

SUPPLEMENTARY APPENDIX 1: ACR/NPF 2018 Psoriatic Arthritis Guideline Methods

Methodology Overview

We developed this guideline following the American College of Rheumatology (ACR) guideline development process (<http://www.rheumatology.org/Practice-Quality/Clinical-Support/Clinical-Practice-Guidelines>). This process includes the Grading of Recommendations Assessment, Development, and Evaluation (GRADE) methodology (www.gradeworkinggroup.org) (1-3).

Teams Involved

This project was a collaboration between the American College of Rheumatology (ACR) and the National Psoriasis Foundation (NPF); all participating teams included representation from both organizations. A Core Leadership Team (5 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 5).

The Literature Review Team (5 members) conducted a systematic search, 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, composed of 12 content experts, was to provide consultation and feedback on the project scope, design, and PICO questions, and to participate in manuscript preparation.

The Voting Panel (16 members) included rheumatologists, one dermatologist, one dermatologist-rheumatologist and one rheumatology physician assistant, internal medicine specialists with expertise and clinical experience in treating psoriatic arthritis (PsA), and 2 patient representatives. 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 and evidence. The Voting Panel used the results of the patient meeting to 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 2 for team/panel rosters.

Patient Panel

The patient panel consisting of 9 adults with PsA was convened on April 24, 2017. The median age of the participants was 50 years (range of 31 to 67), 7 of the 9 were female and the median duration of disease was 14 years (range of 3 to 41). The majority of the panel had previously used or were currently using a biologic drug. Six of the 9 patients had previously used methotrexate. Eight of the nine patients had axial symptoms or spondylitis. One member

of the Core Leadership Team, one member of the Voting Panel, and one ACR staff person facilitated the day-long discussion.

The participants, all of whom had completed research and guideline methodology webinars prior to meeting, were presented with the background and scope of the guideline project. The patients were specifically queried on the relevant importance the onset of drug action, route of administration, relative importance of beneficial and adverse events of drug classes, and importance of non-pharmacologic therapies and potential drawbacks of these therapies. The patient panel 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 and judge the importance of the outcomes accordingly. One core team member, who facilitated the patient panel meeting, presented the values and preferences of the patient panel and the voting results to the Voting Panel by during the two-day Voting Panel meeting held May 20-21, 2017.

Disclosures and Management of Conflicts of Interest

Per ACR policy, everyone who was intellectually involved in the project (i.e., considered for guideline authorship) disclosed all relationships (<https://www.rheumatology.org/Practice-Quality/Clinical-Support/Clinical-Practice-Guidelines/Psoriatic-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 PsA 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. 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 both pharmacologic and non-pharmacologic treatment of patients with active psoriatic arthritis. Active psoriatic arthritis was defined as the presence of any of: actively inflamed joints, active spondylitis, enthesitis, dactylitis, active psoriasis or nail lesions, as well as extra-articular features such as uveitis or inflammatory bowel disease. Clinical situations not addressed by this guideline include specific measures of patient assessment, severity of the disease, presence of oligoarticular disease, therapy in the

setting of concomitant conditions other than inflammatory bowel disease, diabetes, and serious infections, specific treatment for psoriasis (the latter is being defined by the American Academy of Dermatology in conjunction with the National Psoriasis Foundation). The panel did not consider outcomes that were felt to be important but not crucial or for which there was insufficient data. The target audience for this guideline includes health care providers and patients who are at risk for or have PsA. The ACR and the NPF plan to develop derivative products to facilitate implementation of this guideline.

Establishing Key Principles and PICO Development

The Core Team collaborated with the Voting and Expert Panel members to develop the initial set of PICO-formatted clinical questions for the guideline (4). The critically important outcomes included the ACR response criteria 20% improvement (ACR20), the Health Assessment Questionnaire Disability Index (HAQ-DI) (achievement of the minimal clinically important difference of 0.35) (4), the Psoriasis Area and Severity Index 75% improvement (PASI75) and adverse effects of treatments, in particular the incidence of serious infections. The Core Team held weekly conference calls, convened an initial face-to-face meeting of the Core 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, an electronic voting took place, followed by a face to face meeting of the voting panel, where voting on the PICO questions was finalized. As a few questions remained, an additional WebEx call took place for the final vote.

Framework for the PsA Guideline Development

At the initial scoping meeting, the Core Team, Voting Panel and Expert Panel members agreed that the scope of the populations to be addressed would include patients with active

psoriatic arthritis causing symptoms at an unacceptably bothersome level as reported by the patient, and judged by the examining clinician to be due to PsA based on ≥ 1 of the following: swollen joints; tender joints; dactylitis; enthesitis; axial disease; active skin and/or nail involvement; and extra-articular inflammatory manifestations such as uveitis, or inflammatory bowel disease. The examining health care provider may take into account inflammatory markers (ESR, CRP), and imaging.

After defining population risk groups, interventions and comparators were specified for each PICO question (see list of PICO questions in Supplementary Appendix 5). The Core Team agreed that the guideline should include both pharmacologic and non-pharmacologic treatment, in both treatment-naïve patients and in patients treated with various levels of treatment. The Core Team elected to include medications that had completed phase III trials and had at least one approved indication in the United States at the time of drafting the questions in September 2016 (e.g., abatacept, tofacitinib, ixekizumab and brodalumab) for consideration in the PICO questions. PICO questions removed from the list were kept to be addressed in a follow up PsA guideline.

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. We searched 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)) from the beginning of each database through November 15, 2016 (Supplementary Appendix 3), and updated searches were conducted on

May 2, 2017, and again on March 8, 2018. For PICO questions for which we found no direct evidence in the PsA field, we sought indirect evidence: in particular, meta-analyses of randomized trials in non-PsA populations. For PICO questions for which we could not find systematic reviews, we sought individual RCTs or observational studies of non-PsA populations.

Study Selection

We used DistillerSR software (<https://distillercer.com/products/distillersr-systematic-Yreview-software>) 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 4 for details related to the study selection process.

Data Extraction and Analysis

We extracted data from RCTs for each PICO question 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 (ACR20, PASI75, HAQ-DI MCID, and MDA) 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.

Since the majority of RCTs comprised drug-placebo comparisons rather than drug-drug comparisons, network meta-analysis was used to generate risk ratios (RR) from indirect comparison of different drug classes for the outcomes ACR20, HAQ-DI, and PASI-75. In Stata, we used the “network” suite of commands for meta-analysis, which utilize mvmeta command and methods (5). Each network meta-analysis was conducted on the logarithm of proportions,

and preserved randomization by comparing medications to placebo. For PICO questions that had a smaller evidence base (fewer studies), we performed drug-drug comparisons using the adjusted indirect comparison method (6).

In clinical scenarios not addressed by RCT data (e.g., certain special populations, such as patients with diabetes), we used data from observational cohort studies to estimate relative effects. In situations in which the intervention had not been tested in PsA but had been tested in a non-PsA population, we applied the relative risk values from that study, postulating that that the effect was generalizable but rating down the quality of evidence for indirectness.

Evidence Report Formulation

We exported RevMan files into GRADEpro software to formulate a GRADE summary of findings (SoF) table for each PICO question (7). Data from network meta-analyses and adjusted indirect comparisons were manually entered into GRADE SoF tables. The quality of evidence for each outcome was evaluated in duplicate by two independent reviewers using GRADE quality assessment criteria (1) with discordance resolved by discussion. We compiled the resulting SoF tables in an evidence report (Supplementary Appendix 5). The Core Team reviewed the evidence report and addressed possible evidence gaps prior to presentation to the Voting Panel.

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 an essential element of decision making.

We are unaware of published literature exploring patient values and preferences regarding these issues in the context of PsA. Our 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 two patient panel members and, to a considerable extent, on the results of discussion with our 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 a two-day, face-to-face meeting held May 20-21, 2017, and a subsequent conference call and e-mails, the Voting Panel, for each PICO question, reviewed the evidence and provided votes on the direction and strength of the recommendations. The initial voting process was conducted using Poll Everywhere software (<http://www.polleverywhere.com/>) with a follow-up conference call to vote on 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.

In some instances, the Voting Panel decided to split statements into their more granular components (e.g., oral small molecules [OSM] vs. TNFi was split into APR vs. TNFi, MTX vs. TNFi, etc.). 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 addition to journal peer reviews, the manuscript was reviewed by the following committees and subcommittees of the ACR and NPF: ACR Guideline Subcommittee; ACR Quality of Care Committee; ACR Board of Directors; and NPF Medical Board. These ACR and NPF 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 help health care providers work with patients in selecting therapies. The presence or absence of conditions such as inflammatory bowel disease, uveitis, diabetes, and serious infections and the knowledge of previous therapies will help guide this process. The physical examination in the context of PsA, also required for selecting therapy, includes assessment of the peripheral joints (including dactylitis), the entheses, and the skin. Knowledge of inflammatory spine disease and/or spine symptoms is also important.

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. Mease PJ, Woolley JM, Bitman B, Wang BC, Globe DR, Singh A. Minimally important difference of Health Assessment Questionnaire in psoriatic arthritis: relating thresholds of improvement in functional ability to patient-rated importance and satisfaction. *The J Rheumatol* 2011;38(11):2461-5.
5. White IR, Barrett JK, Jackson D, Higgins JPT. Consistency and inconsistency in network meta-analysis: model estimation using multivariate meta-regression. *Res Synth Methods* 2012;3(2): 111-125.
6. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *J Clin Epidemiol* 1997;50(6): 683-691.
7. 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.