

SUPPLEMENTARY APPENDIX 1: Methods

2020 American College of Rheumatology Guideline for the Management of Reproductive Health in Rheumatic and Musculoskeletal Diseases

Methodology Overview

This guideline followed the American College of Rheumatology (ACR) guideline development process (<http://www.rheumatology.org/Practice-Quality/Clinical-Support/Clinical-Practice-Guidelines>). This process includes using GRADE methodology (www.gradeworkinggroup.org) to rate the quality of the available evidence and to develop the recommendations (1-3). GRADE methodology specifies that panels make recommendations based on the balance of benefits and harms, the quality of the evidence (i.e., confidence in effect estimates), and patients' values and preferences. In this case, however, the panel did not adhere to GRADE guidance on two issues. First, the panel made some GRADE recommendations without a formal rating of the quality of the evidence that came from non-RMD populations because the magnitude of indirect evidence that was considered was excessive. Second, the panel chose to make a number of strong recommendations based on low or very low quality evidence that could not be justified using GRADE criteria (3) because, despite the limitations of the available evidence, the potential harm of not proceeding with the recommendation was thought to outweigh other potential harms. The potential harms of not proceeding with these recommendations included potentially catastrophic negative outcomes such as an organ- or life-threatening disease flare or a potentially fatal thromboembolic event. While unlikely, these are severe and unpredictable risks in many situations and were felt to warrant strong recommendations despite the low quality evidence.

This work involved four teams selected by the ACR Quality of Care Committee and Guideline Subcommittee after reviewing individual and group volunteer applications to an open call for interested participants: 1) a Core Leadership Team, which supervised and coordinated the project and drafted the clinical questions, recommendation statements and manuscript; 2) a Literature Review Team, which completed the literature screening, data abstraction and synthesis; 3) an Expert Panel, composed of rheumatologists who helped develop the clinical questions and decide on the project scope; and 4) a Voting Panel, which included rheumatologists, a reproductive immunologist, obstetricians/gynecologists, epidemiologists, and two patients (one a pharmacist). Additionally, a Patient Panel consisting of 12 female RMD patients with varied experiences related to reproductive health provided input on their values and preferences, which was reviewed before discussion of each section of the guideline (e.g. contraception, fertility therapies and preservation, pregnancy management, menopause, medications) and was incorporated into discussions and formulation of recommendations. Supplementary Appendix 5 presents rosters of all guideline development team members. In accordance with ACR policy, the principal investigator and the Literature Review Team leader were free of conflicts, and all teams had >50% members free of conflicts. ACR policy guided disclosures and the management of conflicts of interest (<https://www.rheumatology.org/Practice-Quality/Clinical-Support/Clinical-Practice-Guidelines/Reproductive-Health-in-Rheumatic-Diseases>).

Framework for the Reproductive Health Guideline Development and Scope of the Guideline

At the scoping meeting, the Core Leadership Team, Voting Panel and Expert Panel decided that the guideline would address several areas, including medications and pregnancy (safety of paternal medication exposure, medication safety during pregnancy, and medication safety during lactation); management of pregnancy (counseling and medication transition in anticipation of pregnancy, pregnancy management issues, and management of anti-Ro/SSA and/or anti-La/SSB antibody as well as antiphospholipid antibody positive patients); fertility (including assisted reproductive technology in RMD patients, and fertility preservation with cyclophosphamide therapy); and other reproductive health issues (contraception in RMD patients, menopause/hormone replacement therapy). Human papillomavirus (HPV) prevention and screening and treatment of gynecologic cancers in RMD patients were initially considered as part of this project, but were removed during the scoping meeting due to the already broad spectrum of topics covered as well as the consideration that these topics might be better grouped with potential future recommendations concerning vaccination and malignancy in RMD patients. Recommendations regarding long-term outcomes of offspring were originally considered as well, but ultimately were felt to be beyond the scope of the current guideline.

Systematic Synthesis of the Literature

Direct evidence in RMD patient populations relating to reproductive health questions was obtained through systematic searches of the published English-language literature,

including OVID Medline, PubMed, Embase, and the Cochrane Library (including Cochrane Database of Systematic Reviews, Database of Abstracts of Reviews of Effects, Cochrane Central Register of Controlled Trials, and Health Technology Assessments) from the beginning of each database through November 15, 2017 (Supplementary Appendix 3, Search Strategies); updated searches were conducted on May 9, 2018. Duplications were identified via DistillerSR software (<https://distillercer.com/products/distillersr-systematic-reviewsoftware/>) (Supplementary Appendix 2, Study Selection Process).

All retrieved articles were screened in duplicate and the lead methodologist resolved any conflicts. For all included papers, reviewers entered extracted data describing details of the population, interventions (if any), and results into RevMan v.5.3 software (<http://tech.cochrane.org/revman>) which was used to calculate summary effect sizes (4) and evaluate risk of bias with the Cochrane risk of bias tool (<http://handbook.cochrane.org/>). RevMan files were exported into GRADEpro software to formulate a GRADE summary of findings table (Supplementary Appendix 6) for each PICO question (5). Evidence-based models use the PICO process for framing a question; PICO elements include Population, Intervention, Comparison, and Outcome. For data not appropriate for RevMan (e.g., non-comparative data), reviewers abstracted data describing details of the population, interventions (if any), and results into Word tables. GRADE criteria provided the framework for judging the overall quality of evidence (1).

Additional literature reviews in non-RMD populations were conducted by Core Team members to identify relevant indirect evidence not captured in the project's overall

systematic literature review, which focused specifically on evidence in RMD patient populations. This information was summarized and then reviewed by other Core Team members to assess relevance before being finalized and considered by the Voting Panel in their deliberations about the final recommendations (Supplementary Appendix 4). Evidence derived from the additional literature reviews was not graded as it was not part of the systematic literature review which focused on RMD patients.

Moving from Evidence to Recommendations

Given that GRADE methodology specifies that panels make recommendations based on the balance of benefits and harms, the quality of the evidence, and patients' values and preferences, deciding on the balance between desirable and undesirable outcomes requires estimating the relative value patients place on those outcomes. When the literature provided very limited guidance, the experience of the Voting Panel members in managing the relevant patients and problems also provided an important source of evidence. Patient values and preferences were crucial to all recommendations made and derived from input from the members of the Patient Panel; these were particularly salient in situations with limited literature.

The Voting Panel made every effort to adhere to GRADE guidance that specifies that strong recommendations should in general be based on high or moderate quality evidence, and that there are a restricted set of circumstances that warrant strong recommendations based on low or very low quality evidence (3). Much of the evidence from non-RMD populations was likely of moderate quality; however, this evidence was not formally graded due to inadequate time and resources. Given the low and very quality evidence specific to RMD populations, the data from non-RMD populations

ultimately played an important role in the decision-making process. Furthermore, it should be noted that many of the strong recommendations with low or very low quality evidence relied on the possibility of potentially catastrophic negative outcomes such as an organ- or life-threatening disease flare or a potentially fatal thromboembolic event, unlikely but real risks that were felt to warrant strong recommendations despite the low quality evidence.

Consensus Building

During a two-day face-to-face meeting and subsequent webinars and group emails, Voting Panel members voted on the direction (for or against) and strength (conditional or strong) of the recommendations related to the PICO questions. Not all PICO-generated recommendations were voted upon, although all were presented to the Voting Panel; as a result, numbering of guideline statements is not sequential. Some recommendation statements were dropped due to lack of data or relevance, and others were combined or changed to good practice statements based on level of evidence and Voting Panel discussion. A 'good practice' statement is one in which a large and compelling body of indirect evidence, made up of linked evidence using several indirect comparisons, strongly supports the net benefit of the recommended action. In many situations the case for a good practice statement rather than a GRADE recommendation is the use of time in collecting and summarizing the relevant evidence. Recommendations required a 70% level of agreement as used previously in other similar processes (6). If 70% agreement was not achieved during an initial vote, the panel members held additional discussions before re-voting. For all conditional

recommendations, a written explanation is provided, describing the reasons for this decision.

Moving from Recommendations to Practice

These recommendations are designed to help health care providers, caregivers, and patients engage in shared decision-making regarding disease management. Level of disease activity, comorbidities, response and tolerance of prior therapies, and patient-specific factors, values and preferences at the given point in their reproductive lifespan, should all be taken into consideration in choosing optimal therapy.

REFERENCES

1. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336(7650):924-926. doi:10.1136/bmj.39489.470347.AD
2. Andrews J, Guyatt G, Oxman AD, et al. GRADE guidelines: 14. Going from evidence to recommendations: the significance and presentation of recommendations. *J Clin Epidemiol*. 2013;66(7):719-725. doi:10.1016/j.jclinepi.2012.03.013
3. Andrews JC, Schünemann HJ, Oxman AD, et al. GRADE guidelines: 15. Going from evidence to recommendation—determinants of a recommendation's direction and strength. *J Clin Epidemiol*. 2013;66(7):726-735. doi:10.1016/j.jclinepi.2013.02.003
4. Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.
5. Guyatt GH, Oxman AD, Kunz R, et al. GRADE guidelines: 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol*. 2011;64(4):395-400. doi:10.1016/j.jclinepi.2010.09.012
6. Jaeschke R, Guyatt GH, Dellinger P, et al. Use of GRADE grid to reach decisions on clinical practice guidelines when consensus is elusive. *BMJ*. 2008;337:a744. <http://www.ncbi.nlm.nih.gov/pubmed/18669566>.