Osteoarthritis and Cartilage 27 (2019) 1253–1256

Osteoarthritis and Cartilage

Editorial

When enough is enough - How to determine when the evidence for the effectiveness of a treatment is sufficient?

The over-replication of trials is unethical, expensive and potentially harmful to patients

The concept of replication in research is widely accepted in most disciplines[1]; however, there is no consensus about when the evidence is sufficient to confidently state that a treatment works. One good example is a study by Fergusson *et al.*, in 2005 who published a cumulative meta-analysis of the efficacy of Aprotinin to stop bleeding in cardiac surgery[2]. They found that the estimates for its effectiveness were consistent after the first 12 studies; nevertheless, an additional 52 similar randomised controlled trials (RCTs) were carried out afterwards including more than 5,000 patients. The consequence of over-replicating experiments is not only wasteful economically, but also unethical, and potentially harmful by not exposing patients in the comparator groups to the best treatment available. In this editorial, using the example of RCTs comparing exercise to a control treatment in people with knee osteoarthritis (OA), we discuss the issue of when to stop performing unnecessary research.

The example of exercise trials in knee osteoarthritis

Exercise is now considered medicine for people with knee OA[3] and universally recommended in treatment guidelines [4].Since the first RCT of exercise for knee OA was performed in 1989, the number of trials has reached 82, with a peak of 18 trials published in the period between 2008 and 2010 (Fig. 1). The authors of a cumulative meta-analysis published in 2019 by Verhagen *et al.* in this issue of Osteoarthritis and Cartilage[5] argue that exercise therapy RCTs aimed at reducing knee pain in participants with knee OA have been over-replicated. They updated the two Cochrane systematic reviews assessing the effect of land-based[6] and aquatic-based[7] therapeutic exercise for knee pain and selected for analysis the studies comparing exercise to minimal or no treatments. They concluded that sufficient evidence for a beneficial effect of exercise on knee pain was already present in 1998 after the publication of five trials with consistent results. Additionally, they reported that the studies performed since 2010 (from trial number 23 published in 2010 and onwards) did not report changes to these effect estimates, and by performing an extended funnel plot analysis, they showed that future studies will not change our confidence in these results. These findings are in line with the Cochrane systematic reviews which, by using the GRADE approach and having broader inclusion criteria, suggested that future exercise trials are likely not to change the estimates of the meta-analyses[6,7]. Similarly, Uthman *et al.* performed a trial sequential analysis, including RCTs of adults with knee or hip pain published up to March 2012 and found that by 2002, sufficient evidence had already been accumulated to demonstrate a beneficial effect of exercise therapy (versus no exercise) are unlikely to change these estimates[8]. Of note is that these studies used different statistical methods and inclusion/exclusion criteria which may have caused the discrepancy in the year suggested to be the cut-off for 'enough evidence'. Nevertheless, the overall conclusions that future trials were unlikely to change the estimates for the effect of exercise therapy on pain and functional limitations are consistent (Table I).

More than statistics are required to determine when 'enough is enough'

Cumulative forest plots, extended forest plots, prediction intervals and trial sequential analyses are examples of statistical methods used to assess whether a meta-analysis is providing a conclusive answer [9]. On the contrary, when researchers would like to determine if more studies are needed, then the value of a new study should be based upon the end-user's perspective (e.g. is the new study asking a relevant question?) and the results from a high-quality systematic review (including a grading process) of earlier similar studies (is the new study necessary?). Furthermore, the implication for practice should be based on several elements including the confidence interval of the mean estimate and not the mean estimate itself, the grading of the evidence and the specific context of the intervention[9,10].

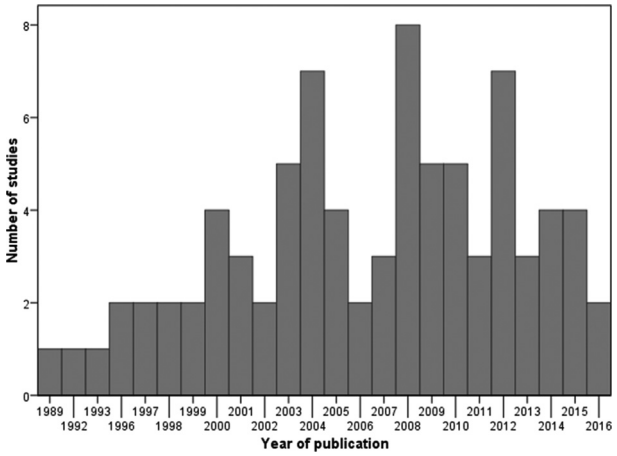

Fig. 1. The 82 randomised controlled trials (RCTs) on the effect of exercise for people with knee OA published per year from 1989 to 31 August 2016. *X*-axis = year of publication; *y*-axis number of studies.

https://doi.org/10.1016/j.joca.2019.06.003

Table I
Overview of the systematic reviews investigating when 'enough is enough' for the effect of exercise on knee osteoarthritis

	Aim	Studies included	Inclusion criteria (PICO)	Statistical method	Results	Conclusions
Uthman et al. 2013	To determine whether there is sufficient evidence to conclude that exercise interventions are more effective than no exercise control and to compare the effectiveness of different exercise interventions in relieving pain and improving function in patients with lower limb osteoarthritis.	RCTs published up to March 2012	*P* = Adults with an established clinical or radiographic diagnosis of knee or hip osteoarthritis as defined by the ACR criteria. I = Any therapeutic exercise intervention (land or water based). *C* = Other forms of exercise or no exercise control group. *O* = Self-reported pain and function	Trial sequential analysis and network meta-analysis	60 trials (44 knee, two hip, 14 mixed). For pain relief, strengthening, flexibility plus strengthening, flexibility plus strengthening plus aerobic, aquatic strengthening, and aquatic strengthening plus flexibility, exercises were significantly more effective than no exercise control. A combined intervention of strengthening, flexibility, and aerobic exercise was also significantly more effective than no exercise control for improving limitation in function. When we limited the analysis to trials focusing on knee osteoarthritis, the cumulative rankings did not change much but effect estimates tended to be larger compared with the overall analysis, which also included trials focusing on hip osteoarthritis and trials investigating osteoarthritis in any joint	As of 2002 sufficient evidence had accumulated to show significant benefit of exercise over no exercise in patients with osteoarthritis. Further trials are unlikely to overturn this result.
Fransen et 2015 (updated of 2008)	To determine whether land-based therapeutic exercise is beneficial for people with knee OA in terms of reduced joint pain or improved physical function and quality of life.	RCTs or quasi-RCTs, up to May 2013	*P* = Adults with knee OA (self-reported or as defined by the ACR criteria) I = Any land-based non-perioperative therapeutic exercise regimens. *C* = An active (given any non-exercise intervention) or no treatment (including waiting list) group. *O* = Knee pain, Self-reported physical function and Quality of life.	Meta-analysis and GRADE assessment	54 studies included. High-quality evidence from 44 trials (3537 participants) indicates that exercise reduced pain (standardised mean difference (SMD) −0.49, 95% confidence interval (CI) −0.39 to −0.59) immediately after treatment. Moderate-quality evidence from 44 trials (3913 participants) showed that exercise improved physical function (SMD -0.52, 95% CI -0.39 to −0.64) immediately after treatment.	High-quality evidence from 13 studies (1073 participants) revealed that exercise improved quality of life (SMD 0.28, 95% CI 0.15 to 0.40) immediately after treatment. High-quality evidence indicates that land-based therapeutic exercise provides short-term benefit that is sustained for at least two to 6 months after cessation of formal treatment in terms of reduced knee pain, and moderate-quality evidence shows improvement in physical function among people with knee OA. This reflects our belief that further research in this area is unlikely to change the findings of our review.
Bartels et al. 2016 (update of 2007)	To evaluate the effects of aquatic exercise for people with knee or hip osteoarthritis, or both, compared to no intervention.	RCTs published up to 28 April 2015	*P* = Participants with knee or/and hip OA as defined by ACR criteria. I =All types of exercises (e.g., ROM, strength, and aerobics) performed in a therapeutic/heated indoor pool. *C* = Control group (e.g., usual care, education, social attention, telephone call, waiting list for surgery) *O* = Pain, Disability, Quality of life and Radiographs	Meta-analysis and GRADE assessment	9 RCTs. Aquatic exercise caused a small short term improvement compared to control in pain (SMD −0.31, 95% CI −0.47 to −0.15; 12 trials, 1076 participants) and disability (SMD −0.32, 95% CI −0.47 to −0.17; 12 trials, 1059 participants). Ten trials showed a small effect on quality of life (QoL) (SMD −0.25, 95% CI −0.49 to −0.01; 10 trials, 971 participants).	Moderate quality evidence that aquatic exercise may have small, short-term, and clinically relevant effects on patient-reported pain, disability, and QoL in people with knee and hip OA. The conclusions of this review update does not change those of the previous published version of this Cochrane review.
Verhagen et al. 2019	To investigate if we need additional trials on exercise in knee OA to accept a certain effect size to be a 'true' effect size, and new	RCTs published up to 31 August 2016.	*P* = Adults with knee OA (self-reported or as defined by the ACR criteria) I = Exercise therapy (no Tai Chi or home exercises) *C* = No treatment (e.g., waiting list), a	Extended funnel plot Secondary analysis of: Bartels *et al.* Cochrane Database Syst Rev.	42 studies. a) there is no clear publication bias, b) subgrouping did not affect the overall effect estimate, c) the effect estimate of exercise is more consistent (no heterogeneity) in	Exercise is effective and clinically worthwhile in reducing pain immediately post treatment compared to no or minimal interventions in patients with knee

Table I (*continued*)

Aim	Studies included	Inclusion criteria (PICO)	Statistical method	Results	Conclusions
studies are not needed anymore.		minimal intervention (e.g., medication), or non-supervised exercise therapy (e.g., home-based exercise therapy). *O* = Pain immediately post treatment	2016 and Fransen *et al.* Cochrane Database Syst Rev. 2015	the studies of low Risk of Bias, d) the benefit of exercise was clear since 2010 e) the extended funnel plot suggests that an additional study has a none or very limited impact to change the current effect estimate	OA and adding new data will unlikely change this conclusion.

RCT = randomised controlled trial; OA = osteoarthritis; *P* = participant; I = intervention; *C* = comparison; *O* = outcome; SMD = standardised mean difference; QoL = quality of life; ACR = American college of rheumatology.

Are exercise RCTs still needed in OA research?

High quality head-to-head RCTs assessing outcomes beyond pain and physical function are still needed to estimate the relative benefits and harms of different OA treatments. For example, studies comparing exercise to surgery or exercise to pharmacological pain relievers, and the comparison of different exercise doses or types of exercise, particularly in people at risk of OA, or having early or end-stage OA seems to be lacking. Similarly, the inclusion of outcomes such as adherence to physical activity recommendations, the influence of physical and mental fatigue which may potentially help to identify relevant subgroups seems to require further investigation.

Methodological challenges and possible solutions

Methodological limitations in exercise trials are still common particularly regarding the difficult blinding of participants and also regarding recruitment and attrition rates[5–7]. For example, Bennell *et al.* 2016[11] and Sandal *et al.* 2017[12] blinded participants to the study hypothesis, which may be considered an appropriate strategy for blinding participants in exercise interventions. The involvement of end-users, such as patients, in designing RCTs seems to improve trial quality by recruiting a larger proportion of eligible participants, lowering attrition rates and selecting research questions that are clinically meaningful[13]. However, better strategies for systematically and transparently involving end-users in research are required.

Conclusions

There is no consensus on how to assess when ‘enough is enough’. The use of statistics can retrospectively help researchers to identify whether there is redundancy in the published studies. On the contrary, for assessing the need of future studies additional considerations need to be taken into account and these include: the end-user’s perspective, the quality of the available evidence and the specific context of the intervention.

Author contributions

Study conception and design. Bricca, Lund, Roos and Juhl.

Drafting the article or revising it critically for important intellectual content. Bricca, Lund, Roos and Juhl.

Final approval of the article. Bricca, Lund, Roos and Juhl.

Conflict of Interest

Dr. Roos is deputy editor of Osteoarthritis and Cartilage, the developer of the Knee injury and Osteoarthritis Outcome Score (KOOS) and several other freely available patient-reported outcome measures and co-founder of Good Life with Osteoarthritis in Denmark (GLA:D®), a not-for profit initiative hosted at University of Southern Denmark aimed at implementing clinical guidelines for osteoarthritis in clinical practice.

The authors report no other conflict of interest.

Acknowledgements

No funding was received for this editorial.

References

1. Open Science Collaboration. Estimating the reproducibility of psychological science. Science 2015;349, https://doi.org/10.1126/science.aac4716. aac4716.

2. Fergusson D, Glass KC, Hutton B, Shapiro S. Randomized controlled trials of aprotinin in cardiac surgery: could clinical equipoise have stopped the bleeding? Clin Trials 2005;2(3):218–29, https://doi.org/10.1191/1740774505cn085oa. discussion 229-32.
3. McAlindon TE, Bannuru RR, Sullivan MC, Arden NK, Berenbaum F, Bierma-Zeinstra SM, *et al*. OARSI guidelines for the non-surgical management of knee osteoarthritis. Osteoarthr Cartil 2014;22(3):363–88.
4. Nelson AE, Allen KD, Golightly YM, Goode AP, Jordan JM. A systematic review of recommendations and guidelines for the management of osteoarthritis: the chronic osteoarthritis management initiative of the U.S. bone and joint initiative. Semin Arthritis Rheum 2014 Jun;43(6):701–12.
5. Verhagen P, Ferreira M, Reijneveld-van de Vendel E, Teirlinck A, Runhaar J, van Middelkoop M, *et al*. Do We Need Another Trial on Exercise in Patients with Knee Osteoarthritis? Osteoarthritis and Cartilage 2019.
6. Fransen M, McConnell S, Harmer AR, Van der Esch M, Simic M, Bennell KL. Exercise for osteoarthritis of the knee. Cochrane Database Syst Rev 2015;1, CD004376.
7. Bartels EM, Juhl CB, Christensen R, Hagen KB, Danneskiold-Samsoe B, Dagfinrud H, *et al*. Aquatic exercise for the treatment of knee and hip osteoarthritis. Cochrane Database Syst Rev 2016;3. CD005523.
8. Uthman OA, van der Windt DA, Jordan JL, Dziedzic KS, Healey EL, Peat GM, *et al*. Exercise for lower limb osteoarthritis: systematic review incorporating trial sequential analysis and network meta-analysis. BMJ 2013 Sep 20;347:f5555, https://doi.org/10.1136/bmj.f5555.
9. Schmid C, Senn S, Sterne J, Kulinskaya E, Posch M, Roes K, et al. Should Cochrane Apply Error-Adjustment Methods When Conducting Repeated Meta-Analyses? https://methods.cochrane.org/sites/default/files/public/uploads/tsa_expert_panel_guidance_and_recommendation_final.pdf.
10. Treweek S, Bevan S, Bower P, Briel M, Campbell M, Christie J, *et al*. Trial Forge Guidance 2: How to Decide if a Further Study within A Trial (SWAT) Is Needed 18 October 2018, https://doi.org/10.21203/rs.1.1/v1. PREPRINT (Version 1) available at Research Square.
11. Bennell KL, Ahamed Y, Jull G, Bryant C, Hunt MA, Forbes AB, *et al*. Physical therapist-delivered pain coping skills training and exercise for knee osteoarthritis: randomized controlled trial. Arthritis Care Res 2016;68(5):590–602.
12. Sandal LF, Thorlund JB, Moore AJ, Ulrich RS, Dieppe PA, Roos EM. Room for improvement: a randomised controlled trial with nested qualitative interviews on space, place and treatment delivery. BrJSports Med 2019;53(6):359–67.
13. Crocker JC, Ricci-Cabello I, Parker A, Hirst JA, Chant A, Petit-Zeman S, *et al*. Impact of patient and public involvement on enrolment and retention in clinical trials: systematic review and meta-analysis. BMJ 2018;363:k4738.

A. Bricca*
Aberdeen Health Psychology Group, Institute of Applied Health Sciences, University of Aberdeen, Foresterhill Campus, Aberdeen, UK

H. Lund
Centre for Evidence-Based Practice, Western Norway University of Applied Sciences, Bergen, Norway

E.M. Roos
Research Unit for Musculoskeletal Function and Physiotherapy, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

C.B. Juhl
Research Unit for Musculoskeletal Function and Physiotherapy, Department of Sports Science and Clinical Biomechanics, University of Southern Denmark, Odense, Denmark

Department of Physiotherapy and Occupational Therapy, University Hospital of Copenhagen, Herlev and Gentofte, Denmark

* Address correspondence and reprint requests to: A. Bricca, Aberdeen Health Psychology Group, Institute of Applied Health Sciences, University of Aberdeen, Foresterhill Campus, Aberdeen, UK.
E-mail address: alessio.bricca@abdn.ac.uk (A. Bricca).