


Outcome Domains, Outcome Measures, and Characteristics of Randomized Controlled Trials Testing Nonsurgical Interventions for Osteoarthritis

Miso Krsticevic, Svjetlana Dosenovic, Daiana Anne-Marie Dimcea, Dominika Jedrzejewska, Ana Catarina Marques Lameirão, Eliana Santos Almeida, Antonia Jelcic Kadic, Milka Jeric Kegalj, Krste Boric, and Livia Puljak 

ABSTRACT. Objective. Core outcome set (COS) is the minimum set of outcome domains that should be measured and reported in clinical trials. We analyzed outcome domains, prevalence of use of COS published by Outcome Measures in Rheumatology (OMERACT) initiative, outcome measures for outcome domains recommended by OMERACT COS, duration and size of randomized controlled trials (RCT) testing nonsurgical interventions for osteoarthritis (OA).

Methods. We searched PubMed and analyzed RCT about nonsurgical interventions for OA published from June 2012 to June 2017. We extracted data about trial type, use of OMERACT COS, efficacy outcome domains, safety outcome domains, outcome measures used for COS assessment, duration, and sample size.

Results. Among 334 analyzed trials, complete OMERACT-recommended COS was used by 14% of trials. Higher median prevalence of using OMERACT COS was found in trials explicitly described as phase III, and trials of pharmacological interventions with followup ≥ 1 year, but both with wide range of COS usage. Trialists used numerous different outcome measures for analyzing core outcome domains: 50 different outcome measures for pain, 74 for physical function, 9 for patient's global assessment, and 5 for imaging.

Conclusion. Suboptimal use of recommended COS and heterogeneity of outcome measures is reducing quality and comparability of OA trials and hinders conclusions about efficacy and comparative efficacy of nonsurgical interventions. Interventions for improving study design of trials in this field would be beneficial. (First Release June 15 2019; J Rheumatol 2020;47:126–31; doi:10.3899/jrheum.180985)

Key Indexing Terms:

OSTEOARTHRITIS RANDOMIZED CONTROLLED TRIAL OUTCOME TREATMENT

From the Department of Orthopedics, and the Department of Anesthesiology and Intensive Care Medicine, and the Department of Pediatrics, University Hospital Split, Split; Department of Dermatovenerology, General Hospital Zadar, Zadar; Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Zagreb, Croatia; Faculty of Medicine, Transilvania University of Brasov, Brasov, Romania; Medical School, Medical University of Lodz, Lodz, Poland; Medical School, Abel Salazar Institute of Biomedical Sciences, Porto; Faculty of Medicine, University of Coimbra, Coimbra, Portugal.

M. Krsticevic, MD, Department of Orthopedics, University Hospital Split; S. Dosenovic, MD, Department of Anesthesiology and Intensive Care Medicine, University Hospital Split; D.A. Dimcea, Medical Student, Faculty of Medicine, Transilvania University of Brasov; D. Jedrzejewska, Medical Student, Medical School, Medical University of Lodz; A.C. Marques Lameirão, Medical Student, Medical School, Abel Salazar Institute of Biomedical Sciences; E.S. Almeida, Medical Student, Faculty of Medicine, University of Coimbra; A. Jelcic Kadic, MD, PhD, Department of Pediatrics, University Hospital Split; M. Jeric Kegalj, MD, PhD, Department of Dermatovenerology, General Hospital Zadar; K. Boric, MD, PhD, Department of Orthopedics, University Hospital Split; L. Puljak, MD, PhD, Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia.

Address correspondence to Prof. L. Puljak, Center for Evidence-Based Medicine and Health Care, Catholic University of Croatia, Ilica 242, 10000 Zagreb, Croatia. E-mail: livia.puljak@gmail.com, livia.puljak@unicath.hr
Accepted for publication February 13, 2019.

Core outcome set (COS) is the minimum set of outcome domains that should be measured and reported in clinical trials with patients having a specific condition. By using COS, trials can be easily compared, and their results can be included in metaanalyses as appropriate¹.

In 1997, the Outcome Measures in Rheumatology (OMERACT) initiative published recommendations for the COS of phase III clinical trials in knee, hip, and hand osteoarthritis (OA). The COS for OA included pain, physical function, patient's global assessment (PtGA), and joint imaging for studies lasting 1 year or longer². Previous analysis has shown that even when COS is defined and used, different outcome measures can be used, which contributes to heterogeneity of trials that are not comparable and whose results cannot be combined³. Analysis of Summary of Findings tables from Cochrane systematic reviews about interventions for chronic painful musculoskeletal conditions showed that in the 57 analyzed tables, 56 included pain as the outcome domain. Those 56 tables reported pain intensity as a measure of pain, which was assessed with as many as 20

different instruments. The visual analog scale (VAS) was the most frequently used instrument (45%). Pain was measured both as continuous and dichotomous outcome. Pain interference and pain frequency were also included in these tables. Some of the tables did not specify which measure of pain was used³. This analysis highlighted the need for defining recommendations for specifying outcome measures that are supposed to be used in conjunction with the COS^{3,4}.

In 2015, the OMERACT Hand Osteoarthritis Working Group published a set of core domains. They included 8 main domains (with 6 subdomains): pain, physical function, PtGA, joint activity (tender joints and soft swollen joints), hand strength, health-related quality of life, structural damage (deformity, radiographic, aesthetic, and bony damage), and hand mobility⁵.

In addition to the COS, it is also important for clinical trials to have adequate duration and sample size. In a recent Cochrane systematic review on OA, it has been suggested that only trials lasting 4 weeks with at least 50 participants per group should be considered for inclusion in high-quality systematic reviews⁶.

The aim of our study was to analyze outcome domains, prevalence of use of OMERACT COS for knee, hip, and hand OA recommended in 1997²; outcome measures for outcome domains recommended by OMERACT COS; and duration and size of randomized controlled trials (RCT) testing nonsurgical interventions for OA. We intended to focus on trials published between 2012 and 2017 (i.e., more recently published trials), to record current trialists' use of recommended outcomes, and to ensure that there was adequate lead-in time to ensure authors of those trials could have implemented the recommendations.

MATERIALS AND METHODS

Study design. A primary retrospective cross-sectional methodological study was conducted.

Database searching. We searched PubMed to find RCT of interventions for OA published from June 2012 to June 2017. The following search strategy was used: ((osteoarthritis) AND "randomized controlled trial"[Publication Type]) AND ("2012/06/01"[Date - Publication]: "2017/06/01"[Date - Publication]), as well as filter for humans.

Types of studies included. We analyzed RCT of any type of nonsurgical intervention for the treatment of OA. We included only studies published in the English language.

Study screening. Two authors (MK and AJK) screened titles and abstracts of retrieved records and the third author (SD) verified titles and abstracts that were retained for inclusion. In case it was not clear from the title and abstract that a study was eligible, the full text was retrieved and screened. After completing the list of included studies, we obtained full texts and extracted data.

Data extraction. Two pairs of authors (DAMD and DJ, plus ACML and ESA) independently extracted data. Two authors (MJ and KB) checked for discrepancies among them. The following data were extracted: bibliographic details (first author, yr), journal name, type of intervention, joint affected by OA, phase of the clinical trial, whether they mention that they used COS, efficacy outcome domains, safety outcome domains, outcome measures for OMERACT COS, final study followup and sample size (total no. patients

randomized and total no. patients who completed the study). Extracted data were analyzed and categorized by MK and SD.

We analyzed prevalence of use of outcome domains recommended by OMERACT COS, including pain, physical function, PtGA, and joint imaging for studies lasting 1 year or longer², and how many non-COS outcome domains were used by the trials. Further, we analyzed which outcome measures were used by the trials that used OMERACT COS.

Data synthesis. We performed descriptive data analysis and presented data as frequencies and percentages. We used Fisher's exact test to calculate differences in proportions. We conducted analyses using MedCalc statistical software, version 15.2.1 (MedCalc Software bvba). Statistical significance was set at $p < 0.05$.

RESULTS

Our search of PubMed retrieved 1243 bibliographic records. After excluding studies that were not RCT published in English, studies that were not focused on participants with OA, and trials that analyzed surgical interventions, we included 334 RCT about nonsurgical interventions for OA that were indexed in PubMed from June 2012 to June 2017.

Study characteristics. We placed interventions that were tested in analyzed trials into 8 groups; the majority of trials tested complementary and alternative therapies ($n = 97$; 29%), followed by various types of physical therapy ($n = 96$; 29%), pharmacological therapies ($n = 56$; 17%), assistive devices ($n = 15$; 4.5%), behavioral interventions ($n = 11$; 3.3%), psychological interventions ($n = 5$; 1.5%), and interventions involving genetic engineering ($n = 3$; 0.9%). In 51 trials (15%), various combinations of alternative, physical, pharmacological, and behavioral interventions were tested.

Only 16 (5%) manuscripts reported phase of a clinical trial: one as phase I, one as phase I/II, three as phase II, seven as phase III, and four as phase IV. Explicit mention of the COS was found in 6 trials (1.8%); in 5 trials the authors explicitly indicated that they used OMERACT COS; 1 mentioned Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials (IMMPACT) COS. None of the trials have reported a reason for not using COS. More detailed information about included studies are presented in Supplementary Table 1 (available with the online version of this article).

Outcome domains used in OA trials. The median number of all outcome domains in the analyzed OA trials was 4 (range 1–10). The median number of the OMERACT outcome domains was 2 (range 0–4), and this result was the same in trials with followup longer than 1 year ($n = 70$) and in shorter trials ($n = 264$; Table 1). This distinction between trials with followup under or over 1 year is relevant because one of the 4 OMERACT-recommended outcome domains is "joint imaging for studies lasting 1 year or longer," which was not applicable in all analyzed RCT if they had shorter duration.

The most commonly used OMERACT outcome domain was pain (97%), followed by physical functioning (84%) and PtGA (17%) among all analyzed trials. Among 70 trials lasting 1 year or longer, one-third used OMERACT outcome domain "joint imaging," which is recommended for trials of

Table 1. Frequency of using Outcome Measures in Rheumatology (OMERACT) core outcome set, and other outcome domains, in analyzed trials.

Group of Analyzed Trials	OMERACT Core Outcome Set, Median (range)	Other Outcome Domains, Median (range)	All Outcome Domains, Median (range)
All trials, n = 334	2 (0–4)	2 (0–8)	4 (1–10)
Trials with followup ≥ 1 yr, n = 70	2 (0–4)	2 (0–6)	5 (1–8)
Trials with followup shorter than 1 yr, n = 264	2 (0–3)	2 (0–8)	4 (1–10)
Phase III trials; all < 1 yr, n = 7	3 (2–3)	2 (1–5)	5 (1–6)
Trials of pharmacological interventions, n = 56	2 (1–4)	3 (0–8)	5 (2–10)
Trials of pharmacological interventions with followup < 1 yr, n = 47	2 (1–3)	3 (0–8)	6 (2–10)
Trials of pharmacological interventions with followup ≥ 1 yr, n = 9	3 (2–4)	1 (1–2)	4 (3–5)
Trials of other interventions; nonpharmacological or combinations, n = 278	2 (0–4)	2 (0–7)	4 (1–9)
Trials of other interventions with followup < 1 yr, n = 217	2 (0–3)	2 (0–7)	4 (1–9)
Trials of other interventions with followup ≥ 1 yr, n = 61	2 (0–4)	2 (0–6)	5 (1–8)

such duration (Table 2). Among trials shorter than 1 year, 43/264 (16%) used all 3 outcome domains recommended by OMERACT for such trials; among trials lasting 1 year or longer, only 4/70 (5.7%) used all 4 recommended outcome domains. The difference in proportions between these 2 groups was significant ($p = 0.0167$). Therefore, overall 47/334 trials (14%) used OMERACT-recommended outcome domains that were relevant for their duration.

Subgroup analysis was conducted for the 7 trials described as phase III; all these trials had followup shorter than 1 year.

Table 2. Type and frequency of outcome domains in randomized controlled trials about conservative interventions for osteoarthritis.

Outcome Domains	N (%)
OMERACT core outcome domains	
Pain	323 (97)
Physical functioning	282 (84)
Patient's global assessment	58 (17)
Joint imaging for studies lasting 1 yr or longer	21 (30)*
Non-OMERACT outcome domains	
Safety	216 (65)
Arthritis symptoms	168 (50)
Quality of life	82 (25)
Medication (analgesic) use	73 (22)
Physician's global assessment or surrogate ratings of global improvement	68 (20)
Biological markers	50 (15)
Emotional functioning	24 (7.2)
Body mass index	22 (6.6)
Clinical improvement	16 (4.8)
Clinical effectiveness	16 (4.8)
Patient satisfaction	15 (4.5)
Adherence to treatment	15 (4.5)
Participation (work status, role functioning, sporting ability, lifestyle, friend and family support)	9 (2.7)
Progress of osteoarthritis	7 (2.1)
Imaging other than joint imaging	6 (1.8)
Other	22 (6.6)

* Percent calculated from 70 trials with followup longer than 1 year, for which this outcome domain is relevant. OMERACT: Outcome Measures in Rheumatology.

In these 7 trials, median number of OMERACT COS domains was higher than in the total sample, (i.e., 3, range 2–3), and median number of all outcomes that were used was 5 (range 1–6; Table 1).

As shown in Table 1, we also compared usage of core outcome domains in trials of pharmacological interventions ($n = 56$) versus all other trials ($n = 278$). Median use of OMERACT COS in trials of pharmacological interventions was 2, just as in the overall sample. Higher use of OMERACT COS domains was found only in trials of pharmacological interventions with duration ≥ 1 year (median 3, range 2–4; Table 1).

Separate analysis of 15 trials that included joints other than knee, hip, and hand showed that those trials also used a median of 2 OMERACT-recommended core outcome domains, and median of 4 total outcome domains. Additional subgroup analysis was conducted on 5 trials that explicitly indicated that they used OMERACT COS; 2 of those trials had 1-year duration and both used 3 recommended outcome domains instead of 4, while only one of the remaining 3 that were shorter than 1 year used 3 outcome domains relevant for trials of that duration; the other two used 2 recommended outcome domains.

The most commonly used non-OMERACT outcome domains were related to safety, arthritis symptoms, quality of life, and medication (analgesic use; Table 2). List of outcome domains that were most commonly used in analyzed trials is shown in Table 2. Beyond the listed outcome domains, in 22 trials authors used other outcomes that were found in fewer than 5 trials each; those outcomes were temperature, dietary intake assessment, perceived helpfulness, treatment progression, therapeutic efficacy, postoperative discomfort, gait analysis, vital signs, number and type of comorbidities, and device usage.

Outcome measures used for OMERACT-recommended outcome domains. Analyzed trials used numerous outcome measures for OMERACT-recommended outcome domains: they used 50 different outcome measures for pain, 74 for physical function, 9 for PtGA, and 5 for imaging. The

majority of trials used VAS and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) for measuring pain; WOMAC and muscle strength for measuring physical function; Likert scales with various numbers of points and VAS scales for measuring PtGA; and joint space width/narrowing and cartilage volume/thickness for joint imaging. Five most commonly used outcome measures for each OMERACT-recommended outcome domain are shown in Table 3, while all outcome measures that were identified are presented in Supplementary Table 2 (available with the online version of this article).

DISCUSSION

The main finding of our study about nonsurgical interventions for OA is that the use of OMERACT-recommended COS in those trials was suboptimal and that trialists use a highly heterogeneous set of outcome measures. Higher median prevalence of using OMERACT COS was found in trials explicitly described as phase III, and trials of pharmacological interventions with followup ≥ 1 year, but both have a wide range of COS usage. These data indicate serious deficiencies in the evidence base about treatments tested for OA.

While pain was analyzed in almost all the analyzed trials, 84% analyzed physical function, only one-fifth of the trials analyzed PtGA and only one-third of trials that lasted a year or longer utilized joint imaging. These findings are worrisome because OMERACT recommended this COS in

1997², so trialists publishing RCT from 2012 to 2017 have had ample time to adjust their study designs to the recommended COS, and to plan them and report them accordingly.

Trialists used a number of different outcome domains and outcome measures in the analyzed studies. Existence of a COS does not mean that trialists should use only outcome domains recommended within a COS; instead, trialists should use COS and also other outcome domains that they find relevant. Sometimes it may not be pertinent to use COS, but in those cases, trialists should explicitly explain why a COS was not used¹. In our cohort of analyzed trials, not a single manuscript mentioned a reason for not using the relevant COS. Even among those few trials that explicitly mentioned that they used OMERACT COS, only 2 out of 5 actually used all the recommended core outcomes relevant for their trial duration. It has already been shown that citing a COS does not necessarily mean that the authors have actually used it⁷.

Lack of adherence to recommended COS reduces comparability of trials and the possibility of meaningful meta-analysis. However, simply following the COS is not sufficient because even usage of the same outcome domains may not yield comparable data if trialists do not use the same outcome measures^{3,4}. For that purpose, we also analyzed which outcome measures were analyzed within the 4 OMERACT outcome domains and we found that the authors used a myriad of different outcome measures, which can make comparisons of trials very challenging.

Table 3. Type and frequency of outcome measures most commonly used for core outcome set recommended by Outcome Measures in Rheumatology (OMERACT) initiative in randomized controlled trials about conservative interventions for osteoarthritis.

Outcome Domains	Outcome Measures	N (%)
Pain, n = 323	VAS	187 (58)
	WOMAC pain	181 (56)
	11-point NRS	23 (7.1)
	KOOS questionnaire	23 (7.1)
	Lequesne Index score	19 (5.9)
Physical functioning, n = 282	WOMAC function	171 (61)
	Muscle strength	42 (15)
	Range of motion (active/passive)	37 (13)
	TUG test	29 (10)
	Lequesne Index	26 (9.3)
PtGA, n = 58	Likert scales	16 (28)
	VAS	11 (19)
	5-point scale	10 (17)
	Patient Global Impression of Change	8 (14)
	PtGA score	5 (8.6)
Joint imaging for studies lasting 1 yr or longer, n = 21	Joint space width/narrowing	16 (76)
	Cartilage volume/thickness	7 (33)
	Bone marrow lesion	5 (24)
	Effusion synovitis score	2 (9.5)
	Synovial thickness	1 (4.8)

VAS: visual analog scale; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index; NRS: numerical rating scale; KOOS: Knee Injury and Osteoarthritis Outcome Score; TUG: timed up-and-go; PtGA: patient's global assessment.

Improved usage of COS has been described in other research fields. Kirkham, *et al* showed that outcome selection in RCT on rheumatoid arthritis (RA) has improved over the last 50 years⁸. Usage of full set of RA COS was higher after its publication. They also showed that use of RA COS was lower in trials shorter than 52 weeks⁸. Bautista-Molano, *et al* analyzed uptake of the Assessment of Spondyloarthritis international Society (ASAS)/OMERACT core set and response criteria for ankylosing spondylitis in RCT⁹. They also found improvement in usage of the analyzed core set after the publication of ASAS/OMERACT⁹. Both these studies indicate that despite observed improvements, there is still room for improvement regarding uptake of relevant COS. Interventions for increasing uptake of recommended outcome domains and measures would be welcome, for example prompting authors to use COS and recommended outcome measures at the time of trial registration¹⁰.

Safety outcomes were reported in 65% of the trials, which is difficult to comprehend. Monitoring of patient safety during clinical trials is a critical component of intervention testing. The aim of safety monitoring in trials is to identify, evaluate, minimize, and appropriately manage risks. For testing of pharmaceutical interventions, regulators are putting increased focus on risk management plans, risk evaluation, and minimization strategies¹¹.

A limitation of our study is reliance on analyzing outcome domains and measures that were reported in manuscripts published in peer-reviewed literature. It is possible that some trialists originally planned different outcome domains and outcome measures, but eventually failed to report them. In this regard, one could analyze registered protocols for these trials, and compare outcome domains and outcome measures that were registered with those that were published in the manuscript. A potential limitation of this approach is our finding that only 57% of analyzed trials had reported in the manuscript that the study was registered. Another limitation could be that we included only trials published in English, focused on a recent 5-year period, and we did not assess the COS for hand OA. We chose this period because we wanted to analyze a cohort of more recently published trials. Going back in time and analyzing trials that were published in the past 20 years since publication of the COS might provide a picture of the changing uptake of the COS over time, but this was not the aim of our study. Analysis of the COS for hand OA was not included because the COS recommendations from 1997 are not in use now and the new recommendations were published in 2015. Because we included trials published until July 2017, we considered that a period of 2 years is not enough time to design, conduct, and publish trial results. It has been shown that, on average, it takes 2 years from completion of a trial to publication of trial results¹². Taking into account additional time needed to design and conduct a trial, it becomes apparent that it may take a very long time

from publication of a COS to the realistic reflection of its uptake in the published literature.

We analyzed all trials described as RCT regardless of the phase because very few trials described the phase of the clinical trials. Only 16 out of 334 analyzed trials reported trial phase, and only 7 were described as a phase III trial. Therefore, it could be argued that the COS we used as a reference standard may not be applicable in all the trials in our cohort, which may be a valid point. Our study can also be taken as analysis of outcomes and relevant outcome measures in published trials, and not only in phase III trials. If one would aim to analyze only RCT specified as phase III trials, that analysis would be limited to a negligibly small number of trials.

Further, we searched PubMed to retrieve eligible trials, and it is possible that some trials were missed; we did not hand-search all journals that might have published such trials.

Suboptimal use of recommended core outcome domains and usage of a wide variety of outcome measures is reducing quality and comparability of OA trials. These practices hinder conclusions about efficacy and comparative efficacy of interventions. Suggestions for improving the design of trials in this field would be beneficial.

ONLINE SUPPLEMENT

Supplementary material accompanies the online version of this article.

REFERENCES

1. COMET. Core outcome measures in effectiveness trials initiative. [Internet. Accessed April 16, 2019.] Available from: www.comet-initiative.org
2. Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, et al. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. *J Rheumatol* 1997;24:799-802.
3. Maxwell LJ, Wells GA, Simon LS, Conaghan PG, Grosskleg S, Scrivens K, et al. Current state of reporting pain outcomes in Cochrane Reviews of chronic musculoskeletal pain conditions and considerations for an OMERACT research agenda. *J Rheumatol* 2015;42:1934-42.
4. Dobson F, Hinman RS, Roos EM, Abbott JH, Stratford P, Davis AM, et al. OARSI recommended performance-based tests to assess physical function in people diagnosed with hip or knee osteoarthritis. *Osteoarthritis Cartilage* 2013;21:1042-52.
5. Kloppenburg M, Boyesen P, Visser AW, Haugen IK, Boers M, Boonen A, et al. Report from the OMERACT Hand Osteoarthritis Working Group: set of core domains and preliminary set of instruments for use in clinical trials and observational studies. *J Rheumatol* 2015;42:2190-7.
6. Puljak L, Marin A, Vrdoljak D, Markotic F, Utrobicic A, Tugwell P. Celecoxib for osteoarthritis. *Cochrane Database Syst Rev* 2017;5:CD009865.
7. Barnes KL, Kirkham JJ, Clarke M, Williamson PR. Citation analysis did not provide a reliable assessment of core outcome set uptake. *J Clin Epidemiol* 2017;86:153-9.
8. Kirkham JJ, Boers M, Tugwell P, Clarke M, Williamson PR. Outcome measures in rheumatoid arthritis randomised trials over the last 50 years. *Trials* 2013;14:324.
9. Bautista-Molano W, Navarro-Compan V, Landewe RB, Boers M,

- Kirkham JJ, van der Heijde D. How well are the ASAS/OMERACT Core Outcome Sets for Ankylosing Spondylitis implemented in randomized clinical trials? A systematic literature review. *Clin Rheumatol* 2014;33:1313-22.
10. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;13:132.
11. Yao B, Zhu L, Jiang Q, Xia HA. Safety monitoring in clinical trials. *Pharmaceutics* 2013;5:94-106.
12. Ross JS, Mocanu M, Lampropulos JF, Tse T, Krumholz HM. Time to publication among completed clinical trials. *JAMA Intern Med* 2013;173:825-8.