**SUPPLEMENTARY APPENDIX 1: Methods**

# 2023 American College of Rheumatology (ACR) and American Association of Hip and Knee Surgeons (AAHKS) Clinical Practice Guideline for the Optimal Timing of Elective Hip or Knee Arthroplasty for Patients with Symptomatic Moderate to Severe Osteoarthritis or Advanced Symptomatic Osteonecrosis with Secondary Arthritis for Whom Nonoperative Therapy is Ineffective

## *Methodology Overview*

This guideline was developed following the American College of Rheumatology (ACR) guideline development process (www.rheumatology.org/Portals/0/Files/ACR%20Guideline%20Manual_Appendices_updated%202015.pdf). This process includes the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) methodology (www.gradeworkinggroup.org) (1-4).

## *Teams Involved*

This project was a collaboration between the American College of Rheumatology (ACR) and the American Association of Hip and Knee Surgeons (AAHKS); all participating teams included representation from both organizations. A Core Leadership Team (6 members) met weekly to supervise the project and was responsible for confirming the scope and clinical (Patient/Intervention/Comparator/Outcomes – PICO) questions (Supplementary Appendix 2), coordinating with the Literature Review Team, overseeing the voting process, and drafting the manuscript. The Core Team, together with the Literature Review Team, was comprised of individuals who had content and methodological expertise, and included a GRADE methodologist who advised on

the process of developing and presenting the evidence and provided input on the quality assessment of evidence and summary of findings (SoF) tables (provided in Supplementary Appendix 3).

The Literature Review Team (19 members) conducted a systematic search with the assistance of an experienced medical librarian, screened papers for relevance, assessed study quality, extracted data, computed pooled estimates of outcomes, graded the quality of evidence, generated an evidence summary for each PICO, and compiled an evidence report.

The Voting Panel consisted of 13 people, including rheumatologists, orthopaedic surgeons, methodologists, and 2 patient representatives. The role of the Voting Panel was to vote on the drafted recommendation statements derived from the PICO questions, keeping the evidence report, their expertise and experience, as well as patient values and preferences in mind.

A Patient Panel was convened to discuss patient values and preferences related to outcomes, evidence, and drafted recommendation statements. The two patients on the Voting Panel also participated in the Patient Panel discussions. The Voting Panel used the input from the Patient Panel meeting to help guide their votes in balancing tradeoffs between the harms and benefits of the alternative management strategies.

The ACR provided training for everyone involved in the development of this guideline, which included explanations of the ACR guideline process and GRADE methodology. See Supplementary Appendix 4 for team/panel rosters.

***Patient Panel***

The Patient Panel, consisting of two adult men and 6 adult women who had osteoarthritis (OA) or osteonecrosis (ON) and had a total hip arthroplasty (THA) or total knee arthroplasty (TKA), or were candidates for THA or TKA, was convened virtually on August 24, 2022. Linda Russell, MD, a rheumatologist who has experience in guideline development and this patient population, facilitated the four-hour discussion. Two ACR staff members were also present. The participants were first

presented with the background and scope of the guideline project.

The Patient Panel then reviewed the evidence synthesized by the Literature Review Team as each PICO question was discussed. The participants were encouraged to consider their personal experiences relevant to the questions, judge the importance of the impact on outcomes, and make comments on the drafted recommendation statements accordingly. Two patients on the Voting Panel, who had been at the Patient Panel meeting, presented the values and preferences of the Patient Panel to the Voting Panel during the one-day Voting Panel meeting held August 29, 2022.

***Disclosures and Management of Conflicts of Interest***

Per ACR policy, everyone who was intellectually involved in the project (i.e., considered for guideline authorship) was required to disclose all relationships (https://www.rheumatology.org/Practice-Quality/Clinical-Support/Clinical-Practice-Guidelines/Indications-for-Total-Hip-and-Knee-Replacement. Disclosures were evaluated to determine if any relationships were considered potential conflicts of interest for purposes of this project. Individuals whose primary employment (> 51% of work time/effort) was with a company that manufactured or sold therapeutics, diagnostics, or total hip or total knee arthroplasty implants were not eligible to participate.

The project's co-principal investigators (PIs) and the co-Literature Review Team leaders had no relevant conflicts of interest for the full 12 months before this project began, and a majority of guideline development team members had no relevant conflicts of interest for the duration of the project. Intellectual conflicts, such as a prior publication or scientific presentation on THA or TKA, were recognized as important and were required to be disclosed, but because they were ubiquitous among this panel of academic physicians, were not counted as members being conflicted toward the allowed threshold.

Participant disclosures were initially shared in the project plan, which was posted online for public comment as the project began. Disclosures were updated and shared again with each project participant via email prior to the Voting Panel meeting. Updated participant disclosures are included online with this manuscript.

***Scope and Target Audience***

The scope of this project included the development of evidence-based recommendations regarding the indications for when to delay arthroplasty for additional non-operative treatments in patients who have advanced symptomatic moderate to severe osteoarthritis or advanced symptomatic osteonecrosis with secondary arthritis of the hip or knee for whom THA or TKA has been indicated. In addition, the project aimed to provide evidence-based recommendations regarding the timing of arthroplasty, in patients who have specific medical comorbidities, with the aim of decreasing surgical risk.

The target audience for this guideline are symptomatic adults who are candidates for THA or TKA who have radiographically moderate to advanced symptomatic osteoarthritis or advanced symptomatic osteonecrosis with secondary arthritis and their health care providers. Derivative products may be developed in the future to facilitate implementation of this guideline to these audiences.

***Establishing Key Principles and PICO Development***

The Core Leadership Team collaborated with Literature Review Team and Voting Panel members to develop the initial set of PICO-formatted clinical questions for the guideline, as well as to identify pre-specified outcomes that were considered critical for each PICO question (Supplementary Appendix 2).

The Core Leadership Team held weekly conference calls, convened an initial virtual meeting of the Core Leadership Team, Literature Review Team, and Voting Panel, in which the scope of the guideline was determined, and then developed the PICO questions. The PICO questions were posted for

30 days on the ACR website for public comment and revised accordingly.

***Systematic Synthesis of the Literature***

***Literature Searches***

To identify relevant evidence for the PICO questions, a medical librarian, in collaboration with the Core Team, performed systematic searches of the published English language literature. Ovid MEDLINE and Ovid Embase original searches were performed from the beginning of each database to September 27, 2021, and updated searches were later performed from September 28, 2021 to June 19, 2022 (see Supplementary Appendix 5, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25175).

***Study Selection***

DistillerSR software (https://distillercer.com/products/distillersrsystematic-review-software/) was used to aid screening the literature search results. Teams of two independent reviewers performed duplicate screening of each title and abstract with articles identified as potentially eligible passing to review of full text. Eligible articles underwent full-text screening by two independent reviewers. Selected manuscripts were matched to PICO questions. See Supplementary Appendix 6, available on the *Arthritis Care & Research* website at http://onlinelibrary.wiley.com/doi/10.1002/acr.25175, for details related to the study selection process.

***Data Extraction and Analysis***

Comparative data (e.g., from RCTs and some controlled observational studies) for each PICO question was extracted into RevMan software (http://tech.cochrane.org/revman). Risk of bias of each primary study was assessed using the Cochrane risk of bias tool (http://handbook.cochrane.org/). The critical outcomes selected for this guideline were mostly binary, and they were reported as relative

risks with 95% confidence intervals. Continuous outcomes were reported as mean differences. Data not suitable for RevMan was extracted into Word tables.

In clinical scenarios not addressed by RCT or controlled observational data, data from observational uncontrolled studies were used to estimate effects. In situations in which evidence for a specific intervention in a patient population with moderate to severe osteoarthritis or osteonecrosis of the hip or knee was sparse or absent, evidence for the intervention in a non-RMD population was included. In these cases, the effect sizes in non-RMD patients were postulated to be generalizable, but the quality of evidence was lowered by rating down for indirectness.

***Evidence Report Formulation***

RevMan files were exported into GRADEpro software to formulate a GRADE Summary of Findings (SoF) table for each PICO question (2), when possible. The quality of evidence for each outcome was evaluated by one literature review team member, then verified by the co-literature review leaders (JS, CH) using GRADE quality assessment criteria (1) with discordance resolved by discussion. The resulting SoF tables were compiled in an evidence report (Supplementary Appendix 3). The Core Leadership Team reviewed the evidence report prior to presentation to the Voting Panel.

***Moving from Evidence to Recommendations***

The GRADE methodology specifies that Voting Panels make recommendations based on a consideration of the balance of benefits and harms/burdens of the treatment options under consideration, the quality of the evidence (i.e., confidence in the effect estimates), and patients' values and preferences. Keys to the recommendation are the trade-off between desirable and undesirable outcomes; recommendations require estimating the relative value patients place on the outcomes. Although costs of additional conservative treatments was considered part of GRADE methodology and raised in the recommendations, the literature is limited in terms of formal cost effectiveness analyses (CEA).

A recommendation could be either in favor of or against the proposed intervention and either strong or conditional. According to GRADE, a recommendation was categorized as strong if the panel was very confident that the benefits of an intervention clearly outweigh the harms (or vice versa); a conditional recommendation denotes uncertainty regarding the balance of benefits and harms, such as when the evidence quality is low or very low, or when the decision is sensitive to individual patient preferences, or when costs are expected to impact the decision. Thus, conditional recommendations refer to decisions in which shared decision-making incorporating patient preferences is essential.

Judgments made in this guideline were based on the experience of the clinician panel members in shared decision-making with their patients, on the experience and perspectives of this guideline's Patient Panel members, and to a considerable extent, on the results of discussion with the Patient Panel.

***Consensus Building***

The Voting Panel received the evidence report for review before it met to discuss and decide on the final recommendations. Individual online voting took place first, to ascertain any existing consensus on drafted recommendation statements that were based on the PICOs. This was followed by a 1-day virtual webinar meeting of the Voting Panel, where they reviewed the evidence and provided final votes on the direction and strength of each recommendation. The webinar voting process was conducted using Poll Everywhere software (www.polleverywhere.com). A 70% consensus was used as the threshold for a recommendation; if 70% consensus was not achieved during an initial vote, the panel members held additional discussions before re-voting until at least 70% consensus was achieved.

***Final Review and Approval of the Manuscript by the ACR***

In addition to journal peer reviews, the manuscript was reviewed by the ACR Guideline Subcommittee, the ACR Quality of Care Committee, and the ACR Board of Directors, as well as the AAHKS Evidence Based Medicine Committee, and AAHKS Board of Directors. These ACR and AAHKS

oversight groups did not make or mandate that specific recommendations be made within the guideline, but rather, served as peer reviewers.

***Moving from Recommendations to Practice***

These recommendations are designed to support health care providers and patients in selecting treatment options. Health care providers and patients must take into consideration disease severity, pain, and functional limitations, as well as comorbidities, surgical risks, as well as patient's values and preferences, in choosing the timing of optimal therapy for an individual patient at the given point in treatment.

**REFERENCES**


1. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008;336(7650):924-6.
2. Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. J Clin Epidemiol 2011;64:395-400.
3. Andrews J, Guyatt G, Oxman AD, Alderson P, Dahm P, Falck-Ytter Y, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. J Clin Epidemiol. 2013;66(7):719-25.
4. Andrews JC, Schunemann HJ, Oxman AD, Pottie K, Meerpohl JJ, Coello PA, et al. GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. J Clin Epidemiol. 2013;66(7):726-35.