Article:
Hadi, MA orcid.org/0000-0003-0108-7833, McHugh, GA orcid.org/0000-0002-5766-5885 and Conaghan, PG orcid.org/0000-0002-3478-5665 (2017) Quality of reporting of harms in randomised controlled trials of pharmacological interventions for rheumatoid arthritis: a systematic review. BMJ Evidence-Based Medicine, 22 (5). pp. 170-177. ISSN 2515-446X

https://doi.org/10.1136/ebmed-2017-110715

Quality of reporting of harms in randomized controlled trials of pharmacological interventions for Rheumatoid Arthritis: A Systematic Review

Muhammad Abdul Hadi,*[a] Gretl A. McHugh,[b] Philip G. Conaghan,[c,d]

[a] Leicester School of Pharmacy, De Montfort University, Leicester, United Kingdom

[b] School of Healthcare, University of Leeds, Leeds, United Kingdom

[c] Leeds Institute of Rheumatology and Musculoskeletal Medicine, University of Leeds, Leeds, United Kingdom

[d] NIHR Leeds Musculoskeletal Biomedical Research Unit, Leeds, United Kingdom

***Corresponding author:**

Muhammad Abdul Hadi

Lecturer in Pharmacy Practice

Leicester School of Pharmacy

Room 00.11, Hawthorn Building

De Montfort University

LE1 9BH Leicester, United Kingdom

Tel: 0044 116 250 6771

Running title: Quality of harms reporting in RCTs

Word count: 3323

Abstract

Background: The quality of reporting of harms data in randomized controlled trials (RCTs) has been reported to be suboptimal. Rheumatoid arthritis (RA) has seen a massive growth in novel pharmacotherapies in the last decade. The aim of this study was to assess the quality of reporting of harms-related data in RCTs evaluating pharmacological interventions for Rheumatoid Arthritis according to the Consolidated Standard of Reporting Trials (CONSORT) statement harms-reporting extension.

Methods: RCTs published between January 2011 and August 2016 in the five highest impact factor journals in general medicine and two in rheumatology subject categories as per 2015 journal citation reports were included. Reports of secondary, supplementary or exploratory analyses of RCTs and non-inferiority trials were excluded. Two reviewers independently extracted data using a structured, pilot-tested 18-item questionnaire developed based on CONSORT harms-extension recommendations.

Results: 68 RCTs were included in the review. Out of a maximum harms reporting score of 18, the mean (SD) score was 8.51 (3.5) (Range = 0 to 15). More than half (56.5%) of the RCTs reported less ≤ 50% of items and only 3 (4.3%) RCTs reported more than 70% (score ≥14) of the items. Multilinear regression analyses found that region of trial origin (P = 0.01), sample size (P = 0.001) and whether the study was a long-term extension (LTE) of a trial or not (P = 0.04) were independent predictors associated with higher total harms reporting score.

Limitations: The study findings may not be generalizable to non-RA RCTs published in included journals.

Conclusions: The adherence to CONSORT harms-extension was poor in recently published RCTs of pharmacological interventions for RA. There is a need to improve quality of harms reporting in RCTs to allow transparent and balanced assessment of the benefit-risk ratio in clinical decision-making.

Key words: Rheumatoid Arthritis; Harms-reporting; Pharmacovigilance; randomized controlled trials; adverse drug reactions.

INTRODUCTION:

Adequate and transparent reporting of both effectiveness and harms data in randomized controlled trials (RCTs) are critical to allow clinicians to make an informed and balanced decision about the benefit-risk ratio of a particular drug/treatment. Sub-optimal reporting of adverse events (AEs) may create false perceptions of drug safety among clinicians leading to medication errors.[1]Transparent, comprehensive, and accurate reporting of AEs is not only important to ensure that clinicians make the appropriate decision for their patients, but also for the patients (the consumers of medicines) in understanding the risk associated with treatment, and for the regulatory agencies in approving and/or withdrawing a drug from market.

In 2004, the harms extension of Consolidated Standards of Reporting of Trials (CONSORT) was first published with an aim to improve quality of reporting of AEs in RCTs.[2] The original CONSORT statement[3] did not provide any specific guidance on reporting of harms-related data, however the 2001 revision[4] included a single item, still inadequate given the critical importance of harms data in clinical decision making. This prompted the development of the 2004 CONSORT harms extension. The CONSORT statement has improved quality of reporting of RCTs[5-8]. However, the reporting of AEs in RCTs still remain suboptimal.[1,9-14] A number of studies have found deficiencies and inaccuracies in relation with reporting of harms data in RCTs. Alarmingly, even important AEs are often under-reported.[1, 9-14]

Rheumatoid arthritis (RA) is a common inflammatory arthritis where pharmacological therapy with disease modifying anti-rheumatic drugs (DMARDs) is paramount and where a large number of new agents have been introduced in the last decade. Many of these are biologic agents with potential for serious adverse effects. The aim of this

systematic review was to evaluate the quality of reporting harms-related data in RCTs evaluating the effectiveness of pharmacological interventions for RA published in the top-tier medical journals.

METHODS:

Study Selection

We selected the top five highest impact factor journals in General and Internal medicine subject category which included The New England Journal of Medicine (NEJM), the Lancet, Journal of American Medical Association (JAMA), The British Medical Journal (BMJ) and The Annals of Internal Medicine and the top two journals in Rheumatology which included Annals of Rheumatic Diseases and Arthritis & Rheumatology subject category as per the 2015 Journal Citations Reports® (JCR). Although, Nature Reviews Rheumatology has the 2nd highest impact factor in the Rheumatology subject category in 2015 JCR, it was not included in this review as the journal does not publish primary research. Subsequently, Arthritis & Rheumatology, ranked as number 3 in the Rheumatology category was included. These journals were selected based on the assumption that the quality of reporting of RCTs in these top-tier journals is likely to be the best due to their rigorous peer-review and high-standard editorial checks during the submission and publication processes. The implications of using only a certain number of journals are discussed in the limitations section.

We searched Medline (via OVID) for RCTs evaluating the effectiveness of pharmacological interventions for the treatment of RA. To identify the recent RCTs we selected studies published between January 2011 and August 2016 in order to look at the latest trends in AEs reporting. The search was performed in August 2016 and a highly sensitive filter for detecting RCTs, developed by an expert health sciences

librarian, was used together with the key word “rheumatoid arthritis”. The following studies were excluded: Phase I or II RCTs; secondary, exploratory or pooled analysis of RCTs; non-RA RCTs; non-inferiority RCTs; RCTs of non-pharmacological interventions and systematic reviews and meta-analysis., Reports of long-term extensions (LTEs) of RCTs were included in the review.

All search results were transferred to an Endnote® file. After de-duplication, the title and abstracts were screened by one reviewer and full-text of the relevant articles were retrieved. The RCTs meeting the inclusion/exclusion criteria were included in the review.

Data Extraction and Development of Harms Reporting Scoring System

An 18-item checklist was developed based on the 2004 CONSORT extension on harms-reporting [2] and previously published literature.[9, 12, 15] Various checklists have been used in the literature with items ranging from 10 to 25.[9-12] The original CONSORT extension has only 10 recommendations[2] but with multiple items of interest within a single recommendation. Scoring the multiple items within a single recommendation would have been not only difficult but also misleading, therefore, where appropriate, the single CONSORT harms extension-items were split into two or three items resulting in 18-item checklist. Since subgroup analysis for AEs is rarely performed and reported, the ninth recommendation of the 2004 CONSORT extension was excluded from the checklist.[2] Each item of the 18-item checklist was scored individually and weighted with equal importance in line with CONSORT recommendations. Each item was scored as “1” if the item was adequately reported or “0” if it was not clearly reported or not reported at all. Adequacy of reporting of an item was judged on the recommendations of 2004 CONSORT extension [2]. The total harms reporting score

(THRS) was calculated by summing up all the individual scores with maximum and minimum scores of 18 and 0 respectively. In addition, data about trial characteristics (e.g. year of publication, journal, funding agency, trial origin) were also extracted using a structured form.

Both the trial characteristics questionnaire and 18-item checklist were piloted on six of the included RCTs. For all included studies, data were extracted by two reviewers independently. Discrepancies were resolved through discussion until a census was reached. In addition to full-text, where available, supplementary files and data associated with included trials were also used to extract any relevant data. For this review, the terms 'harms' and 'adverse events (AEs)' have been used interchangeably, as appropriate.

Data analysis:

Data were entered and analyzed using Statistical Package for Social Sciences (SPSS)^{TM} version 23. All continuous (ratio/interval) data were expressed as mean (±SD). A t-test or One-way ANOVA were used as appropriate to compare mean THRS across various trial characteristics. Specifically, a t-test was used to assess if the quality of reporting of harms data varied across: journal class (rheumatology journals vs general medicine journals); type of drug (biological vs non-biological); region of RCT origin (intercontinental vs regional); funding (industry vs non-commercial funding); the results of primary outcome (positive vs neutral) and long-term extension (Yes vs No). A One-way ANOVA was used to compare THRS across: sample size (≤ 200 vs ≤ 500 vs > 500); toxicity profile (comparable vs investigational arm more toxic vs not clear) and role of funding agency (funding only vs role in trial design vs not clear). A multivariate linear regression model was developed using a step-wise

forward approach to identify predictors associated with higher total harms-reporting score. Statistical assumptions for multilinear regression modelling including homoscedasticity, multicollinearity and multivariate normality were also tested and satisfied. All statistical tests were two-sided and a P-value less than 0.05 was considered statistically significant.

RESULTS:

Characteristics of RCTs included

Of the 347 records identified from the database search, the titles and abstract of 283 studies were screened after de-duplication. Full-texts of 146 studies were assessed for eligibility and 68 studies were included in the review (Appendix 1). The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) flow diagram including the reasons for exclusions are provided in Figure 1.

No RCTs were identified from JAMA and BMJ but more than two-thirds of the included trials (53; 77.9%) were published in Annals of Rheumatic Diseases and were funded by pharmaceutical industry (53; 77.9%) (Table 1). Almost all RCTs were multi-center (66; 97.1%) and more than half had intercontinental origin (37; 54.4%). Most of the RCTs (57; 82.6%) evaluated the effectiveness of a biological agent alone or in combination with other drugs used in the treatment of RA. Results of the primary efficacy outcome were reported to be positive (investigational arm more effective than the control arm) by most of the RCTs included in the review (56; 82.3%). The toxicity profile of the investigational arm was concluded by authors to be equivalent/comparable to the control arm in more than half of the RCTs (40; 58.0%).

In 27 (39.7%) trials, the funding agency had a role in designing or reporting of the trial.

Further details of the characteristics of the RCTs are presented in Table 1.

Table 1: Characteristics of included RCTs

Trial characteristics	**N (%)** **Total N = 68**
Year of Publication	
2011	12 (17.4)
2012	15 (21.7)
2013	18 (26.1)
2014	15 (21.7)
2015	6 (8.8)
2016	2 (2.9)
Journal	
Ann Rheum Dis	53 (77.9)
Arth & Rheumatol	2 (3.0)
Lancet	6 (8.8)
NEJM	5 (7.3)
Ann Intern Med	2 (3.0)
Type of drug under investigation	
Biological	56 (82.3)
Non-Biological	12 (17.7)
Region of RCT origin	
Intercontinental	37 (54.4)
Europe	24 (35.2)
Asia	4 (5.8)
Others	3 (4.3)
Sample Size	
≤ 200	20 (29.4)
201- 500	20 (29.4)
> 500	28 (41.2)
Funding	
Industry-funded	53 (77.9)
Non-industry funded	5 (7.4)
Mixed	9 (13.2)
Unclear	1 (1.5)
Results of primary outcome	
Positive	56 (82.3)
Neutral	12 (17.7)

Toxicity profile	
Equivalent/No now safety signal/comparable	39 (57.3)
Investigational arm more toxic	6 (8.8)
Control arm more toxic	0 (0)
No conclusion	23 (33.8)
Study sites	
Single centre	2 (2.9)
Multi-centre	66 (97.1)
Blinding	
Single	5 (7.2)
Double	40 (58.8)
Double blind followed by open label	12 (17.5)
Open label followed by double blind	9 (13.2)
Open label/Open label extension	2 (2.9)
Role of funding agency	
Funding only	17 (25.0)
Funding plus design and conception of trial	3 (4.3)
Funding plus analysis and reporting	7 (10.2)
Funding plus design, conception, analysis and reporting	17 (25.0)
unclear	24 (34.8)
Long-term extension (LTE)	
Yes	16 (23.6)
No	52 (76.4)

Ann Rhem Dis =Annals of Rheumatic Diseases; Arth &Rheumatol = Arthritis and Rheumatology; NEJM = The New England Journal of Medicine; Ann Intern Med = Annals of Internal Medicine

Reporting of the expanded CONSORT harms items

For the total 18 items, more than half of the RCTs reported less than 50% (n=9) of the CONSORT items and only 3 (4.3%) RCTs reported more than 70% of the items. The number (and percentages) of the RCTs fulfilling each of CONSORT harms recommendation are presented in Table 2. More than two thirds of the RCTs (53; 76.8%) mentioned AEs in title or abstract (CONSORT recommendation 1). However, only a few trials provided information on AEs in the introduction section (9; 13.0%), used a validated scale to measure severity of AEs (8; 11.6%) and gave a definition of

AEs (9; 13.0%) (CONSORT recommendations 2 and 3). Less than one third of the trials (20; 29.0%) described how AE related data were collected (CONSORT recommendation 4 (4a). Less than a quarter of the trials (16; 23.2%) described methods of presenting and/or analyzing AEs (CONSORT recommendation 5). A vast majority of RCTs described AEs leading to death (60; 87.0%) and number of withdraws caused by AEs in each arm (59; 85.5%). However, slightly more than a quarter (18; 26.1%) of RCTs provided a description of the AEs which resulted in patient withdrawals (CONSORT recommendation 6). The majority of trials presented results for each arm separately (63; 91.3%) and presented a balanced discussion on both safety and efficacy of drug (46; 66.7%) (CONSORT recommendations 8 and 10).

Harms reporting score and associated trials characteristics

Out of a maximum score of 18, the mean total harms reporting score (THRS) was 8.51 (SD 3.5) (range 0 – 15). The mean THRS was significantly higher for those RCTs published in general medicine journals as compared to rheumatology journals (P= 0.02) (Table 3). Similarly, the quality of harms reporting score was significantly higher for RCTs involving biologicals (P= 0.07), of intercontinental origin (P = 0.001) and with a sample size more than 500 participants (P= 0.001) compared with RCTs of non-biologicals, regional/national trials and trials involving ≤ 200 participants respectively. Interestingly, the RCTs where the trial sponsor had no role for any aspect of designing and/or reporting of the trial had lower total scores as compared to the RCTs in which the trial sponsor played a role in at least one aspect of trial design or reporting (P= 0.004) (Table 3)

Multiple linear regression analysis was undertaken to identify predictors of higher total harms reporting score (Table 4). Region of trial origin (intercontinental vs others) (P =

0.01), sample size (≤ 200 vs > 200) (P = 0.001) and whether the study was a long-term extension (LTE) of a trial or not (P = 0.04) were found to be independent predictors associated with higher THRS. Intercontinental RCTs had a THRS on an average 1.9 points higher than national/regional RCTs (95% CI = 0.3 to 3.5). Similarly, trials with a sample size > 200 had THRS on an average 2.9 points higher than RCTs with ≤ 200 patients (95% CI = 1.2 to 4.6). Journal type, funding source and type of drug were not significantly associated with higher THRS. The model explained 36% variation in total harms reporting score (Table 4).

Table 2: Items adequately reported against quality of reporting criteria.

Recommendations of 2004 CONSORT- HARMS extension	Quality of reporting criteria	Item reported N (%)
1. If the study collected data on harms and benefits, the title of abstract should state so.	1. AE mentioned in the title or abstract	53 (76.8)
	AEs mentioned in the title	20 (29.0)
	AEs mentioned in the abstract	33 (47.8)
2. If the trial addresses both harms and benefits, the introduction should so state.	2. Information on AEs mentioned in introduction	9 (13.0)
3. List addressed adverse events with definitions for each (with attention, when relevant, to grading, expected vs. unexpected events, reference to standardized and validated definitions, and description of new definitions).	3a. If article mentioned use of validated instrument to report AE severity	8 (11.6)
	3b. If article mentioned definition of AE	9 (13.0)
4. Clarify how harms-related information was collected (mode of data collection, timing, attribution methods, intensity of ascertainment, and harms-related monitoring and stopping rules, if pertinent).	4a. Description of how harms data were collected (e.g. diaries, phone interviews, f-2-f interviews)	20 (29.0)
	4b. Description of when AE data were collected	32 (46.4)
	4c. Whether or not AEs were attributed to trial drug (e.g. how AEs were attributed to drugs)	10 (14.5)
5. Describe plans for presenting and analyzing information on harms (including coding, handling	5. Description of methods for presenting and/or analyzing AEs	16 (23.2)

of recurrent events, specification of timing issues, handling of continuous measures, and any statistical analyses).		
6. Describe for each arm the participant withdrawals that are due to harms and the experience with the allocated treatment.	6a. If the article reported number of withdraws caused by AEs in each arm	59 (85.5)
	6b. Description of AEs leading to withdrawals	18 (26.1)
	6c. Description of AEs leading to death	60 (87.0)
7. Provide the denominators for analyses on harms.	7a. If the article provided denominators for AEs	45 (65.2)
	7b. If the article provided definitions used for analysis set (ITT, Per protocol, safety data available, unclear)	39 (56.5)
8. Present the absolute risk of each adverse event (specifying type, grade, and seriousness per arm), and present appropriate metrics for recurrent events, continuous variables and scale variables, whenever pertinent.	8a. Results presented separately for each arm	63 (91.3)
	8b. Separate reporting of severe AEs (grade> 2 or serious AEs)	61 (88.4)
	8c. Provided both number of AEs and number of patients with AEs	18 (26.1)
9. Describe any subgroup analyses and exploratory analyses for harms.	--------	
10. Provide a balanced discussion of benefits and harms with emphasis on study limitations, generalizability, and other sources of information on harms.	10a.. If the discussion was balanced with regards to efficacy and AEs	46 (66.7)
	10b. Limitations of the study specifically in relation to AEs discussed?	21 (30.4)

Table 3: Comparison of total quality of harms reporting score across trial characteristics

Trial characteristics	N (%)	Total quality of reporting score Mean (SD)	P-value
Journal class			0.02*
Rheumatology journals	55 (80.8)	8.1 (3.6)	
General Medicine journals	13 (19.2)	10.3 (2.6)	
Type of drug under investigation			0.07*
Biological	56 (82.3)	9.1 (2.9)	
Non-Biological	12 (17.7)	6.0 (5.1)	
Region of RCT Origin			0.001*
Intercontinental	37 (54.4)	10.1 (2.4)	
National/Regional	31 (45.6)	6.5 (3.8)	
Sample Size			0.001‡
≤ 200	20 (29.4)	5.9 (3.9)	
≤ 500	20 (29.4)	8.7 (3.3)	
> 500	28 (41.2)	10.1 (2.4)	
Funding			0.03*
Industry-funded	53 (77.9)	8.9 (3.8)	
Non-commercially funded/Mixed	15 (22.1)	6.8 (3.4)	
Results of primary outcome			0.28*
Positive	56(82.3)	8.8 (3.1)	
Neutral	12 (17.7)	7.0 (5.1)	
Toxicity profile			0.01‡
Comparable/No new safety signal	39 (57.3)	8.8 (2.7)	
Investigational arm more toxic	6 (8.8)	11.5 (1.5)	
No conclusion	23 (34.2)	7.1 (4.6)	

Blinding	5 (7.2)	7.4 (2.3)	0.51‡
Single	40 (58.8)	8.8 (3.7)	
Double	12 (17.3)	9.1 (3.4)	
Double blind followed by open label	9 (13.0)	6.7 (3.9)	
Open label followed by double blind	2 (2.8)	8.0 (1.4)	
Open label/Open label extension			
Role of funding agency			0.004‡
Funding only	17 (24.6)	6.4 (4.5)	
Involvement in any aspect of trial	27 (39.7)	9.9 (2.4)	
design	24 (34.8)	8.2 (3.3)	
Unclear			
Long-term Extension (LTE)			
No	52 (76.4)	7.6 (4.1)	0.26*
Yes	16 (23.2)	8.7 (3.4)	

*P-value calculated using t-test. ‡ P-value calculated using One-way ANOVA.

Table 4: Multiple linear regression analysis of total CONSORT harms score

Variable (coding)	B (95% CI)	P-value
Region	1.9 (0.3 to 3.5)	0.01
(0 = others; 1 = Intercontinental)		
Journal type	- 0.5 (-2.3 to 1.3)	0.56
(0 = Medicine; 1 = Rheumatology)		
Type of drug	1.8 (- 0.8 to 3.8)	0.06
(0 = Non-biological; 1=biological)		
LTE	-1.7(-3.4 to -0.05)	0.04
(0 = No; 1 = Yes)		
Sample size	2.9 (- 1.2 to -4.6)	0.001
(0 = ≤200; 1= > 200)		
Funding	0.18 (-2.4 to 1.3)	0.84
(0 =Others; 1 = Industry funded)		

Adjusted R^2 for the model = 0.36. B= Unstandardized Beta-coefficient for the regression model; CI= Confidence Interval

Methods of Presenting Harms data

More than two-thirds of the RCTs (54; 79.7%) presented harms data in text and tables. More than one third of the RCTs (21; 30.4%) reported frequent and serious AEs only. The definition used by authors to describe and present “frequent AEs” varied among RCTs (AEs affecting 1% to 5% of patients). Most of the studies (59; 86.7%) did not describe the scale used to measure the severity of AEs. In most instances, AEs data was presented as frequencies only (44; 64.7%) without any statistical comparison between the investigational and control arms (Table 5).

Table 5: Presentation of Harms data in included trials

Presentation of AEs	N (%)
Mode of presentation	
Text only	3 (4.3)
Text and table	54 (79.4)
Text and figure	1 (1.4)
Text, table and figure	5 (7.3)
Not presented	5 (7.3)
Attribution of AEs to trial drugs	
Yes	19 (27.9)
No	5 (7.3)
Unclear	44 (64.7)
Selection of AEs reported	
Sever/Serious only	15 (21.9)
Frequent including sever	21 (30.3)
AEs selected by investigator	7 (10.2)
Unclear	24 (35.3)
Not reported	1 (1.4)
Scale used to report AE severity	
CTCAE	5 (7.4)
WHO	1 (1.4)
Others	3 (4.4)
No or unknown scale	59 (86.7)
Statistical comparison of AE rates between trial arms	
Yes	4 (5.8)
No	44 (64.7)
Partial*	18 (26.5)

*Only selected AEs were compared statistically. CTCAE =Common Terminology Criteria for Adverse Events; WHO = World Health Organization; AE = Adverse event

DISCUSSION

This review assessed the quality of reporting of harms-related data in RCTs evaluating pharmacological interventions for the treatment of RA in seven top-tier general medicine (five) and rheumatology (two) journals. In general, the adherence to CONSORT harms-extension was poor in recently published RCTs. Previously, a review has evaluated the quality of harms reporting in RCTs of non-pharmacological interventions for rheumatic diseases published between 1999 and 2005 and found it to be suboptimal.[16]

We found great variations in the reporting of individual items of CONSORT harms extension across trials. Some of the recommendations including: mentioning AEs in the title or abstract (CONSORT item 1); providing number of withdrawals in each arm (item 6a); describing of AEs leading to death (item 6b); separate reporting of results for each arm (item 8a); and separate reporting of severe/serious AEs (item 8b) were adequately reported by the majority of RCTs. However, certain critical elements including: description of AEs leading to withdrawals (item 6b); process of attributing AEs to trial drug (item 4c); description of methods for presenting and analyzing AEs (item 5); and using a validated instrument to report AE severity were poorly reported. It is possible that authors might have collected this information but could not report it due to restrictions on manuscript length. Manuscript length can be one of the reasons for not adequately reporting AEs.[9] The option of 'online only' supplement is offered by almost all journals now which could be used to report additional harms related data.

The mean THRS for RCTs that were industry-funded was significantly higher than for non-industry funded RCTs. However, industry funding was not found to be an independent predictor in multiple linear regression. Country of trial origin, sample size

and whether the study was a long-term extension of a previously published trial or not were independent predictors significantly associated with higher THRS. Reviews evaluating quality of harms reporting in oncology[9] and analgesic[10] RCTs also reported better reporting of AEs in trials funded by industry. This may be explained by tighter control by regulatory agencies for industry funded trials, better and thorough data collection capabilities and soliciting services of professional medical writers in manuscript writing.[9]

Lower THRS for LTEs (long-term extensions) found in the present study can be attributed to authors' assumption that AEs have already been sufficiently reported in the primary paper and need not be reported again in LTEs. However, this is a false assumption as thorough and transparent collection and reporting of all AEs is critically important for LTEs as well to establish the long-term safety of drugs and to identify rare AEs.[1] We found that AEs reporting was significantly better for RCTs evaluating biologicals compared to non-biologicals. This is perhaps because the use of biologicals for the treatment of RA is relatively new and there are reservations regarding their safety, especially long-term safety. On the other hand, safety profiles are well established for traditional DMARDs and this may have prompted authors to focus primarily on the efficacy rather than safety.

In line with our study findings, a number of reviews assessing the quality of AE reporting in various medical specialties have reported critical inadequacies in reporting of AEs.[1, 9-12, 15] Under-detection and inaccurate reporting of AEs not only in RCTs but also in clinical practice can have serious negative consequences in relation to ensuring patient safety.[1] Empirical research exploring the reasons of underreporting of harms-related data in RCTs is almost non-existent. Various reasons including manuscript length, neglecting accurate collection, interpretation and presentation of harms data,

and lack of authors' interest in reporting harms data have been cited in the literature for underreporting of harms data.[1, 9,10, 17] Another challenge commonly encountered by clinicians is to establish a cause-effect relationship between a particular side effect and investigational drug especially when the patient is taking multiple drugs with overlapping toxicity profiles and when disease symptoms/complications are similar to adverse effects. There is a need to do more advocacy by creating awareness of the CONSORT harms extension recommendations through endorsement, undertaking more research about quality of harms reporting and implementing stringent editorial processes to ensure adherence to CONSORT harms recommendations. Only one of the seven included journals, the Lancet, explicitly recommends authors to follow the CONSORT harms extension in their instructions to authors. However, all these journals require all submitted RCT reports to follow the 2010 CONSORT statement.[18] Furthermore, since 2004 the CONSORT harms extension has received only 604 citations in SCOPUS to date (Date of search 09-11-2016) compared to over 3000 citations for the 2010 CONSORT statement which was published in multiple medical journals. This clearly shows that adequate and accurate reporting of harms data, despite its critical importance, is less emphasized and left to the discretion of authors. We suggest that journals should make it compulsory for authors of RCTs to submit a CONSORT harms extension checklist at manuscript submission, with the page number identified where an item has been reported. Although burdensome, peer-reviewers could counter check the methods for collecting, presenting and analyzing harms data against the protocol and trial registry.[1] These measures are likely to improve reporting of AEs in RCTs allowing clinicians and patients to understand risks associated with a particular treatment.

Limitations

There are several methodological limitations which should be considered when interpreting the review findings. Firstly, the findings of the review may not be generalizable to non-RA RCTs published in the same journals and/or RCTs published in other journals. There is also potential for publication bias in relation to selection of journals as only seven journals were searched. However, it should be noted here that the selected journals are among the most cited and respected journals in their respective subject categories. Therefore, the quality of reporting of AEs in these journals is likely to be at least comparable to other journals, if not superior. Secondly, more than two-thirds of the RCTs included in the review were published in Annals of Rheumatic Diseases which might have significantly influenced the overall findings of the review as journals usually a "pre-specified" style of reporting RCTs. However, given that none of the RCTs reported all the CONSORT harms-extension items and only three RCTs reported more than 70% (13 out of 18) of the items clearly indicates that journal's "pre-specified" style is unlikely to play a major role in suboptimal reporting of harms-related data in RCTs. Furthermore, underreporting of harms-related data documented previously in other medical specialties further support the argument that there are other reasons beyond the journals' style contributing to underreporting. Thirdly, we only reviewed published trial report and supplementary files but did not review trial protocol and trial registry. It is possible that authors might have defined and described certain elements of CONCORT harms recommendations (e.g. definition of AEs, methods of presenting and analyzing harms data) in the protocol but not in the final paper. However, the CONSORT statement recommends that this information should be included in the final report as well. Finally, each item in the checklist was weighted and scored equally irrespective of its importance in clinical decision making. For example, AEs information mentioned in the introduction (item 1) is far less

important than the description of AEs leading to death (item 6) in order to assess benefit-risk ratio. Although previously used frequently in the literature, reporting a cumulative score (THRS) (by adding each item score) is not recommended by the CONSORT harms-extension. However, weighing each item equally was in line with CONSORT recommendations[2] as it does not give priority to any item over the other because these recommendations provide generic guidance on harms reporting not a scoring system to assess quality of harms reporting. Therefore, THRS may not truly reflect deficiencies in reporting of individual items of CONSORT recommendations. Until the CONSORT harms-extension is revised and updated, the currently available version should be implemented as "minimum acceptable standard" for reporting harms data irrespective of the impact individual items on the assessment of benefit-risk ratio in clinical decision making., In the present study, in order to ensure transparency of study findings, data on both individual items and overall score have been presented (Table 2).

CONCLUSION:

The adherence to CONSORT recommendations for harms reporting in RCTs evaluating pharmacological interventions for RA in leading medical and rheumatology journals is suboptimal. It is a moral, ethical and scientific duty of all stakeholders involved in designing, conducting and reporting of RCTs that harms-related data is adequately collected and accurately reported without any personal or commercial interest. Peer-reviewers and editors should carefully review that the AEs have been reported in RCTs in compliance with the recommendations of the CONSORT harms extension. This is especially important for RCTs evaluating safety and efficacy of new drugs and drugs with a narrow therapeutic index. Without transparent, fair and responsible reporting of AEs in RCTs, evidence based clinical decisions may fail to

achieve therapeutic goals and compromise patient safety. Although, there is no single ideal solution to the challenges associated with collecting and reporting of harms-related data in RCTs, the adherence to CONSORT guidance can help authors and reviewers to improve transparency and ensure completeness of reporting of harms-related data.

Grant support: This study received no funding from any public or private agency.

Conflicts of Interest: MAH and GM has no conflicts of interest to disclose. PGC is on the speaker's bureau of or has acted as a consultant for AbbVie, Bioiberica, BMS, Flexion, Lilly, Merck Serono, Novartis, Pfizer and Roche. PGC is supported in part by the UK National Institute for Health Research (NIHR) Leeds Musculoskeletal Biomedical Research Unit. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Author Contributions:

Conception and design: M.A. Hadi, G. A. McHugh, P. Conaghan

Data extraction: M.A. Hadi, G.A. McHugh

Analysis and interpretation of the data: M.A. Hadi, G.A. McHugh

Drafting of the article: M.A. Hadi, G.A McHugh

Critical revision of the article for important intellectual content: P. G. Conaghan

Final approval for publication: M.A. Hadi, G.A. McHugh, P. G. Conaghan

References

1. Seruga B, Templeton AJ, Badillo FEV, et al. Under-reporting of harm in clinical trials. The Lancet Oncol 2016;**17**(5):e209-e19
2. Ioannidis JP, Evans SJ, Gøtzsche PC, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Ann Intern Med 2004;**141**(10):781-88
3. Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. BMJ1996;**313**(7057):570
4. Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001;**134**(8):663-94
5. Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. Med J Aust 2006;**185**(5):263
6. Kane RL, Wang J, Garrard J. Reporting in randomized clinical trials improved after adoption of the CONSORT statement. J Clin Epidemol 2007;**60**(3):241-49
7. Devereaux P, Manns BJ, Ghali WA, et al. 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**(4):380-88
8. Han C, Kwak K-p, Marks DM, et al. The impact of the CONSORT statement on reporting of randomized clinical trials in psychiatry. Contemp Clin Trials 2009;**30**(2):116-22

9. Péron J, Maillet D, Gan HK, et al. Adherence to CONSORT adverse event reporting guidelines in randomized clinical trials evaluating systemic cancer therapy: a systematic review. J Clin Oncol 2013;**31**(31):3957-63
10. Smith SM, Chang RD, Pereira A, et al. Adherence to CONSORT harms-reporting recommendations in publications of recent analgesic clinical trials: an ACTTION systematic review. Pain 2012;**153**(12):2415-21
11. Hodkinson A, Kirkham JJ, Tudur-Smith C, et al. Reporting of harms data in RCTs: a systematic review of empirical assessments against the CONSORT harms extension. BMJ Open 2013;**3**(9):e003436
12. Carmichael K, Nolan SJ, Weston J, et al. Assessment of the quality of harms reporting in non-randomised studies and randomised controlled studies of topiramate for the treatment of epilepsy using CONSORT criteria. Epilepsy Res 2015;**114**:106-13
13. Turner L-A, Singh K, Garritty C, et al. An evaluation of the completeness of safety reporting in reports of complementary and alternative medicine trials. BMC Complement Altern Med 2011;**11**(1):1
14. Pitrou I, Boutron I, Ahmad N, et al. Reporting of safety results in published reports of randomized controlled trials. Arch Intern Med 2009;**169**(19):1756-61
15. Trotti A, Bentzen SM. The need for adverse effects reporting standards in oncology clinical trials. J Clin Oncol 2004;**22**(1):19-22
16. Ethgen M, Boutron I, Baron G, et al. Reporting of harm in randomized, controlled trials of nonpharmacologic treatment for rheumatic disease. Ann Intern Med 2005;**143**(1):20-25
17. Scharf O, Colevas AD. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. J Clin Oncol 2006;**24**(24):3933-38

18. Schulz KF, Altman DG, Moher D, et al. Consort 2010 statement: Updated guidelines for reporting parallel group randomized trials. Ann Intern Med 2010;**152**(11):726-32