SUPPLEMENTARY APPENDIX 1: Methods

2021 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis

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) (1-3).

Teams Involved

A Core Leadership Team (five members) supervised the project and was responsible for defining the scope, drafting the clinical (Patient/Intervention/Comparator/Outcomes – PICO) questions, 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 with 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 2).

The Literature Review Team (17 members) conducted a systematic search, screened papers for relevance, assessed study quality, extracted data, computed pooled estimates of outcomes, graded the quality of evidence, generated the SoF tables, and compiled an evidence report.

The role of the Expert Panel, comprised of 10 content experts, was to provide consultation and feedback on the project scope, design, and PICO questions.

The Voting Panel consisted of 15 people, including rheumatologists, a physician assistant and two patient representatives diagnosed with rheumatoid arthritis (RA). The role of the Voting Panel was to participate in the development of the scope and PICO questions, including making judgments regarding the relative importance of the outcomes, and vote on the PICO questions, keeping the evidence report, their expertise and experience, and 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 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 sessions on the ACR guideline process and GRADE methodology. See Supplementary Appendix 5 for team/panel rosters.

Patient Panel

The Patient Panel, consisting of one adult male and nine adult females diagnosed with RA, was convened on January 9, 2020. A member of the Core Team and one ACR staff person facilitated the day-long discussion.

The participants were first presented with the background and scope of the guideline project. They were then specifically queried on the relative importance of beneficial and adverse events of drugs and drug classes, including but not limited to efficacy, route of administration, and side effects, with particular attention paid to how values and preferences might differ. The Patient Panel reviewed the evidence synthesized by the Literature Review Team as several PICO questions were discussed. The participants were encouraged to consider

their personal experiences relevant to the questions and judge the importance of the outcomes and vote on the drafted recommendation statements accordingly. Two patients on the Voting Panel, who had been at the patient meeting, presented the values and preferences of the Patient Panel and their voting results to the Voting Panel during the two-day Voting Panel meeting held January 10-11, 2020.

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 (www.rheumatology.org/Practice-Quality/Clinical-Support/Clinical-Practice-Guidelines/Rheumatoid-Arthritis). Disclosures were compared against a previously drafted list of "affected companies" (i.e., companies or organizations that were considered reasonably likely to be positively or negatively affected by care delivered in accordance with the guideline) to determine which relationships were considered potential conflicts of interest for purposes of this project. Individuals were also asked to explicitly highlight relationships with any companies *not* on the affected companies list that related to the topic of the guideline. Individuals whose primary employment (> 51% of work time/effort) was with a company that manufactured or sold therapeutics or diagnostics were not eligible to participate.

The project's principal investigator (PI) and the literature review leader had no relevant conflicts of interest for the full 12 months before this project began, and the majority of the guideline development team members had no relevant conflicts of interest for the duration of the project. A participant who had any relationship with an affected company was counted as conflicted (i.e., toward the allowed threshold) regardless of the type or subject of the relationship. Intellectual conflicts, such as a prior publication or scientific presentation on RA therapy, were recognized as important and were required to be disclosed, but because

they were ubiquitous, intellectual conflicts were not counted as conflicted toward the allowed threshold.

Participant disclosures were included in the project plan that was posted online for public comment (see description below). In addition, disclosures of all participants were shared, in writing, with each project participant. At the face-to-face Voting Panel meeting, verbal disclosures were provided before the content discussion began, and the same information was provided via tent cards and in a written summary provided to all participants at the beginning of the meeting. Updated participant disclosures, as well as ACR committee reviewer disclosures, are included online with this manuscript. Finally, author disclosures are also included in this paper.

Scope and Target Audience

The scope of this project included pharmacologic treatment of patients with RA. The target audience for this guideline includes health care providers and patients with RA. The ACR plans to develop derivative products to facilitate implementation of this guideline.

Establishing Key Principles and PICO Development

The Core Leadership Team collaborated with the Voting and Expert Panel members to develop the initial set of PICO-formatted clinical questions for the guideline, as well as identify pre-specified outcomes as critical or important for each PICO question (see Supplementary Appendix 3).

The Core Leadership Team held weekly conference calls, convened an initial face-to-face meeting of the Core Leadership Team, Voting Panel and Expert 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. Once the PICO questions were finalized, individual online voting took place to ascertain any existing

consensus, followed by a face-to-face meeting of the Voting Panel, where voting on the PICO questions was finalized. Following the meeting, additional clarifying questions were discussed by email and related voting took place via online survey.

Framework for the RA Guideline Development

At the initial scoping meeting, the Core Leadership Team, Voting Panel and Expert Panel members agreed that the guideline should include pharmacologic treatment approaches and elected that recommendations would be developed to address treatment with: conventional synthetic DMARDs (csDMARDs), biologic DMARDs (bDMARDs), and targeted synthetic DMARDs (tsDMARDs); glucocorticoids, and use of these medications in certain high-risk populations. The use of vaccines and non-pharmacologic treatment approaches was originally discussed as a part of this guideline, but instead will be covered in separate, future ACR guidelines. After defining population risk groups, interventions and comparators were specified for each PICO question (see list of PICO questions in Supplementary Appendix 2).

Systematic Synthesis of the Literature

Literature Searches

To identify relevant evidence for the PICO questions, a medical librarian, in collaboration with the Literature Review Team, performed systematic searches of the published English language literature. OVID Medline, PubMed, Embase, and the Cochrane Library (including Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (DARE); Cochrane Central Register of Controlled Trials (CENTRAL); and Health Technology Assessments (HTA)) were searched from the beginning of each database through December 11, 2019 (Supplementary Appendix 7).

Study Selection

DistillerSR software (https://distillercer.com/products/distillers-systematic-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 then matched to PICO questions. See Supplementary Appendix 6 for details related to the study selection process.

Data Extraction and Analysis

Data from RCTs 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/important outcomes selected for this guideline were binary, and they were analyzed using the Mantel-Haenszel method in a random effects model and reported as relative risks with 95% confidence intervals.

In clinical scenarios not addressed by RCT data, data from observational cohort studies was used to estimate relative effects. In situations in which the intervention had not been tested in RA but had been tested in a non-RA population, the effect sizes from that study were applied, postulating that the effect was generalizable but rating down the quality of evidence 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 (4). The quality of evidence for each outcome was evaluated by one reviewer (either SY or LK) then verified by the lead reviewer (EAA) using GRADE quality assessment criteria (1) with discordance resolved by discussion. The resulting

SoF tables were compiled in an evidence report (Supplementary Appendix 2). The Core Leadership Team reviewed the evidence report and addressed possible evidence gaps prior to presentation to the Voting Panel. In addition, existing cost-effectiveness studies were included with the evidence reports and cost estimates were retrieved from Lexicomp (see Supplementary Appendix 4).

Moving from Evidence to Recommendations

GRADE methodology specifies that panels make recommendations based on a consideration of the balance of benefits and harms of the treatment options under consideration, the quality of the evidence (i.e., confidence in the effect estimates), and patients' values and preferences. Key to the recommendation is the trade-off between desirable and undesirable outcomes; recommendations require estimating the relative value patients place on the outcomes.

A recommendation could be either in favor of or against the proposed intervention and either strong or conditional. According to GRADE, a recommendation is categorized as strong if the panel is 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 incorporation of patient preferences is a particularly essential element of decision making.

Judgments are based on the experience of the clinician panel members in shared decision making with their patients, on the experience and perspectives of the Patient Panel members and, to a considerable extent, on the results of discussion with the patient focus group.

Consensus Building

The Voting Panel received the evidence report for review before it met to discuss and decide on the final recommendations. During the two-day, face-to-face meeting, the Voting Panel, for each PICO question, reviewed the evidence and provided votes on the direction and strength of the recommendations. The in-person voting process was conducted using Poll Everywhere software (www.polleverywhere.com), with a follow-up online survey to vote on clarifications/unresolved questions. 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.

Consistent with GRADE guidance, in some instances, the Voting Panel chose to provide a strong recommendation despite a low or very low quality rating of evidence (3). In such cases, a written explanation is provided describing the reasons behind this decision with reference to GRADE guidance on the matter (3).

Final Review and Approval of the Manuscript by the ACR

In additional to journal peer reviews, the manuscript was reviewed by the following committees and subcommittees of the ACR: ACR Guideline Subcommittee; ACR Quality of Care Committee; and ACR Board of Directors. These ACR oversight groups did not mandate that certain 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 who work with patients in selecting therapies. Health care providers and patients must take into consideration not only clinical phenotype and level of disease activity, but also comorbidities, response and tolerance of prior therapies, patient's values and preferences, and patient's functional status

and functional goals in choosing the 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. 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.
3. 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.
4. 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.