multi-group or cluster randomization) ($\beta = 5.81$, 95% CI 2.88–8.73; $p < 0.001$); and multicentre versus single-centre trials ($\beta = 7.26$, 95% CI 3.25–11.27; $p < 0.001$). There was no difference in trial adherence to the updated CONSORT (2010) versus the original (1996) CONSORT criteria, and temporal trends in adherence to both sets of criteria were similar, likely due to overlap between the two sets of criteria. Trials with greater adherence to CONSORT were published in higher impact factor journals, with a positive correlation ($r = 0.312$; $p < 0.001$).

Conclusion

The quality of reporting in HF RCTs, as measured by CONSORT adherence, has improved over time but remains suboptimal.

*Corresponding author: Baim Institute for Clinical Research, 20 Copeland Avenue, David Braley Research Building, Suite C3-117, Boston Massachusetts L8L 0A3, USA. Tel: +1 905 5242712, Email: harriette.vanspall@phri.ca

Graphical Abstract



Among 7913 criteria in 221 heart failure (HF) trials published 2000–2020, the mean (standard deviation) adherence to CONSolidated Standards Of Reporting Trials (CONSORT) criteria was 69.7 (11.5)%. CONSORT reporting improved over time ($p < 0.001$). Trial type, year of publication, number of participating centres, and trial funding disclosure were associated with adherence to CONSORT criteria. CI, confidence interval; RCT, randomized controlled trial.

Keywords

Heart failure • CONSORT guidelines

Introduction

Heart failure (HF) is a leading cause of morbidity and mortality among older adults worldwide.¹ Clinical care practices for these individuals depend on evidence generated from randomized controlled trials (RCTs) documented in the literature. While there are a variety of important measures of quality of a clinical trial, the interpretation of trial findings and appropriate translation to care is contingent on transparent reporting and the evaluation of internal and external validity of RCTs. The clarity and comprehensiveness of trial reporting also has implications for meta-analyses which deduce prevalent findings across a breadth of literature. Clear and consistent reporting of RCTs is vitally important for the integrity of HF investigation(s) and care practices thereof.

The CONSolidated Standards Of Reporting Trials (CONSORT) statement, first introduced in 1996 and updated in 2001 and 2010, is an evidence-based guide for reporting RCTs endorsed

by high-impact journals and widely accepted by the medical community. The statement includes checklist items that offer a standard way of reporting clinical trials and serves as a guide for research protocols and reporting of trial results.^{2–4} This guidance aims to reduce the influence of bias and ensure that trial protocols and reports are clear, complete, and transparent. Initially conceived for standard two-arm parallel RCTs, the CONSORT statement has since been modified and extended to accommodate various trial designs (e.g. non-inferiority trials), interventions (e.g. non-pharmacologic interventions), and data types (e.g. CONSORT-PRO).⁵ Adherence to CONSORT reporting standards allows readers to evaluate whether the trial results are valid, and who the results may be applied to.⁶

Despite the recognition of the importance of using CONSORT reporting criteria in clinical trials, they continue to be underutilized.^{6–8} Although more than 600 journals have endorsed the CONSORT statement,⁶ it is unclear whether HF trials,

including those containing practice-changing findings published in high-impact journals, adhere to CONSORT guidelines. Limited data exist on how well HF trials adhere to CONSORT 2010 guidelines. While previous evaluations of RCT reporting quality have focused on a small number of pharmacological intervention trials for HF with preserved ejection fraction (HFpEF) over narrow chronological periods,⁸ a comprehensive assessment of reporting standards across a breadth of HF trials has yet to be conducted. In this systematic bibliometric review, we assessed the reporting of HF RCTs published in journals with an impact factor ≥ 10 between 2000 and 2020. We assessed RCT adherence to CONSORT criteria, temporal trends in trial reporting, and trial factors independently associated with improved adherence.

Methods

Identification of clinical trials

We undertook a systematic literature search in MEDLINE, EMBASE and CINAHL using the following criteria: (1) search terms 'heart failure' and 'randomized controlled trial'; (2) English language publications; (3) publication dates between 1 January 2000 and 17 June 2020; (4) adults with HF 18 years of age or older; (5) journal impact factor ≥ 10 in 2019 based on Web of Science 2019 classification report⁹; (6) full-text manuscripts reporting primary outcomes; and (7) HF trials that enrolled ≥ 100 participants. We used journal impact factors ≥ 10 in 2019 to identify trials published in highly cited journals that are likely to influence standard of care.

To avoid double counting, we excluded publications that described post hoc, interim, exploratory, or secondary analyses. Two authors independently screened all abstracts from the search for eligibility according to pre-defined inclusion criteria and selected eligible studies for inclusion. If the two authors disagreed, a third reviewer resolved the discrepancy by consensus.

Data extraction and reporting assessment

Three reviewers (M.B.J., V.A.B., and C.Z.B.) independently extracted data. Disagreements were resolved by consensus and adjudicated by a fourth assessor (H.G.C.V.) when required.

The revised CONSORT 2010 statement explanation and elaboration was used to guide the assessment of criteria.² Each criterion was assigned a 'Yes', 'No', or 'Not applicable (NA)' response indicating if it was reported, not reported, or not relevant for the checked item. A "Yes" response was awarded 1 point. We counted the total number of points and divided the total number of points available (excluding 'NA') for that article and computed the percentage adherence for each study (excluding 'NA').⁸ A total mean (standard deviation [SD]) CONSORT adherence was tabulated across all full-text HF RCTs.

In addition to the CONSORT criteria, we extracted available data regarding the year of publication, region of coordinating centre, type and level of randomization (cluster vs. individual randomization, multi-arm vs. two-arm, parallel vs. crossover), type of intervention (drug, device, surgical, or other), number of participating centres, funding source, trial size, and journal impact factor.

Outcomes

The primary outcome was adherence to CONSORT criteria, estimated as the proportion of CONSORT criteria that individual trials adhered to. We assessed overall adherence, temporal trends in adherence, and trial characteristics independently associated with improved adherence.

Data analysis and synthesis

We performed descriptive statistical analysis and reported the mean and SD for normally distributed continuous variables, and the median and interquartile range for non-normally distributed variables in 221 full-text HF RCT articles. We reported numbers and proportions for categorical variables. Adherence to CONSORT was estimated based on the proportion (%) of the relevant criteria that were appropriately reported.

We used multivariable linear regression to explore independent associations between pre-specified trial characteristics and mean CONSORT adherence. Explanatory variables were selected based on their relevance to the analysis performed and included: type of trial design (two-arm parallel individual randomization vs. others), type of intervention (drugs, device/surgery vs. others [health service and exercise/rehabilitation]), year of publication (after 2010 vs. during or before 2010), number of participating centres (multicentre vs. single-centre), and funding sources (public, missing vs. industry). The results were reported as coefficients (estimated β) along with 95% confidence intervals (CI) and *p*-values. Tests of significance were performed at 5% significance level. We applied the Jonckheere–Terpstra trend test (two-tailed testing, $\alpha = 0.05$) to analyse temporal trends in reporting of HF RCTs. Data were analysed using R (version 4.2.1) software.

Sensitivity analysis

We assessed mean adherence to the original (1996) CONSORT criteria, and using the Jonckheere–Terpstra trend test, assessed temporal trends in RCT adherence to them.

Our study is registered in the International Prospective Register of Systematic Reviews (PROSPERO; ID CRD42023446327). The results of this systematic review were reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.¹⁰

Results

A total of 18 828 articles were screened in EMBASE, CINAHL, and MEDLINE. Among all unique HF RCT full-text articles assessed for eligibility, 3189 were excluded. A total of 221 articles satisfied eligibility criteria and were included in this study (Figure 1).

Baseline characteristics

Table 1 summarizes the characteristics of the 221 articles included in this systematic review. A majority of trials were coordinated in North America (50.7%) and Europe (41.6%), and a small proportion in Asia (3.6%), Australia (2.3%) and Central South America (1.4%), and Africa (0.5%). Among the included HF RCTs, there were five cluster-randomized trials, which included

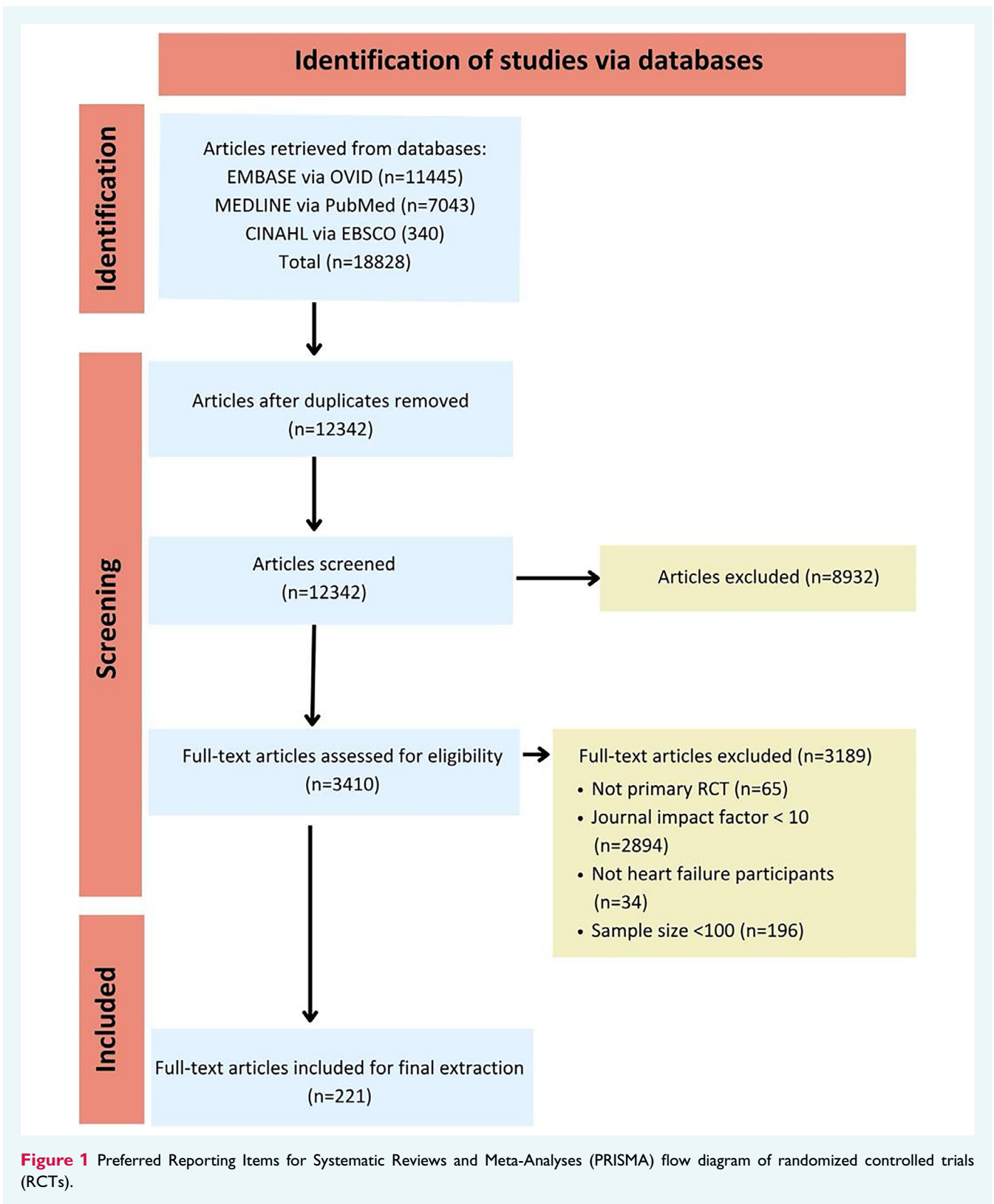


Table 1 Baseline characteristics of 221 heart failure randomized controlled trials included in the study

Trial characteristic	Total trials, n (%)	Adherence to 2010 CONSORT update, mean (SD) %
Trial type		
Cluster randomized trials		
Parallel-group cluster	3 (1.4)	60.4 (2.1)
Stepped-wedge cluster	2 (0.9)	66.0 (9.7)
Individual randomized trials		
Multi-arm parallel-group	36 (16.3)	64.3 (11.8)
Non-inferiority or equivalence	7 (3.2)	70.1 (8.8)
Two-group parallel	173 (78.3)	71.1 (11.2)
Coordinating centre region		
Europe	92 (41.6)	69.3 (12.8)
North America	112 (50.7)	69.6 (10.4)
Asia	8 (3.6)	66.2 (9.8)
Australia	5 (2.3)	67.9 (10.6)
Africa	1 (0.5)	74.3 (NA)
Central and South America	3 (1.4)	83.7 (2.0)
Trial size (number of participants)		
100–500	142 (64.3)	67.0 (12.3)
501–1000	24 (10.9)	76.0 (8.8)
>1000	55 (24.9)	73.4 (7.6)
Intervention type		
Drug	143 (64.7)	68.5 (12.0)
Health service	46 (20.8)	71.9 (11.1)
Device/surgery	27 (12.2)	70.4 (9.3)
Exercise/rehabilitation	5 (2.3)	70.7 (8.8)
Number of participating centres		
Multicentre	195 (88.2)	70.1 (10.9)
Single-centre	26 (11.8)	64.3 (14.0)
Funding type		
Industry	117 (52.9)	69.8 (10.7)
Public	79 (35.7)	73.1 (10.1)
Unreported	25 (11.3)	58.9 (12.6)
Journal impact factor		
10–35	136 (61.5)	66.1 (11.7)
36–60	47 (21.3)	77.1 (9.0)
>60	38 (17.2)	72.6 (7.8)
Total	221 (100.0)	69.7 (11.5)

CONSORT, CONSolidated Standards Of Reporting Trials; SD, standard deviation.

two stepped-wedge trials and three parallel-arm cluster trials. Additionally, there were 216 individual-level RCTs, including 36 multi-arm trials, 173 two-arm trials, and 7 non-inferiority trials among those analysed. Most of the trials (78.3%) were two-arm parallel individual-level RCTs. A majority (88.2%) of trials were conducted across multiple centres. Over half of all trials (52.9%) received funding from industry, while 35.7% were funded through public sources. Most trials (64.3%) included a sample population of 100 to 500 participants. A majority of RCTs (58.4%) were published on or before 2010. Journals with an impact factor between

10 and 35 accounted for most (61.5%) of the publications included in this systematic review.

Adherence to CONSORT

The mean (SD) adherence of trials to CONSORT criteria was 69.7 [11.5]% (5513/7913 criteria). Trials across regions had similar adherence to CONSORT, with slightly greater adherence in Central and South American trials (82/98 criteria, 83.7 [2.0]%) relative to the rest, although there were only three trials from this region. Two-arm parallel individual-level RCTs exhibited greater mean adherence to CONSORT (4187/5888 criteria, 71.1 [11.2]%) than other trial types. Trials with 501–1000 participants (651/862 criteria, 76.0 [8.8]%) and >1000 participants had a greater adherence (1447/1971, 73.4 [7.6]%) to CONSORT than trials with <500 participants (3416/5100 criteria 67.0 [12.3]%). Adherence to CONSORT was similar across intervention type (1153/1603 criteria, 71.9 [11.1]%) for healthcare service; 3562/5204, 68.5 [12.0]%) for drug; 561/797, 70.4 [9.3]%) for device/surgery; and 118/167 criteria, 70.7 [8.8]%) for exercise/rehabilitation interventions.

Trials fully adhered to CONSORT criteria pertaining to trial background and objectives (criteria 2a-i, 2b-i, and 2b-ii), participant eligibility (4a-i), outcomes (6a-ii and 6b), randomization (8b-ii), baseline table (15a), generalizability (21a), and research ethics review subsection for stepped-wedge cluster trials (extension item 26) (Table 2 and online supplementary Table S1, Figure S1).

In contrast, poor reporting was observed among title and abstract extension items 1a-ii, 1a-iii, and 1a-v, which identified RCTs as having cluster, non-inferiority or equivalence, and stepped-wedge cluster designs, respectively (online supplementary Table S2). There was suboptimal reporting of criteria related to sample size and statistical power for cluster RCTs (extension item 7a-ii pertaining to cluster number, size, balance, the coefficient of intra-cluster correlation and indication of its uncertainty). The implementation of cluster extension criteria including the description of allocation sequence, enrolment in clusters and assignment of interventions (10b-i), mechanism of inclusion in clusters for the purpose of the trial (10b-ii), and from whom and when consent was obtained (10b-iii) were seldom presented (online supplementary Table S2 and Figure S1).

Trials with lower CONSORT adherence were published in journals with a lower impact factor than those with greater adherence. Trials published in journals with impact factor ≤ 35 had lower mean (SD) adherence to CONSORT criteria and relevant extensions (3217/4869 criteria, 66.1 [11.7]%) than those published in journals with impact factor 35–60 (1312/1702 criteria, 77.1 [9.0]%) or impact factor >60 (985/1356 criteria, 72.6 [7.8]%) (online supplementary Figure S2). There was a moderate linear correlation between mean adherence to CONSORT 2010 criteria and journal impact factor ($r = 0.312$; $p = 2.195 \times 10^{-6}$) (online supplementary Figure S3).

Sensitivity analysis

Mean (SD) trial adherence to the original (1996) CONSORT criteria was 68.8 (15.8)% (1365/1984 criteria). There was no difference in trial adherence to the 2010 update versus original

Table 2 Proportion of eligible trials that adhere to applicable CONSORT 2010 criteria

Item	Assessment criteria	Adherent RCTs, n/N (%)
Title and abstract		
1a-i	Identification as RCT in the title	107/221 (48.4)
1a-ii	Identification as a cluster randomized trial in the title	0/5 (0.0)
1a-iii	Identification as a non-inferiority randomized trial in the title	0/7 (0.0)
1a-iv	Identification as multi-arm randomized trial in the title and indication of number of treatment groups that the participants were randomly assigned to	1/36 (2.7)
1a-v	Identification as a stepped-wedge cluster randomized trial in the title.	0/2 (0.0)
1b-i	Structured summary of trial design, methods, results, and conclusions outlined in parallel trials; 'abstracts'	13/221 (5.9)
1b-ii	Specify the number of treatment groups; details of any groups added or dropped	30/36 (83.3)
Introduction: background and objectives		
2a-i	Scientific background and explanation of rationale ^a	221/221 (100.0)
2a-ii	Rationale for using a cluster design	3/5 (60.0)
2a-iii	Rationale for using a non-inferiority design	1/7 (14.3)
2a-iv	Rationale for using a multi-arm design	21/36 (58.3)
2a-v	Rationale for using a stepped-wedge design	1/2 (50.0)
2b-i	Specific objectives or hypotheses	221/221 (100.0)
2b-ii	Whether objectives pertain to the cluster level, the individual participant level or both	5/5 (100.0)
2b-iii	Hypothesis concerning non-inferiority, specifying the non-inferiority margin with the rationale for its choice	4/7 (57.1)
2b-iv	Specification of the research question referring to all of the treatment groups; clear statement of hypothesis to be tested and the primary comparisons involved	16/36 (44.4)
2b-v	Specific objectives or hypotheses for stepped-wedge clusters	1/2 (50.0)
Methods: trial design		
3a-i	Description of trial design (such as parallel, factorial) including allocation ratio ^a	68/221 (30.8)
3a-ii	Definition of cluster and description of how the design features apply to the clusters	1/5 (20.0)
3a-iii	Specification of number of treatment groups	33/36 (91.7)
3a-iv	Description and diagram of trial design including definition of cluster, number of sequences, number of clusters randomized to each sequence, number of periods, duration of time between each step, and whether the participants assessed in different periods are the same people, different people, or a mixture	1/2 (50.0)
3b-i	Important changes to methods after trial commencement (such as eligibility criteria), with reasons ^a	31/36 (86.1)
3b-ii	Details of any treatment groups added or dropped (if relevant) with reasons, and or changes to the allocation ratio	4/6 (66.7)
3b-iii	Important changes to methods after stepped-wedge cluster trial commencement (such as eligibility criteria), with reasons	NA
Participants		
4a-i	Eligibility criteria for participants	219/221 (99.1)
4a-ii	Eligibility criteria for clusters	3/5 (60.0)
4a-iii	Whether participants in the non-inferiority trial are similar to those in any trials that established the efficacy of the reference treatment	1/7 (14.3)
4a-iv	Eligibility criteria for clusters and participants	2/2 (100.0)
4b	Settings and locations where the data were collected ^a	189/221 (85.5)
Interventions		
5a-i	The interventions for each group with sufficient details to allow replication, including how and when they were actually administered ^a	219/221 (99.1)
5a-ii	Whether interventions pertain to the cluster level, the individual participant level or both	3/5 (60.0)
5a-iii	Whether the reference treatment in the non-inferiority trial is identical (or very similar) to that in any trials that established efficacy	2/7 (28.6)
Outcomes		
6a-i	Completely defined pre-specified primary and secondary outcome measures, including how and when they were assessed	219/221 (99.1)
6a-ii	Whether outcome measures pertain to the cluster level, the individual participant level, or both	5/5 (100.0)
6a-iii	Specify the non-inferiority outcomes and whether hypotheses for main and secondary outcomes are non-inferiority or superiority. Whether the outcomes in the non-inferiority trial are identical or similar to those in any trials that established the efficacy of the reference treatment	2/7 (28.6)
6b	Any changes to trial outcomes after the trial commenced, with reasons ^a	10/10 (100.0)

Table 2 (Continued)

Item	Assessment criteria	Adherent RCTs, n/N (%)
Sample size		
7a-i	How sample size was determined	193/221 (87.3)
7a-ii	Method of calculation, number of clusters(s) (and whether equal or unequal cluster sizes are assumed), cluster size, a coefficient of intra-cluster correlation (or <i>k</i>), and an indication of its uncertainty	0/5 (0.0)
7a-iii	Whether sample size was calculated using a non-inferiority criterion and if so, what the inferiority margin was	4/7 (57.1)
7a-iv	Planned sample size with details of how it was determined for each primary comparison	20/36 (55.6)
7b-i	When applicable, explanation of any interim analyses and stopping guidelines	75/164 (45.7)
7b-ii	To which outcomes they apply and whether related to a non-inferiority hypothesis	3/7 (42.9)
Randomization		
Sequence generation		
8a	Method used to generate the random allocation sequence	111/221 (50.2)
8b-i	Type of randomization; details of any restriction (such as blocking and block size)	99/ 221 (44.8)
8b-ii	Details of stratification or matching if used	5/5 (100.0)
Allocation and concealment mechanism		
9a-i	Mechanism used to implement the random allocation sequence (such as sequentially numbered containers), describing any steps taken to conceal the sequence until interventions were assigned	96/221 (43.4)
9a-ii	Specification that allocation was based on clusters rather than individuals and whether allocation concealment (if any) was at the cluster level, the individual participant level, or both	2/5 (40.0)
Implementation		
10a	Who generated the random allocation sequence, who enrolled participants, and who assigned participants to interventions ^a	20/221 (9.0)
10b-i	Who generated the random allocation sequence, who enrolled clusters, who assigned clusters to interventions	0/5 (0.0)
10b-ii	Mechanism by which individual participants were included in clusters for the purpose of the trial (enumeration, random sampling)	0/5 (0.0)
10b-iii	From whom consent was sought (representatives from the cluster, individual cluster members), consent before or after randomization	0/5 (0.0)
Blinding		
11a	If done, who was blinded after assignment to interventions (e.g. participants, care providers, those assessing outcomes) and how ^a	151/201 (75.1)
11b	If relevant, description of the similarity of interventions	35/144 (24.3)
Statistical methods		
12a-i	Statistical methods used to compare groups for primary and secondary outcomes	218/221 (98.6)
12a-ii	How clustering was taken into account	1/5 (20.0)
12a-iii	Whether a one- or two-sided confidence interval approach was used	2/7 (28.6)
12a-iv	Explicitly state if no adjustments for multiplicity were applied; if adjustments were applied, state the methods used	9/36 (25.0)
12b	Methods for additional analyses, such as subgroup analyses and adjusted analyses	152/176 (86.4)
Participant flow		
13a-i	For each group, the numbers of participants who were randomly assigned, received intended treatment, and were analysed for the primary outcome	217/221 (98.2)
13a-ii	For each group, the numbers of clusters that were randomly assigned, received intended treatment, and were analysed for the primary outcome	4/5 (80.0)
13b-i	For each group, losses and exclusions after randomization, together with reasons ^a	210/221 (95.0)
13b-ii	For each group, losses and exclusions for both clusters and individual cluster members	4/5 (80.0)
Recruitment		
14a-i	Dates defining the periods of recruitment and follow-up ^a	194/ 221 (87.8)
14a-ii	If periods of recruitment and follow-up are different across treatment groups (e.g. groups were added or dropped), the periods of recruitment and follow-up, reasons should the differences, and any statistical implications should be described	5/36 (13.9)
14a-iii	Dates defining the steps, initiation of intervention, and deviations from planned dates. Dates defining recruitment and follow-up for participants	2/2 (100.0)
14b	Why the trial ended or was stopped ^a	34/221 (15.4)

Table 2 (Continued)

Item	Assessment criteria	Adherent RCTs, n/N (%)
Baseline data		
15a	A table showing baseline demographic and clinical characteristics for each group	221/221 (100.0)
15b	Baseline characteristics for the individual and cluster levels as applicable for each group	1/5 (20.0)
Numbers analysed		
16a	For each group, number of participants (denominator) included in each analysis and whether the analysis was by original assigned groups ^a	217/221 (98.2)
16b	For each group, number of clusters included in each analysis	3/5 (60.0)
Outcomes and estimation		
17a-i	For each primary and secondary outcome, results for each group, and the estimated effect size and its precision (such as 95% confidence interval)	164/221 (74.2)
17a-ii	Results at the individual or cluster level as applicable and a coefficient of intra-cluster correlation (or <i>k</i>) for each primary outcome	1/5 (20.0)
17a-iii	For the outcomes for which non-inferiority was hypothesized, a figure showing confidence intervals and the non-inferiority margin may be useful	4/7 (57.1)
17a-iv	Results for each pre-specified comparison of treatment groups	29/36 (80.6)
17b	For binary outcomes, presentation of both absolute and relative effect sizes is recommended ^a	87/214 (40.7)
Ancillary analyses		
18	Results of any other analyses performed, including subgroup analyses and adjusted analyses, distinguishing pre-specified from exploratory ^a	156/181 (86.2)
Harms		
19	All important harms or unintended effects in each group ^a	155/176 (88.1)
Discussion Limitations		
20	Trial limitations, addressing sources of potential bias, imprecision, and, if relevant, multiplicity of analyses ^a	129/221 (58.4)
Generalizability		
21a	Generalizability (external validity, applicability) of the trial findings	221/221 (100.0)
21b	Generalizability to clusters and/or individual participants (as relevant)	5/5 (100.0)
Interpretation		
22a	Interpretation consistent with results, balancing benefits and harms, and considering other relevant evidence	217/221 (98.2)
22b	Interpret results in relation to the non-inferiority hypothesis. If a superiority conclusion is drawn for outcomes for which non-inferiority was hypothesized, provide justification for switching	4/7 (57.1)
Other information		
Registration		
23	Registration number and name of trial registry ^a	120/221 (54.3)
Protocol		
24	Where the full trial protocol can be accessed, if available ^a	43/221 (19.5)
Funding		
25	Sources of funding and other support (such as supply of drugs), role of funders ^a	196/221 (88.7)
Research ethics review		
26	Whether the study was approved by a research ethics committee, with identification of the review committee(s). Justification for any waiver or modification of informed consent requirements	2/2 (100.0)

CONSORT, CONSolidated Standards Of Reporting Trials; NA, not available; RCT, randomized controlled trial.

White rows: all trials, two-group parallel individual-level randomization trials. Red rows: extension checklist items for non-inferiority and equivalence trials. Blue rows: extension checklist items for multi-arm parallel-group trials. Green rows: extension checklist items for stepped-wedge cluster trials.

^aChecklist items different from or in addition to the original (1996) CONSORT criteria.

(1996) CONSORT criteria (mean difference 0.91%, $p=0.191$), likely due to similarity in the two sets of criteria.

Temporal trends in adherence to CONSORT

Adherence to 2010 CONSORT criteria improved over time ($p < 0.001$ for all comparisons). Trials published after 2010 exhibited a mean (SD) adherence of 76.5 (8.3)%

(2537/3317 criteria). Adherence improved between 2000 and 2003 (1207/1974 criteria, 61.1 [11.0]%) and 2016–2020 (1667/2187 criteria, 76.3 [8.7]%) (Figure 2, online supplementary Table S3).

Sensitivity analysis

There was a temporal increase in trial adherence to the original (1996) CONSORT criteria ($p < 0.001$) (Figure 2).

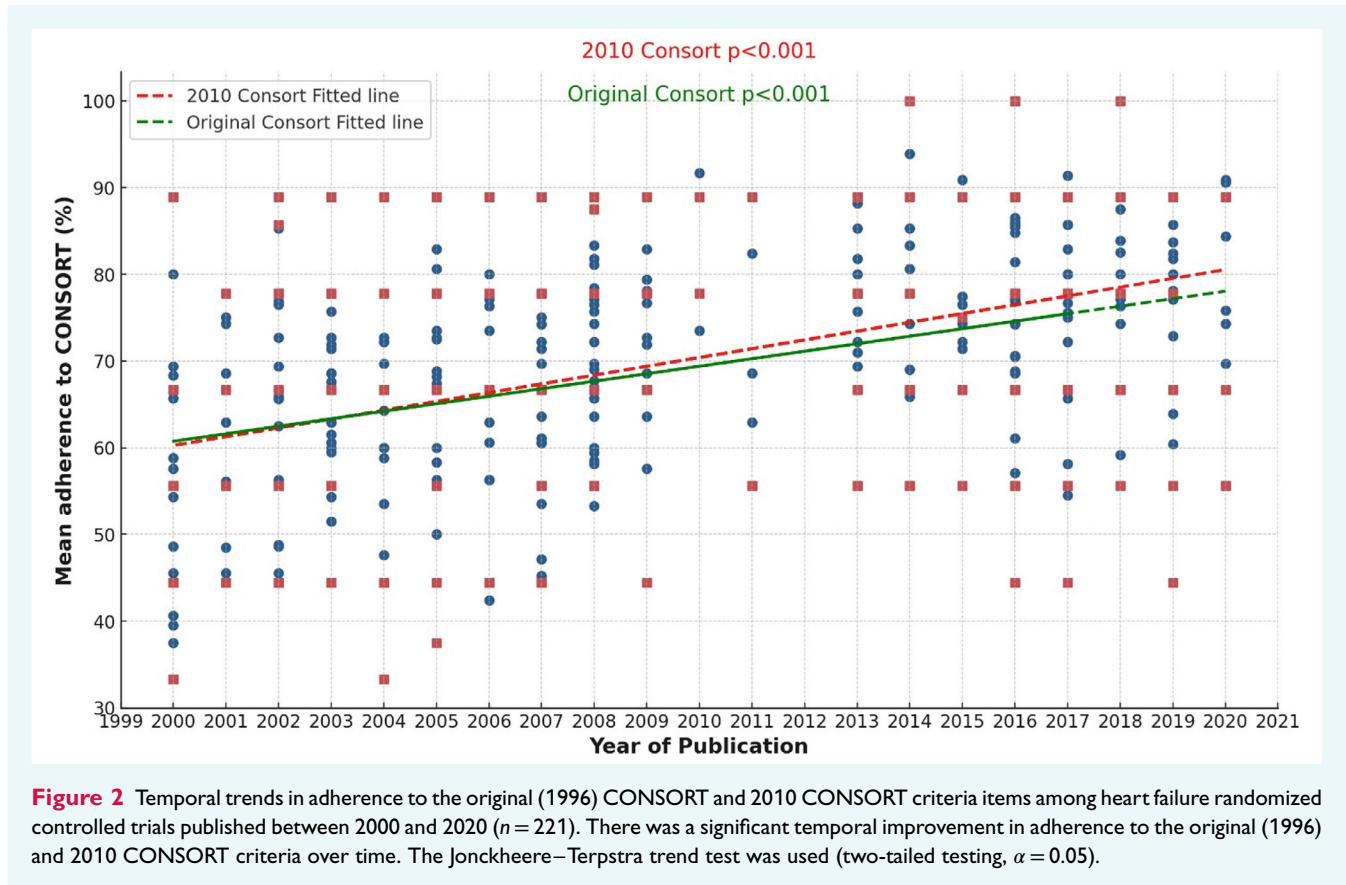


Figure 2 Temporal trends in adherence to the original (1996) CONSORT and 2010 CONSORT criteria items among heart failure randomized controlled trials published between 2000 and 2020 ($n = 221$). There was a significant temporal improvement in adherence to the original (1996) and 2010 CONSORT criteria over time. The Jonckheere–Terpstra trend test was used (two-tailed testing, $\alpha = 0.05$).

Trial characteristics independently associated with adherence

Year of publication, level of randomization, number of participating centres, and transparent reporting of funding sources were significantly associated with mean adherence to CONSORT criteria. Studies published after 2010 were associated with greater adherence to CONSORT reporting than those published during or before 2010 ($\beta = 10.17$, 95% CI 7.64–12.70; $p < 0.001$) (Figure 3). Trials with two-arm parallel individual-level randomization were associated with greater adherence than other trials ($\beta = 5.81$, 95% CI 2.88–8.73); $p < 0.001$). Multicentre trials showed a greater adherence to CONSORT standards compared to single-centre trials (estimated $\beta = 7.26$, 95% CI 3.25–11.27; $p < 0.001$). Lack of reporting of funding source was associated with a lower adherence to CONSORT overall ($\beta = -5.97$, 95% CI: -10.05 to -1.89; $p = 0.004$). There was no significant difference in adherence based on intervention type (Figure 3).

Discussion

In this systematic bibliometric review, we assessed the quality of reporting in 221 HF RCTs published between 2000 and 2020 in high-impact journals, defining quality as adherence to CONSORT reporting standards. We found less than 70% adherence to the CONSORT 2010 update and to the original (1996) CONSORT

criteria, although adherence to both sets increased with time. Suboptimal reporting was noted particularly in cluster RCTs. Trial characteristics independently associated with increased adherence included more recent publication, two-arm parallel individual-level randomization, multicentre trials and transparent reporting of funding source. There was a positive correlation between trial adherence to CONSORT and publication in a higher impact factor journal (*Graphical Abstract*).

Suboptimal adherence to specific CONSORT categories highlights opportunities for improved reporting. Criteria pertaining to trial title and abstract, methodology, and results had suboptimal adherence in all RCTs. Less than 25% of criteria pertaining to randomization provided a rationale for using a non-inferiority trial, and provision of the protocol. Some criteria are less important and may not even be possible, for example including ‘non-inferiority’ in the title if the journal demands a very short title that simply defines the topic. Cluster RCTs, in particular, had suboptimal reporting; cluster trials infrequently specified cluster randomization in the title, and inadequately described the target sample size, the method used to estimate sample size, the number and size of clusters, and the coefficient of intra-cluster correlation and associated uncertainty. Cluster RCTs also sub-optimally reported how clusters were selected for the trial, from whom consent was obtained, and how clusters were randomly allocated to intervention or comparator arm. Non-inferiority and equivalence trials had similar limitations, withholding the rationale for choosing a non-inferiority

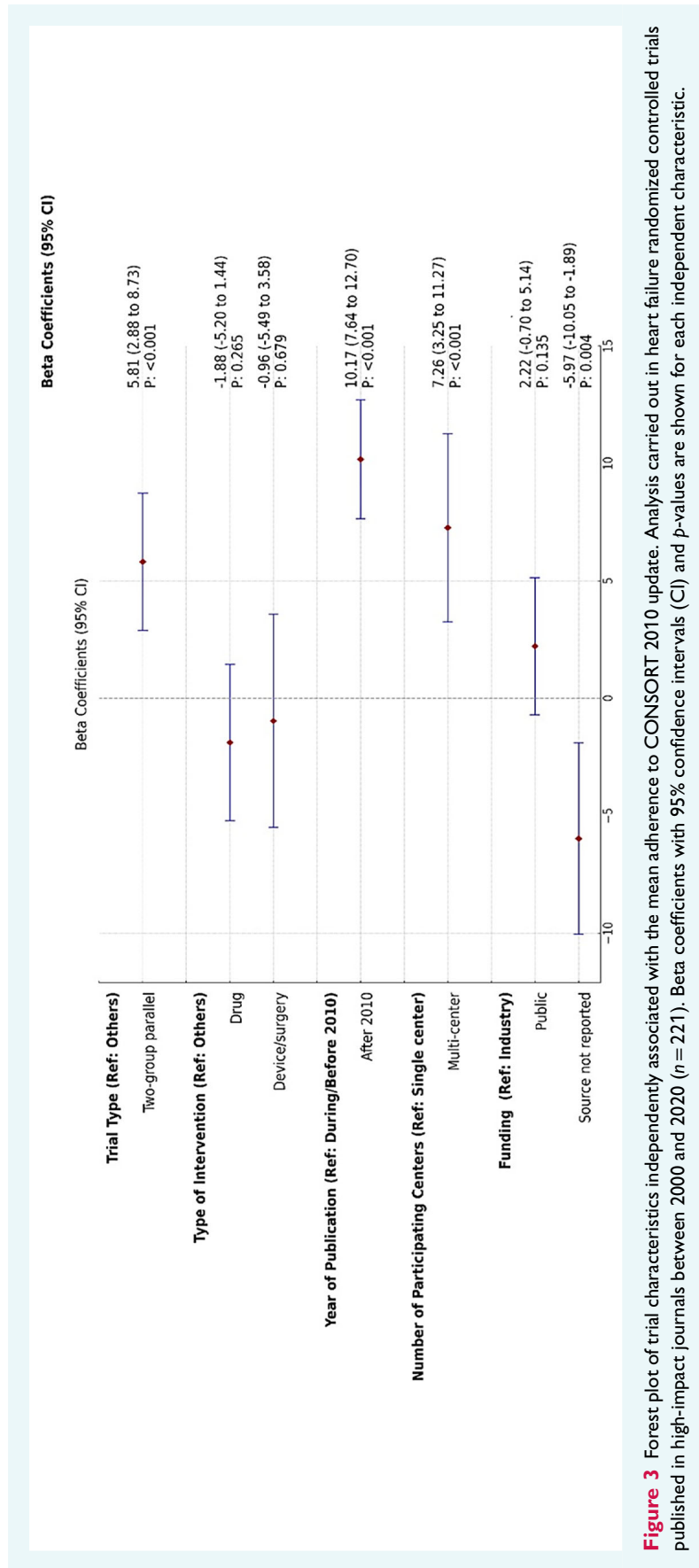


Figure 3 Forest plot of trial characteristics independently associated with the mean adherence to CONSORT 2010 update. Analysis carried out in heart failure randomized controlled trials published in high-impact journals between 2000 and 2020 (n = 221). Beta coefficients with 95% confidence intervals (CI) and p-values are shown for each independent characteristic.

design, under-reporting sample size and power calculations, and providing insufficient detail on randomization procedures. That these limitations were not identified during peer review may reflect reviewers' unfamiliarity with less common trial designs.

We found no significant difference in CONSORT adherence according to type of intervention, but opaque reporting regarding funding source was associated with lower adherence. As many as 11.3% of RCTs did not disclose their funding source. Our findings are consistent with prior studies demonstrating under-reporting of funding sources.^{2,11,12} A survey of PubMed-indexed RCTs indicated that 34% of 519 trials did not disclose the funding source.¹¹ This issue necessitates rigorous policy changes to guarantee the transparency of financial disclosures in publications.

Prior studies have assessed RCT adherence to CONSORT standards, although none assessed for factors independently associated with reporting quality. A prior evaluation of 33 HFpEF RCTs reported a mean CONSORT adherence of 55.4%, lower than our analysis of 221 HF RCTs.⁸ Our review had a larger, more contemporary group of RCTs; included HF across a spectrum of ejection fraction (vs. HFpEF alone), and included a range of drug, device/surgery, service, and exercise/rehabilitation interventions. A review of HF RCTs published 2010 to 2020 in nine high-impact journals documented inconsistent outcome reporting relative to protocols and registration websites in about 10% of trials.¹³ A prior review of RCT abstracts in four high-impact medical journals from 2010 identified inconsistencies and frequent non-adherence to CONSORT guidelines for abstracts, particularly in methodological quality domains.¹⁴ Similarly, another scoping review reported inconsistencies between abstracts and full reports including study design, sample size, designating a primary outcome, demonstrating main results, and drawing a conclusion.¹⁵ While not specific to HF trials, these findings align with our observation that adherence to CONSORT and extension items remain poor, particularly for the description of trial design and allocation concealment.¹⁴

Suboptimal reporting of primary RCT results can introduce bias and may lead to misinterpretation of treatment effect. Pooling such trials for systematic reviews and meta-analyses can amplify biases and impact the quality of care offered to patients with HF.¹⁶ Moreover, suboptimal reporting of title and abstract criteria have implications on searches and indexing in databases.² If authors do not clearly report the information, indexers may not categorize publications accurately.

The temporal increase in adherence to CONSORT confirms gradual knowledge uptake of recommendations. The correlation between adherence and publication in higher impact journals may reflect more rigorous standards or quality of peer review in higher impact factor journals. To further enhance adherence to CONSORT and avoid the problem of missing information at the trial reporting stage, investigators could incorporate elements of the CONSORT criteria in the protocol development. For example, reporting standards regarding cluster size, intra-cluster correlation, and other important aspects of internal validity should be considered when a trial is being designed.¹⁷ Endorsement of CONSORT by leading editorial organizations such as the World Association

of Medical Editors (WAME), International Committee of Medical Journal Editorial (ICMJE), the Council of Science Editors (CSE),¹⁸ European Clinical Research Infrastructure Network (ECIRN)¹⁹ and requirements for use of the checklist by top journals and reviewers may support adherence. Reviewers could use the criteria as a guide during peer review.²⁰

Strengths and limitations of this review

This study includes a systematic review of HF trials published over a 20-year period. Our study conformed to the PRISMA reporting guideline for systematic reviews. Data extraction and assessment of adherence was performed independently and in duplicate by assessors. With a focus on potentially practice-changing trials, we limited the analysis to RCTs published in high-impact factor journals and recognize that adherence may be poorer in trials published in low-impact factor journals. We assessed to both the 2010 update and the original (1996) criteria.

Limitations should be noted. Adherence to CONSORT criteria is a measure of the quality of trial reporting, but does not capture important aspects of trial validity.²¹ We weighted each CONSORT reporting category equally, although we recognize that some have more implications on internal validity than others. Trial registration, publication of the protocol versions, publication of the complete statistical analysis plan, and data sharing are also important means to improve disclosure and transparency. Finally, we could not account for any revision requests from reviewers that may have influenced trial reporting or adherence to CONSORT.

Conclusion

In this analysis of HF 221 RCTs published in high-impact factor journals between 2000 and 2020, we found that the quality of trial reporting was suboptimal, but improved over time. The quality of reporting was particularly suboptimal in uncommon trial designs such as multi-group, cluster randomized, and non-inferiority or equivalence trials. Use of the CONSORT checklist, particularly in uncommon designs, could increase the transparency of trial reporting and better inform the interpretation of results, practice, and policy.

Supplementary Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Conflict of interest: J.B. is a consultant at Abbott, American Regent, Amgen, Applied Therapeutic, AstraZeneca, Bayer, Boehringer Ingelheim, Bristol Myers Squibb, Cardiac Dimension, Cardior, CVRx, Cytokinetics, Edwards, Element Science, Innolife, Impulse Dynamics, Imbria, Inventiva, Lexicon, Lilly, LivaNova, Janssen, Medtronic, Merck, Occlutech, Novartis, Novo Nordisk, Pfizer, Pharmacosmos, Pharmain, Roche, Sequana, SQ Innovation, and Vifor. F.Z. received personal fees from 89Bio, Applied Therapeutics, Bayer, Boehringer, BMS, CVRx, Cardior, Cereno pharmaceutical, Cellprothera, CEVA, KPB, Merck, Novartis, NovoNordisk, Owkin, Pfizer, Otsuka, Roche Diagnostics, Servier, US2.2, having stock options

at G3Pharmaceutical and equities at Cereno pharmaceutical, Cardiorenal, Eshmon Clinical research and being the founder of Cardiovascular Clinical Trialists. C.B.G. received consulting fees from Abbvie, Abiomed, Alnylam Pharmaceuticals, Anthos, Bayer Corporation, Boehringer Ingelheim, Boston Scientific Corporation, Bristol Myers Squibb, Cardionomics, CeleCor Therapeutics, Janssen Pharmaceutical, Merck, Novo Nordisk, Novartis Pharmaceutical Company, Pfizer, Philips, REATA, NephroSynergy; he also has salary funded by Duke grants sponsored by Boehringer Ingelheim, Bristol Myers Squibb, Daiichi Sankyo, Food and Drug Administration, Janssen Pharmaceuticals, Novartis Pharmaceutical Company, Pfizer, and Philips; he has Equity in Tenac.io. H.G.C.V. received research support from the Canadian Institutes of Health Research and Heart and Stroke Foundation of Canada, educational programme grants from Novartis and Boehringer-Ingelheim, and consulting fees from Regeneron and Medtronic. All other authors have nothing to disclose.

References

- Wei S, Miranda JJ, Mamas MA, Zühlke LJ, Kontopantelis E, Thabane L, et al. Sex differences in the etiology and burden of heart failure across country income level: Analysis of 204 countries and territories 1990–2019. *Eur Heart J Qual Care Clin Outcomes*. 2023;**9**:662–672. <https://doi.org/10.1093/ehjqcco/qcac088>
- Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;**340**:c869. <https://doi.org/10.1136/bmj.c869>
- Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996;**276**:637–639. <https://doi.org/10.1001/jama.276.8.637>
- Moher D, Schulz KF, Altman DG, Lepage L. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel group randomized trials. *BMC Med Res Methodol* 2001;**134**:657–662. <https://doi.org/10.7326/0003-4819-134-8-200104170-00011>
- EQUATOR Network. Enhancing the Quality and Transparency Of health Research. <https://www.equator-network.org/reporting-guidelines/consort/> Accessed 4 April 2024.
- Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al.; Cochrane Methodology Review Group. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev* 2012;**11**:MR000030. <https://doi.org/10.1002/14651858.MR000030.pub2>
- McErlean M, Samways J, Godolphin PJ, Chen Y. The reporting standards of randomised controlled trials in leading medical journals between 2019 and 2020: A systematic review. *Ir J Med Sci* 2023;**192**:73–80. <https://doi.org/10.1007/s11845-022-02955-6>
- Zheng SL, Chan FT, Maclean E, Jayakumar S, Nabeebaccus AA. Reporting trends of randomised controlled trials in heart failure with preserved ejection fraction: A systematic review. *Open Heart* 2016;**3**:e000449. <https://doi.org/10.1136/openhrt-2016-000449>
- Web of Science. Journal Citation Report (JCR). Announcing the 2019 Journal Citation Reports. <https://clarivate.com/blog/announcing-the-2019-journal-citation-reports>. Accessed 29 March 2024.
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 2021;**372**:n71. <https://doi.org/10.1136/bmj.n71>
- Chan AW, Altman DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005;**365**:1159–1162. [https://doi.org/10.1016/S0140-6736\(05\)71879-1](https://doi.org/10.1016/S0140-6736(05)71879-1)
- Als-Nielsen B, Chen W, Gluud C, Kjaergard LL. Association of funding and conclusions in randomized drug trials: A reflection of treatment effect or adverse events? *JAMA* 2003;**290**:921–928. <https://doi.org/10.1001/jama.290.7.921>
- Siddiqi TJ, Shahid I, Arshad MS, Greene SJ, Pandey A, Vaduganathan M, et al. Inconsistent outcome reporting in heart failure randomized controlled trials. *J Card Fail* 2023;**29**:425–433. <https://doi.org/10.1016/j.cardfail.2022.11.008>
- Ghimire S, Kyung E, Kang W, Kim E. Assessment of adherence to the CONSORT statement for quality of reports on randomized controlled trial abstracts from four high-impact general medical journals. *Trials* 2012;**13**:77. <https://doi.org/10.1186/1745-6215-13-77>
- Li G, Abbade LPF, Nwosu I, Jin Y, Leenus A, Maaz M, et al. A scoping review of comparisons between abstracts and full reports in primary biomedical research. *BMC Med Res Methodol* 2017;**17**:181. <https://doi.org/10.1186/s12874-017-0459-5>
- Higgins JP, Altman DG. Assessing risk of bias in included studies. In: JPT H, Green S, eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: John Wiley & Sons; 2008. p187–241.
- Turner L, Shamseer L, Altman DG, Schulz KF, Moher D. Does use of the CONSORT statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. *Syst Rev* 2012;**1**:60. <https://doi.org/10.1186/2046-4053-1-60>
- Devereaux PJ, Manns BJ, Ghali WA, Quan H, Guyatt GH. The reporting of methodological factors in randomized controlled trials and the association with a journal policy to promote adherence to the Consolidated Standards of Reporting Trials (CONSORT) checklist. *Control Clin Trials* 2002;**23**:380–388. [https://doi.org/10.1016/s0197-2456\(02\)00214-3](https://doi.org/10.1016/s0197-2456(02)00214-3)
- Hopewell S, Boutron I, Chan AW, Collins GS, de Beyer JA, Hróbjartsson A, et al. An update to SPIRIT and CONSORT reporting guidelines to enhance transparency in randomized trials. *Nat Med* 2022;**28**:1740–1743. <https://doi.org/10.1038/s41591-022-01989-8>
- Cobo E, Cortés J, Ribera JM, Cardellach F, Selva-O'Callaghan A, Kostov B, et al. Effect of using reporting guidelines during peer review on quality of final manuscripts submitted to a biomedical journal: Masked randomised trial. *BMJ* 2011;**343**:d6783. <https://doi.org/10.1136/bmj.d6783>
- Friedman LM, Furberg CD, DeMets DL, Reboussin DM, Granger CB. *Fundamentals of Clinical Trials*. 5th ed. Cham: Springer; 2015. p550.