Supplement: Applicability of trials in rheumatoid arthritis and osteoarthritis: A systematic review and meta-analysis of trial populations showing adequate proportion of women, but underrepresentation of elderly people

Andriko Palmowski, Thomas Buttgereit, Yannick Palmowski, Sabrina M. Nielsen, Maarten Boers, Robin Christensen, Frank Buttgereit

This version is available at
https://doi.org/10.17169/refubium-26668

Citation details
DOI: 10.1016/j.semarthrit.2018.10.017

Terms of use
This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International license: https://creativecommons.org/licenses/by-nc-nd/4.0/
Applicability of trials in rheumatoid arthritis and osteoarthritis:
A systematic review and meta-analysis of trial populations showing adequate proportion of women, but not enough elderly people

Andriko Palmowski, Thomas Buttgereit, Yannick Palmowski, Sabrina M Nielsen, Maarten Boers, Robin Christensen, Frank Buttgereit

Table of contents

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Protocol</td>
<td>2</td>
</tr>
<tr>
<td>References of included population-based studies</td>
<td>14</td>
</tr>
<tr>
<td>References of included randomized controlled trials</td>
<td>17</td>
</tr>
<tr>
<td>Risk of bias assessment</td>
<td>38</td>
</tr>
<tr>
<td>Meta-analyses</td>
<td>42</td>
</tr>
<tr>
<td>Studies on underrepresentation of the elderly in clinical trials</td>
<td>47</td>
</tr>
<tr>
<td>PRISMA Checklist</td>
<td>49</td>
</tr>
<tr>
<td>MOOSE Checklist</td>
<td>50</td>
</tr>
</tbody>
</table>
January 19th, 2018

Comparability of age distribution and proportion of elderly individuals included in randomized trials and real-world population-based studies: Protocol for a systematic review and meta-analysis of publications in rheumatoid arthritis and osteoarthritis

Andriko B. A. Palmowski¹, Thomas Buttgereit¹ ², Yannick Palmowski³, Dörte Huscher⁴, Maarten Boers⁵ ⁶, Robin Christensen⁷, Frank Buttgereit¹

Affiliations
¹ Department of Rheumatology and Clinical Immunology, Charité – University Medicine, Berlin, Germany.
² Department of Dermatology, Venerology and Allergology, Charité – University Medicine, Berlin, Germany. ³ Center for Musculoskeletal Surgery, Charité – University Medicine, Berlin, Germany.
⁴ Epidemiology Unit, German Rheumatism Research Centre, Berlin, Germany.
⁵ Department of Psychology, Health and Technology, University of Twente, Enschede, The Netherlands.
⁶ Department of Epidemiology & Biostatistics; Amsterdam Rheumatology & Immunology Center, VU University Medical Center, Amsterdam, The Netherlands.
⁷ Musculoskeletal Statistics Unit, The Parker Institute, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark.

Corresponding author: Frank Buttgereit, MD, Prof; Department of Rheumatology and Clinical Immunology, Charité University Medicine, Berlin, Germany; Charitéplatz 1, 10117 Berlin, Germany: Phone: ++49 30 450 513125, Fax: ++49 30 450 513917, e-mail: frank.buttgereit@charite.de.

This project is part of the GLORIA trial (Glucocorticoid low-dose outcome in rheumatoid arthritis study; http://www.gloriatrial.org/; registered on https://clinicaltrials.gov/; identifier NCT02585258) and has received funding from the European Union’s Horizon 2020 Framework Programme for Research and Innovation under grant agreement No. 634886. Musculoskeletal Statistics Unit, The Parker Institute, (Robin Christensen) is supported by grants from The Oak Foundation.
SUMMARY

Rationale: The elderly are underrepresented in clinical trials, but the extent is unknown. Underrepresentation may limit the applicability of such trials as elderly individuals differ from younger adults in multiple aspects. The number of elderly is expected to rise dramatically in countries all over the world, increasing the number of people at risk of treatment with agents insufficiently tested in their age group. This certainly applies to osteoarthritis (OA) and rheumatoid arthritis (RA) where the prevalence in cross-sectional samples increases with age.

Objective: To evaluate the representation of elderly people with OA and RA in randomized controlled trials (RCTs) compared to the age distribution of patients identified in population-based studies.

Data sources: We will conduct two systematic searches on OA and RA in the online database MEDLINE (via PubMed): one on RCTs and one on population-based studies.

Study eligibility criteria: For both OA and RA, we will include all therapeutic RCTs published in 2016 and 2017 (both years included); and all population-based studies published between 2013 and 2017 that present real-world data on age from the two conditions, including publications drawing their data from registries.

Trial interventions: No restrictions.

Outcomes: Age distribution in RCTs and population-based studies including as summary measure mean and standard deviation, and the proportion of people aged 65 years or more. A secondary outcome is the percentage of female participants in RCTs compared to population-based studies.

Critical appraisal: We will assess the population-based studies’ internal validity with the JBI Critical Appraisal Checklist for Prevalence Studies. As we do not draw any conclusions from results or conclusions of included RCTs, we will not perform quality assessment on these studies. This allows for inclusion of all retrieved RCTs.

Synthesis methods: We will conduct two independent meta-analyses (for each major outcome) to estimate (i) the combined mean age, (ii) the pooled standard deviation thereof, and (iii) the proportions of people aged 65 years or more in RCTs and population-based studies, respectively. These independent estimates will subsequently be statistically compared using a 2-sample z-test. The same procedure applies to the percentage of female participants.
INTRODUCTION

Background
The elderly – commonly defined by an age of 65 or more years – are significantly underrepresented in clinical trials, as shown for a wide variety of diseases and throughout various medical specialties.[1-3] This poses a potentially serious problem as older people differ from younger adults in multiple aspects including pharmacodynamics and -kinetics, comorbidity, polypharmacy and physical performance, all of which affect the potential for benefit and usually increase the overall chance of drug-related adverse events.[4-7] In addition, the number of elderly is expected to rise dramatically in countries all over the world,[8, 9] leading to higher numbers of elderly (likely in need of healthcare). This particularly applies to osteoarthritis (OA) and rheumatoid arthritis (RA) as both have been shown to be age-related. [10, 11]

Applicability (generalizability or external validity) assesses the question of trial results being valid in other patients than those included in the original study population.[12, 13] An inadequate representation of elderly individuals in clinical trials for OA and RA will significantly reduce the applicability of trial results to the general population and lower the credibility of the conclusions that are being drawn from those trials.

Rationale
Evidence synthesis of the representation of elderly people, and thus trial result applicability, in rheumatology trials in general is still lacking to our knowledge. We have chosen RA and OA as exemplars as they are common musculoskeletal conditions with increasing prevalence (and burden) at higher ages.[10, 11, 14]

Objective
To evaluate the representation of elderly people in OA and RA RCTs compared to the age distribution of patients identified in population-based studies. In addition, we will also assess the frequency of explicit exclusion of elderly people; whether representation differs across interventions and the type of funding source; if and how often authors assess whether trial results differ across age groups (e.g., present trial results stratified by age); and if there are differences in the gender proportions in RCTs and population-based studies.
METHODS

Protocol and registration
Our protocol is registered on PROSPERO (https://www.crd.york.ac.uk/PROSPERO); identifier CRD42018085409). The protocol conforms to the Preferred Reporting Items for Systematic Review and Meta-analysis protocols (PRISMA-P) guidelines for reporting a protocol for a systematic review and meta-analysis.[15]

Eligibility criteria
We will include therapeutic RCTs and population-based studies, including those presenting data from registries, on OA and RA that report at least the mean (or the median) age of the specific disease-group. Studies on pediatric patients (e.g., juvenile arthritis) will be excluded. To assess current developments and explore recent trends, and also for reasons of feasibility, we have decided to limit our search to RCTs published in the last two years; i.e., the period from January 1st, 2016 to Dec 31st, 2017. Corresponding to participating GLORIA collaborators and their language skills we will impose a minor restriction concerning languages: Publications in other than English, German, French, Spanish, Portuguese, Dutch, Slovakian, Italian, Romanian and Hungarian will be excluded.

If multiple reports are derived from the same data set, we will attempt to include the latest published findings. Prevention and phase I clinical trials will be excluded. We have come to this decision because our aim is to analyze studies exhibiting direct impact in terms of treatment decisions or the approval of new drugs. Secondary analyses and (ancillary) reports of multiple studies and of studies already included will be excluded. Studies reporting age data from study completers only will be excluded due to the risk of age bias, i.e., we suspect younger patients might have a higher probability of completing a study.

In the search for population-based studies, we will include studies published between January 1st, 2013 and December 31st, 2017 in order to get sufficient data on the real-world age distribution. We will only include studies in which the disease was described as apparently being diagnosed by a physician. As we aim for prevalence data, studies providing data on incidence only (new cases over time) will be excluded. The same holds true for studies reporting only on specific patients, e.g., studies reporting data on depressive patients suffering from RA only.

Information sources
We will search the online biomedical and life science database MEDLINE via PubMed. Note that of the five domains addressed by the PICOS-mnemonic (Patient/Population; Intervention; Comparison; Outcome; Study design) recommended by the Cochrane Musculoskeletal Group (CMSG),[16] only the respective first and last sections of the acronym are applicable parts of our search strategies.
Search strategy
We have constructed an elaborate search strategy for the two diseases and study types we include in our review (Appendix A). The search will be performed in January 2018 in order not to miss recent trials. Additionally, we will perform a hand search for relevant publications including a scan of the references of major guidelines and reviews of the two diseases we address.

Data management, collection, items and analysis
We will use Microsoft Excel 2016 (Microsoft Corp., Redmond, WA, USA), and SAS (SAS Institute, Cary, NC, USA) software for data extraction, management and analysis. We will extract data using predefined data extraction sheets which are derived from the Cochrane Collaboration’s recommendations for data extraction and modified for our purposes.[17] The data extraction items are available in Appendix B at the end of this protocol. If trial or patient characteristics had been described in a preceding publication, we will attempt to extract them from the previous research manuscript.

Study selection
Retrieved articles will be imported into EndNote X8 software (Clarivate Analytics, Philadelphia, PA, USA). Two Authors (AP and TB) will independently screen the articles for inclusion or exclusion. They will eliminate duplicates with the help of EndNote X8 software,[18] screen the articles by title and abstract, and then assess the remaining articles in full text. Afterwards, consensus on study inclusion will be achieved between the two reviewers – if necessary, by consultation of a third reviewer. We will provide flow diagrams as Moher et al. proposed in the PRISMA statement.[19]

Outcomes
Age distribution including as summary measures mean and standard deviation, and the proportion of people aged 65 years or more. The proportions will either be abstracted directly from the research manuscript or be derived from an assumed normal distribution (see Synthesis of results for more information). The proportion of female participants is a considered a minor outcome.

Risk of bias in individual studies
Two authors (AP and TB) will independently perform a bias assessment on the population-based studies with the JBI Critical Appraisal Checklist for Prevalence Studies by the Joanna Briggs Institute.[20] In case of rating discrepancies consensus will be achieved, if necessary, by consultation of a third reviewer. We will not exclude population-based studies by their quality ratings alone, but rather present the results of this appraisal in our final article to allow for bias assessment. We will only extract
participant’s baseline characteristics in included RCTs, but no trial results, so that quality appraisal is not necessary. This allows for inclusion all retrieved RCTs that report the required data.

**Synthesis of results**

First of all, we will present descriptively analyzed study characteristics. We will then conduct meta-analyses based on three effect sizes derived from the major outcomes mean age, standard deviation thereof, and the proportion of people aged 65 or more years, and on the effect size derived from the secondary outcome, which is the percentage of female participants. These analyses will be done for RCTs and population-based studies independently. For each type of random effects meta-analysis, heterogeneity across included studies will be evaluated with Cochran’s Q-statistic[21] and presented as an I² value. I² measures the total percentage of variance across studies due to clinical heterogeneity rather than statistical error.[22] For statistical analyses, a two-sided significance level α is set at .05.

We will estimate the proportion of people aged ≥65 years under the assumption that age is distributed normally. However, we will consider all age distributions of all studies to be singly truncated at a lower age level, even if this is not reported (as a diagnosis of RA implies an age of at least 18 years). Studies that also employ an upper age limit are assumed to have a doubly truncated age distribution. For each study, we will derive individual cumulative distribution functions \( F_a(x) \) from study mean ages and their standard deviations in order to obtain the cumulative integral of individual probability density functions \( f_a(x) \), to obtain the proportion of patients at or below the age of \( x \). This allows for the calculation of the proportion of people aged 65 or more years (\( x \) equals the participant’s age; \( a \) is the respective study ID) by employing the following formula:

\[
(1 - F_a(64)) \times 100 = \text{Percentage of people aged 65 years or more in Study } a
\]

We will follow the instructions of the Cochrane Collaboration’s *Handbook for Systematic Reviews of* to combine the results of multiple study arms into one single group per trial.[17] If studies report median age and interquartile range (IQR) instead of mean age and standard deviation (SD), we will follow the Cochrane Collaboration’s guidelines and assume equality of the median and the mean (i.e., both representing a “central estimate”), and equality of the IQR and 1.35 SDs (i.e., both measures represent dispersion).[17] If it is the case that studies report median age and the range thereof (maximum and minimum), we will follow the method proposed by Hozo et al. to estimate mean age and standard deviation,[23] which includes the following:

Say \( \bar{x} \) is the mean, \( a \) the minimum, \( b \) the maximum, \( R \) the range (i.e., maximum – minimum), \( m \) the median, \( SD \) the standard deviation, and \( n \) the sample size, then:
if \( n \leq 15 \), then
\[
\bar{x} = \frac{a+2m+b}{4} \quad \text{and} \quad SD = \frac{1}{\sqrt{12}} \left( \frac{(a-2m+b)^2}{4} + (b-a)^2 \right)
\]
if \( 15 < n \leq 25 \), then
\[
\bar{x} = \frac{a+2m+b}{4} \quad \text{and} \quad SD = \frac{R}{4}
\]
if \( 25 < n \leq 70 \), then
\[
\bar{x} = m \quad \text{and} \quad SD = \frac{R}{4}
\]
and if \( n > 70 \), then
\[
\bar{x} = m \quad \text{and} \quad SD = \frac{R}{6}
\]

If studies report mean or median age only, we will not use these data for calculation and comparison of proportions of people aged 65 years or more, but only for comparison of mean age in general.

In order to compare the two estimates of the same quantity derived from separate analyses (meta-analyses of RCTs and population-based studies, respectively), we will perform a test known as a test of interaction, based on a z-test. Our two estimates are independent (i.e., not obtained from the same individuals). If \( E_1 \) and \( E_2 \) are the estimates, and \( SE_{E_1} \) and \( SE_{E_2} \) their standard errors, \( d \) the difference between the two estimates, \( V_d \) its variance, then:

\[
V_d = SE_{E_1}^2 + SE_{E_2}^2
\]
\[
Z = \frac{d}{SE_d}
\]

The ratio \( Z \) gives a test of the null hypothesis that in the population the difference \( d \) is zero, by comparing the value of \( Z \) to the standard normal distribution. If \( CI_{95} \) is the confidence interval for the difference between the estimates, then:

\[
CI_{95} = d \pm 1.96 \times SE_d
\]
LIMITATIONS

Our study will serve as an overview of conformities or discrepancies concerning age of trial patients compared with real-world patients as described by population-based studies. However, the study exhibits some limitations. We will search MEDLINE only, which does not include every single study published in the medical field. Yet, MEDLINE is arguably the most important and encompassing database for medical publications, and we think that our retrieved sample of randomized controlled trials will give an authentic picture of the whole clinical trial landscape in rheumatoid arthritis and osteoarthritis over the last years. We will also not test the internal validity of included studies. We have come to this decision as we will only analyze participant’s baseline characteristics. Furthermore, this allows us to include a very high number of randomized controlled trials, i.e., all retrieved RCTs – considerably expanding the size of the picture. Just this year, Kilcher et al. (for the GetReal project, which consists of a consortium of pharmaceutical companies, academia, health technology assessment agencies, regulators and patient organizations) published an analysis of applicability aspects comparing participants of etanercept, rituximab, and tocilizumab trials to patients treated with these biologic agents registered in observational studies.[26] However, this study cannot be considered comprehensive enough to establish sound evidence on the applicability of trial results to the general rheumatoid arthritis or osteoarthritis population regarding age.

ETHICS AND DISSEMINATION

We do not collect any primary data. Thus, no additional formal ethics approval is necessary. Our systematic review will be the first to systematically analyze the adequacy of the representation of elderly people in osteoarthritis and rheumatoid arthritis trials. The results of this review will be published in a peer-reviewed journal according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.[19]

COMPETING INTERESTS

All authors declare that they have no competing interests.
REFERENCES


APPENDIX A – SEARCH STRINGS

Glossary

* = truncation wildcard

RCTs

PubMed (MEDLINE)

RA ("Arthritis, Rheumatoid"[Mesh] OR "Rheumatoid Arthritis"[Title/Abstract] OR (Rheumatoid Arthr*[Title/Abstract] OR “rheumatic arthritis”[Title/Abstract] OR “arthritis deformans”[Title/Abstract] OR “arthrit, rheumatoid”[Title/Abstract] OR “arthrosis deformans”[Title/Abstract] OR “chronic polyarthritis”[Title/Abstract] OR “chronic progressive polyarthritis”[Title/Abstract] OR “inflammatory arthritis”[Title/abstract] OR “rheumatic polyarthritis”[Title/Abstract] OR “Felty Syndrome”[Title/Abstract]) AND ("Randomized Controlled Trial"[Publication Type] OR (randomized[Title/Abstract] AND controlled[Title/Abstract] AND trial[Title/Abstract])) NOT (Review[Title]) NOT (Meta-analysis[Title]) NOT (Letter[Publication Type]) NOT (Case Report[Title]) AND ("2016/01/01"[PDAT] : "2017/12/31"[PDAT]) AND (Species: Humans)

OA ("Osteoarthritis"[Mesh]) OR osteoarthrosis[Title/Abstract] OR osteoarthrosis[Title/Abstract] OR “degenerative joint disease”[Title/Abstract] OR “noninflammatory arthritis”[Title/Abstract] OR osteo-arthritis[Title/Abstract] OR osteoarthrosis[Title/Abstract] OR “primary osteoarthritis”[Title/Abstract] AND ("Randomized Controlled Trial"[Publication Type] OR (randomized[Title/Abstract] AND controlled[Title/Abstract] AND trial[Title/Abstract])) NOT (Review[Title]) NOT (Meta-analysis[Title]) NOT (Letter[Publication Type]) NOT (Case Report[Title]) AND ("2016/01/01"[PDAT] : "2017/12/31"[PDAT]) AND (Species: Humans)

Population-based studies

PubMed (MEDLINE)

RA (Prevalence[Title] OR Epidemiology[Title] OR "Age Distribution"[Title] OR COPCORD[Title] OR (Prevalence[Title] OR Epidemiology[Title] OR "Cross-Sectional Studies"[Mesh] OR "Age Distribution"[Mesh]) AND ("Arthritis, Rheumatoid"[Mesh] OR "Rheumatoid Arthritis"[Title] OR Rheumatoid Arthr*[Title]) OR "rheumatic arthritis"[Title] OR "arthritis deformans"[Title] OR "arthritis, rheumatoid"[Title] OR "arthrosis deformans"[Title] OR "chronic polyarthritis"[Title] OR "chronic progressive polyarthritis"[Title] OR "inflammatory arthritis"[Title] OR "rheumatic polyarthritis"[Title] OR "Felty Syndrome"[Title]) NOT (Letter[Publication Type]) NOT (Case Report[Title]) AND ("2013/01/01"[PDat] : "2017/12/31"[PDat]) AND (Species: Humans)

APPENDIX B – DATA EXTRACTION

RCTs
Extracting author ID, Date of extraction, Study ID (e.g., Smith 2008), Author, Year, Disease, Disease type if specified (e.g., hip OA), Trial acronym, Study funding source, Region/country, Type of intervention/treatment modality, Biologic agent study?, Bioequivalence study?, Duration of study, Lower age limit, Upper age limit, Reason for age limit, Number of baseline (randomized) participants, % females, Age-stratified results?, Number of baseline participants intervention group, Number of baseline participants control group 1, Number of baseline participants control group X, Mean age of intervention group, SD of age intervention group, Mean age control group 1, SD of age control group 1, Mean age control group X, SD of age control group X, Percentage of people aged ≥65 years

Population-based studies
Extracting author ID, Date of extraction, Study ID (e.g., Smith 2008), Author, Year, Disease, Disease type if specified (e.g., hip OA), Study funding source, Region/country, Method of patient acquisition (e.g., phone; mail; registry data; regular patients), Method of condition identification, Patients affected, % females, Mean age, SD of age, Percentage of people aged ≥65 years
## References of included population-based studies


[29] Loyola-Sanchez A, Richardson J, Pelaez-Ballestas I, et al. The impact of arthritis on the physical function of a rural Maya-Yucateco community and factors associated with its


References of included randomized controlled trials


Choe JY, Prodanovic N, Niebrzydowski J, et al. A randomised, double-blind, phase III study comparing SB2, an infliximab biosimilar, to the infliximab reference product Remicade


[137] Lampe F, Fiedler F, Marques CJ, et al. Surgically modifiable factors measured by computer-navigation together with patient-specific factors predict knee society score


Risk of bias assessment

**Questions**

1. Was the sample frame appropriate to address the target population?
2. Were study participants sampled in an appropriate way?
3. Was the sample size adequate? *
4. Were the study subjects and the setting described in detail?
5. Was the data analysis conducted with sufficient coverage of the identified sample?
6. Were valid methods used for the identification of the condition?
7. Was the condition measured in a standard, reliable way for all participants?
8. Was there appropriate statistical analysis?
9. Was the response rate adequate, and if not, was the low response rate managed appropriately?

* This quality appraisal tool is designed for prevalence studies.[1, 2] Estimation of adequate sample sizes in these studies is recommended to be based on suspected prevalences of diseases in the general population. However, our aim was not to assess whether studies are adequate for the examination of the prevalence of a disease in the general population. Our aim was to evaluate whether they are adequate for the assessment of proportions (of elderly people and women) within a population consisting of patients with a specific disease (i.e., RA or OA). Therefore, question number 3 was answered with “no” for studies in the general population with generally high sample size, but a low number of finally diagnosed patients. The number of required participants was estimated based on the equation presented by Munn et al.[2] For more on this formula please consider Naing et al.[3]

**References**


Table S1. Risk of bias assessment.

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>OA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bailey</td>
<td>2014</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Birtwhistle</td>
<td>2015</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>X</td>
</tr>
<tr>
<td>Castano Carou</td>
<td>2015</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>Chiba</td>
<td>2016</td>
<td>+</td>
<td>?</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>Cimmino</td>
<td>2013</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Hernandez-Caceres</td>
<td>2015</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Iidaka</td>
<td>2016</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Loyola-Sanchez</td>
<td>2016</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>Ohsawa</td>
<td>2016</td>
<td>+</td>
<td>?</td>
<td>–</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>Tanishi</td>
<td>2014</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>Wolf</td>
<td>2014</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><strong>RA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Albrecht</td>
<td>2017</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Al-Herz</td>
<td>2017</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Austad</td>
<td>2017</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Baldassari</td>
<td>2016</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Bengtsson</td>
<td>2016</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>X</td>
</tr>
<tr>
<td>Brunier</td>
<td>2017</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Di</td>
<td>2016</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Dougados</td>
<td>2013</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>Ferraz-Amaro</td>
<td>2015</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Haye Salinas</td>
<td>2013</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Hense</td>
<td>2016</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Hifinger</td>
<td>2016</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Jafri</td>
<td>2017</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>X</td>
</tr>
<tr>
<td>Jamshidi</td>
<td>2014</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Jin</td>
<td>2017</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Khan</td>
<td>2017</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Kim</td>
<td>2016</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>Koko</td>
<td>2015</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Littlejohn</td>
<td>2015</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>Loyola-Sanchez</td>
<td>2016</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
</tr>
<tr>
<td>Monk</td>
<td>2013</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>+</td>
<td>X</td>
</tr>
<tr>
<td>Nakajima</td>
<td>2015</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

RA, rheumatoid arthritis; OA, osteoarthritis; +, yes; –, no; ?, unclear; and X, not applicable.
Continuation of Table S1. Risk of bias assessment.

<table>
<thead>
<tr>
<th>First author</th>
<th>Year</th>
<th>Question number</th>
<th>RA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Penserga</td>
<td>2015</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>Rathbun</td>
<td>2014</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ruscitti</td>
<td>2017</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Santiago-Casas</td>
<td>2013</td>
<td>?</td>
<td>+</td>
</tr>
<tr>
<td>Slimani</td>
<td>2017</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Strahl</td>
<td>2017</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Sung</td>
<td>2013</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Tantayakom</td>
<td>2016</td>
<td>–</td>
<td>+</td>
</tr>
<tr>
<td>Tedeschi</td>
<td>2017</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Urruticoechea-Arana</td>
<td>2015</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Widdifield</td>
<td>2014</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Xu</td>
<td>2014</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Yilmaz</td>
<td>2014</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Zhang</td>
<td>2015</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Zhang</td>
<td>2016</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Zlatkovic-Svenda</td>
<td>2014</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

RA, rheumatoid arthritis, +, yes, –, no, ?, unclear, and X, not applicable.
Figure S1. Meta-analysis on the proportion of elderly people in rheumatoid arthritis population-based studies. Box sizes represent random effects weights. CI, confidence interval.

Figure S2. Meta-analysis on the mean age in rheumatoid arthritis population-based studies. Box sizes represent random effects weights. MRAW, mean age (years), and CI, confidence interval.

Figure S3. Meta-analysis on the standard deviation of age in rheumatoid arthritis population-based studies. Box sizes represent random effects weights. MRAW, mean age (years), and CI, confidence interval.

Figure S4. Meta-analysis on the proportion of women in rheumatoid arthritis population-based studies. Box sizes represent random effects weights. MRAW, mean age (years), and CI, confidence interval.
Figure S7. Meta-analysis on the standard deviation of age (years) in rheumatoid arthritis randomized controlled trials. Box sizes represent random effects weights. MRAW, standard deviation of age (years), and CI, confidence interval.

Figure S5. Meta-analysis on the proportion of elderly people in rheumatoid arthritis randomized controlled trials. Box sizes represent random effects weights. CI, confidence interval.

Figure S6. Meta-analysis on the mean age in rheumatoid arthritis randomized controlled trials. Box sizes represent random effects weights. MRAW, mean age (years), and CI, confidence interval.

Figure S8. Meta-analysis on the proportion of women in rheumatoid arthritis randomized controlled trials. Box sizes represent random effects weights. CI, confidence interval.
Figure S9. Meta-analysis on the proportion of elderly people in osteoarthritis population-based studies. Box sizes represent random effects weights. CI, confidence interval.

Figure S10. Meta-analysis on the mean age in osteoarthritis population-based studies. Box sizes represent random effects weights. MRAW, mean age (years), and CI, confidence interval.

Figure S11. Meta-analysis on the standard deviation of age (years) in osteoarthritis population-based studies. Box sizes represent random effects weights. MRAW, standard deviation of age (years), and CI, confidence interval.

Figure S12. Meta-analysis on the proportion of women in osteoarthritis population-based studies. Box sizes represent random effects weights. CI, confidence interval.
**Figure S13.** Meta-analysis on the proportion of elderly people in osteoarthritis randomized controlled trials. Box sizes represent random effects weights. CI, confidence interval.

**Figure S14.** Meta-analysis on the mean age in osteoarthritis randomized controlled trials. Box sizes represent random effects weights. MRAW, mean age (years), and CI, confidence interval.
Figure S15. Meta-analysis on the standard deviation of age (years) in osteoarthritis randomized controlled trials. Box sizes represent random effects weights. MRAW, standard deviation of age (years), and CI, confidence interval.

Figure S16. Meta-analysis on the proportion of women in osteoarthritis randomized controlled trials. Box sizes represent random effects weights. CI, confidence interval.
Studies on underrepresentation of the elderly in clinical trials


# PRISMA Checklist

<table>
<thead>
<tr>
<th>Item No</th>
<th>Recommendation</th>
<th>Reported on Page No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Identify the report as a systematic review, meta-analysis, or both.</td>
<td>1</td>
</tr>
<tr>
<td><strong>Abstract</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.</td>
<td>2</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Describe the rationale for the review in the context of what is already known.</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).</td>
<td>Protocol, 3</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.</td>
<td>Protocol, 3</td>
</tr>
<tr>
<td>6</td>
<td>Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>7</td>
<td>Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>8</td>
<td>Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.</td>
<td>Protocol</td>
</tr>
<tr>
<td>9</td>
<td>State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>10</td>
<td>Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>11</td>
<td>List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>12</td>
<td>Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.</td>
<td>Protocol, 4, 5</td>
</tr>
<tr>
<td>13</td>
<td>State the principal summary measures (e.g., risk ratio, difference in means).</td>
<td>Protocol, 5</td>
</tr>
<tr>
<td>14</td>
<td>Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I²) for each meta-analysis.</td>
<td>Protocol, 5</td>
</tr>
<tr>
<td>15</td>
<td>Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).</td>
<td>5</td>
</tr>
<tr>
<td>16</td>
<td>Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.</td>
<td>5</td>
</tr>
<tr>
<td><strong>Results</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.</td>
<td>5, Figure 1</td>
</tr>
<tr>
<td>18</td>
<td>For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.</td>
<td>6, Table 1</td>
</tr>
<tr>
<td>19</td>
<td>Present data on risk of bias of each study and, if available, any outcome-level assessment (see Item 12).</td>
<td>6, Supplement</td>
</tr>
<tr>
<td>20</td>
<td>For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group and (b) effect estimates and confidence intervals, ideally with a forest plot.</td>
<td>Figure 2, Supplement</td>
</tr>
<tr>
<td>21</td>
<td>Present results of each meta-analysis done, including confidence intervals and measures of consistency.</td>
<td>Figure 2, Supplement</td>
</tr>
<tr>
<td>22</td>
<td>Present results of any assessment of risk of bias across studies (see Item 15).</td>
<td>see Item 15</td>
</tr>
<tr>
<td>23</td>
<td>Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).</td>
<td>6, Table 2</td>
</tr>
</tbody>
</table>
### MOOSE Checklist

<table>
<thead>
<tr>
<th>Item No</th>
<th>Recommendation</th>
<th>Reported on Page No</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Discussion</strong></td>
<td>Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., health care providers, users, and policy makers).</td>
<td>6 - 9</td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>Discuss limitations at study and outcome level (e.g., risk of bias), and at review level (e.g., incomplete retrieval of identified research, reporting bias).</td>
<td>8, 9</td>
</tr>
<tr>
<td>26</td>
<td>Provide a general interpretation of the results in the context of other evidence, and implications for future research.</td>
<td>6 - 9</td>
</tr>
<tr>
<td><strong>Funding</strong></td>
<td>Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.</td>
<td>3, 10</td>
</tr>
<tr>
<td>27</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Item No | Recommendation | Reported on Page No |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reporting of background should include</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Problem definition</td>
<td>2 - 4, 8 - 10</td>
</tr>
<tr>
<td>2</td>
<td>Hypothesis statement</td>
<td>2 - 3</td>
</tr>
<tr>
<td>3</td>
<td>Description of study outcome(s)</td>
<td>2, 3, 5</td>
</tr>
<tr>
<td>4</td>
<td>Type of exposure or intervention used</td>
<td>2 - 4</td>
</tr>
<tr>
<td>5</td>
<td>Type of study designs used</td>
<td>2 - 4</td>
</tr>
<tr>
<td>6</td>
<td>Study population</td>
<td>2 – 4</td>
</tr>
<tr>
<td><strong>Reporting of search strategy should include</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Qualifications of searchers (e.g., librarians and investigators)</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Search strategy, including time period included in the synthesis and key words</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>9</td>
<td>Effort to include all available studies, including contact with authors</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>Databases and registries searched</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>11</td>
<td>Search software used, name and version, including special features used (e.g., explosion)</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>12</td>
<td>Use of hand searching (e.g., reference lists of obtained articles)</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>13</td>
<td>List of citations located and those excluded, including justification</td>
<td>Supplement, Figure 1</td>
</tr>
<tr>
<td>14</td>
<td>Method of addressing articles published in languages other than English</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>15</td>
<td>Method of handling abstracts and unpublished studies</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>Description of any contact with authors</td>
<td>4</td>
</tr>
<tr>
<td><strong>Reporting of methods should include</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Description of relevance or appropriateness of studies assembled for assessing the hypothesis to be tested</td>
<td>3, 4</td>
</tr>
<tr>
<td>18</td>
<td>Rationale for the selection and coding of data (e.g., sound clinical principles or convenience)</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>19</td>
<td>Documentation of how data were classified and coded (e.g., multiple raters, blinding and interrater reliability)</td>
<td>Protocol, 4</td>
</tr>
<tr>
<td>20</td>
<td>Assessment of confounding (e.g., comparability of cases and controls in studies where appropriate)</td>
<td>Table 1</td>
</tr>
<tr>
<td>Item No</td>
<td>Recommendation</td>
<td>Reported on Page No</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>21</td>
<td>Assessment of study quality, including blinding of quality assessors, stratification or regression on possible predictors of study results</td>
<td>Protocol, 4, 5</td>
</tr>
<tr>
<td>22</td>
<td>Assessment of heterogeneity</td>
<td>Protocol, 5</td>
</tr>
<tr>
<td>23</td>
<td>Description of statistical methods (eg, complete description of fixed or random effects models, justification of whether the chosen models account for predictors of study results, dose-response models, or cumulative meta-analysis) in sufficient detail to be replicated</td>
<td>Protocol, 5</td>
</tr>
<tr>
<td>24</td>
<td>Provision of appropriate tables and graphics</td>
<td>see below</td>
</tr>
<tr>
<td>25</td>
<td>Graphic summarizing individual study estimates and overall estimate</td>
<td>Figure 2, Supplement</td>
</tr>
<tr>
<td>26</td>
<td>Table giving descriptive information for each study included</td>
<td>Table 1</td>
</tr>
<tr>
<td>27</td>
<td>Results of sensitivity testing (eg, subgroup analysis)</td>
<td>6, Table 2</td>
</tr>
<tr>
<td>28</td>
<td>Indication of statistical uncertainty of findings</td>
<td>6, Figure 2</td>
</tr>
<tr>
<td>29</td>
<td>Quantitative assessment of bias (eg, publication bias)</td>
<td>-</td>
</tr>
<tr>
<td>30</td>
<td>Justification for exclusion (eg, exclusion of non-English language citations)</td>
<td>4, Figure 1</td>
</tr>
<tr>
<td>31</td>
<td>Assessment of quality of included studies</td>
<td>6, Supplement</td>
</tr>
<tr>
<td>32</td>
<td>Consideration of alternative explanations for observed results</td>
<td>6 - 9</td>
</tr>
<tr>
<td>33</td>
<td>Generalization of the conclusions (ie, appropriate for the data presented and within the domain of the literature review)</td>
<td>6 - 9</td>
</tr>
<tr>
<td>34</td>
<td>Guidelines for future research</td>
<td>6 - 9</td>
</tr>
<tr>
<td>35</td>
<td>Disclosure of funding source</td>
<td>3, 10</td>
</tr>
</tbody>
</table>