Assessing Single Joints in Arthritis Clinical Trials

HELEN I. KEEN, CLIFTON O. BINGHAM III, LAURENCE A. BRADLEY, PHILIP G. CONAGHAN, ALISON E. HEALD, GURJIT S. KAELEY, WALTER P. MAKSYMOWYCH, ROLAND W. MOSKOWITZ, H. RALPH SCHUMACHER, Jr, THASIA E. WOODWORTH, and PHILIP J. MEASE

ABSTRACT. The need to develop validated outcome measures to assess response to therapies in single joints has been recognized. In 2004, a task force was established to assess established and novel outcome measures in accordance with the OMERACT filter (truth, discrimination, and feasibility) for single joint assessment. This report describes the proceedings of the single joint assessment special interest group (SIG) at OMERACT 9, including an updated literature review of imaging of the knee joints, with a focus on the extent to which these modalities fulfill the OMERACT filter. A series of studies are reported that examine patient reported, clinical examination, and imaging outcomes in therapeutic studies in knee arthritis. A summary of discussions from the meeting are presented that raise many of the ongoing challenges in establishing appropriate domains to evaluate a range of conditions and potential therapeutic interventions. Because of emerging drug candidates and modalities targeting individual joints, the ongoing work of this SIG is providing the evidence base that can be used to establish a core domain set to incorporate as outcomes in future studies. (J Rheumatol 2009;36:2092–6; doi:10.3899/jrheum.090364)

Key Indexing Terms:
KNEE ARTHRITIS RELIABILITY VALIDITY
RESPONSIVENESS OUTCOME ASSESSMENT JOINT EXAMINATION

From the School of Medicine and Pharmacology, University of Western Australia, Perth, Australia; Divisions of Rheumatology and Allergy and Clinical Immunology, Johns Hopkins University, Baltimore, MD; Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham, USA; Section of Musculoskeletal Disease, University of Leeds, Leeds, UK; Clinical Affairs, Targeted Genetics Corporation, Seattle, WA; ROAD Clinic, (Rheumatology, Osteoporosis and Arthritic Diseases Clinic), University of Washington, Seattle, WA, USA; Department of Medicine, University of Alberta, Edmonton, Alberta; Case Western Reserve University, Division of Rheumatic Diseases, University Hospitals, Cleveland, Ohio; University of Pennsylvania VA Medical Center, Philadelphia, Pennsylvania, USA; Roche Products, Ltd., Welwyn, UK; Seattle Rheumatology Associates; Division of Rheumatology Research, Swedish Medical Center, University of Washington School of Medicine, Seattle, Washington, USA.

H.I. Keen, MB, BS, FRACP, Senior Lecturer, School of Medicine and Pharmacology, University of Western Australia; C.O. Bingham III, MD, Assistant Professor of Medicine, Divisions of Rheumatology and Allergy and Clinical Immunology, Johns Hopkins University; L.A. Bradley, PhD, Professor of Medicine, Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham; P.G. Conaghan, MB, BS, PhD, FRACP, FRCP, Professor of Musculoskeletal Medicine, Section of Musculoskeletal Disease, University of Leeds; A.E. Heald, MD, Senior Director, Clinical Affairs, Targeted Genetics Corporation; G.S. Kaeley, MBBS, MRCP, ROAD Clinic, Clinical Associate Professor of Medicine, University of Washington; W.P. Maksymowych, MD, FRCPC, FACP, FRCP(UK), Professor of Medicine, Consultant Rheumatologist, Senior Scholar Alberta Heritage Foundation for Medical Research, Department of Medicine, University of Alberta; R.W. Moskowitz, MD, Professor of Medicine, Case Western Reserve University, Division of Rheumatic Diseases, University Hospitals; H.R. Schumacher Jr., MD, Professor of Medicine, University of Pennsylvania VA Medical Center; T.G. Woodworth, MD, Global Clinical Science Leader, Roche Products, Ltd.; P.J. Mease, MD, Seattle Rheumatology Associates, Director, Division of Rheumatology Research, Swedish Medical Center, Clinical Professor, University of Washington School of Medicine.

Address correspondence to Dr. H. Keen. E-mail: H.I.Keen@leeds.ac.uk

Outcome measures for many rheumatology clinical trials evaluate systemic responses to therapies. These validated outcome measures, such as the EULAR disease activity score (DAS) and the American College of Rheumatology (ACR) responder criteria[1,2], are formulated from a composite of global and systemic domains including clinical and patient reported outcomes (PRO). These measures have been invaluable in demonstrating efficacy in randomized clinical trials (RCT) in inflammatory arthritis. However, both in the setting of partially controlled systemic inflammatory arthritis with one or two refractory joints, and in monoarticular or pauciarticular processes [e.g., osteoarthritis (OA), psoriatic arthritis (PsA) and gout] and situations in which an intraarticular (IA) therapy may be contemplated, the utility of these composite outcome measures to detect change is uncertain. Therapeutic advances for the treatment of single joints are likely to be hindered in the absence of validated outcome tools to specifically address single joint responses.

The recognition of the need for validated outcomes to assess responses in single joints led a group of interested individuals to approach the OMERACT Executive in 2004. The expertise of the group included outcome measures (clinical, functional and imaging assessment), rheumatic diseases [RA, OA, ankylosing spondylitis (AS), PsA, and gout], imaging modalities [magnetic resonance imaging (MRI), ultrasonography (US)], biomarkers, and tissue analysis. This group participated in OMERACT 8 as a spe-

cial interest group (SIG), with a longterm objective of: (a) determining the domains of importance in single joint assessment and (b) initially assessing and further developing outcome measures for the knee joint to be evaluated through the OMERACT filter (truth, discrimination, and feasibility)[3]. At OMERACT 8 a comprehensive literature review of clinical and PRO was presented. This review demonstrated that few existing tools met the OMERACT filter across different disease conditions[4]. During the SIG, a research agenda was developed to evaluate instruments used in measuring response to IA steroid injection in RA or knee OA, to pilot clinical and PRO measures in an IA gene therapy RCT in RA, PsA or AS, and to examine databases of systemic therapy RCT to determine the impact of single joint changes on composite outcomes[4].

Since OMERACT 8 and in preparation for OMERACT 9, working groups initiated several studies to evaluate various aspects of clinical, functional, and imaging outcomes in single joints in response to either IA or systemic therapy. Several additional participants were recruited to contribute their expertise to the working group. An updated literature review was conducted by a fellow (HIK) to evaluate the performance of imaging modalities in detecting changes in single joints. Regular teleconferences were conducted to coordinate and update various aspects of the SIG.

The purpose of the SIG at OMERACT 9 was to: (1) Update and expand the literature review in regards to imaging outcomes to assess response to therapy in knee arthritis, using the OMERACT filter; (2) Review results and design of studies of IA and systemic therapies targeting single joints with a focus on the performance of outcome instruments; and (3) Develop a research agenda to determine a core set of domains for single joint assessment in the context of both inflammatory and degenerative arthritis and move toward their validation.

LITERATURE UPDATE

Literature review (presented by H.I. Keen, G.S. Kaeley). A comprehensive literature review was conducted to identify studies that used imaging modalities to measure responses to therapies in adults with knee joint arthritis. The aim was to identify published studies focusing on the knee joint in which imaging outcome measures were examined. These results were then analyzed in context of the OMERACT filter. The review evaluated studies using ultrasound (US), MRI, and scintigraphy. There was great variability in image acquisition methods and variables within each imaging modality. A variety of anatomical structures were examined in these studies with variable definitions of observed pathologies. Even among studies employing the same imaging modality, the scoring systems used to define outcomes were not uniform. Although some studies examined construct validity of the imaging modality used, there was little evidence of criterion validity (e.g., comparison to direct macroscopic or microscopic visualization of the pathology). Discrimination and feasibility were rarely considered in the published manuscripts, and information regarding the reliability of imaging measurements was also lacking in the studies reviewed. These studies generally examined small cohorts of subjects, and focused on synovium, with little examination of other joint structures.

In summary, the literature review demonstrated a paucity of data regarding the extent to which imaging modalities fulfill the OMERACT filter with regards to truth, discrimination, and feasibility in assessing responses at the knee joint. The results of the detailed and comprehensive review of this literature will be published separately.

REVIEW OF STUDIES

Study 1 (H.I. Keen). The first study (manuscript in preparation) aimed to develop a model using US to assess short-term synovial response to IA steroids *in vivo* in subjects with knee OA. Subjects with knee OA had clinical and US assessments at baseline and followup. Those with a clinical need had an IA steroid injection at baseline. The study utilized US to examine multiple pathologies in multiple regions in the knee. The data were then examined to determine which pathologies and regions best discriminated between groups and time.

Ultrasonography was able to differentiate between the treatment group and control group over time, while clinical examination was not. Based on these results, a preliminary US scoring system for knee synovitis will be proposed that may require further examination using the OMERACT filter.

Study 2 (P.J. Mease). At OMERACT 8 in the single joint SIG, a phase 1 study of an intraarticular gene transfer agent was presented. The study involved the gene for an anti-tumor necrosis factor (TNF) soluble receptor protein combined with an adeno-associated virus vector, which was injected into the knees or ankles of subjects with active RA or AS. The results suggested that clinical examination was not necessarily adequate to distinguish the effects of active therapy from placebo[5]. A phase 1/2 study had been initiated to study the efficacy of this agent with or without background systemic anti-TNF therapy in patients with RA, PsA, and AS. In addition to the primary safety and efficacy outcomes conducted in the initial 60 patients, special study procedures were added to an expanded second phase in an additional 60 patients focusing on face and content validity, interobserver agreement of target joint swelling and tenderness, sensitivity to change, and feasibility using the OMERACT filter. Assessments included PRO [target joint global assessment visual analog scale (VAS), target joint function VAS, target joint pain on a Likert scale of 1–5, modified Rheumatoid and Arthritis Outcome Score and Health Assessment Questionnaire (HAQ)] and joint assessments performed by 2 independent assessors, scored for swelling and tenderness, each on a semiquantitative graded scale of

0–3. MRI scans were performed on 10 subjects at baseline and at week 12.

The primary efficacy outcomes included prespecified 30% reductions in patient global assessment, target joint global assessment, and target joint function, and changes in target joint pain. These newly evaluated PRO of their index joint demonstrated good correlation with other validated patient assessments of disease status (HAQ, RAOS) and correlated highly with each other. The physician assessments of the target joint demonstrated very high interobserver agreement with near complete agreement for both joint tenderness and joint swelling. The smallest detectable difference for the clinical joint assessments, however, was wider than for the PRO. In summary, our study highlighted the discordance among different methods of assessing single joints using this gene therapy construct and the need for additional formative research to inform future studies. The complete results of this study will be published as a separate manuscript.

Studies 3 and 4 (W.P. Maksymowych). The protocol and progress of a currently recruiting, prospective open label study of IA yttrium-90 (yt90) injection in refractory knee arthritis was presented. Evaluations include expanded target joint clinical assessments, and patient and physician reported outcomes and correlative MRI. The design and progress of a second study evaluating the efficacy and safety of systemically administered adalimumab in treating inflammatory OA of the knee was also discussed. As with the yt90 study, similar outcomes are being assessed in this study.

Update on the OMERACT gout working group's proposed domains and tools to assess the knee (H.R. Schumacher). Because gout often presents as a monoarticular process for which domains have been developed for RCT, we reviewed the ongoing work from the OMERACT gout group concerning outcomes in this disease. The domains that have been proposed for assessing acute gout are pain, joint swelling, joint tenderness, patient global, physician global, and function. The proposed domains for assessing chronic gout are serum urate, flares, tophi, health related quality of life, function, pain, patient global, physician global, work disability, and joint inflammation[6,7]. Other domains proposed as discretionary or needing research include acute phase markers and joint range of motion. Those for chronic gout include joint damage by imaging, health care utilization, and synovial fluid crystals or leukocyte counts.

Discussion on the assessment of central sensitization in knee arthritis (L.A. Bradley). There is increasing evidence that the experience of arthritis pain can be significantly influenced by central nervous system mechanisms. Individual differences in function of endogenous pain inhibition systems may be especially important in understanding and possibly reducing ethnic disparities in pain responses and management of pain. For example, African-American patients with knee OA and RA report higher pain levels[8] than their non-Hispanic, White counterparts. Similarly, African-American patients in chronic pain management programs display lower pain tolerance levels in response to experimental stimuli[9]. Studies of healthy adults consistently show that higher levels of pain sensitivity and lower pain tolerance levels among African-American, relative to non-Hispanic White, individuals are associated with impairment of endogenous, non-opioid and opioid pain inhibitory functions among the African-Americans[10]). Until recently, however, ethnic differences in pain inhibitory functions have not been evaluated among patients with arthritis.

A new, US National Institutes of Health supported protocol assessing ethnic differences in opioid pain inhibitory function in African-American and non-Hispanic White patients with knee OA was presented. The study aims to evaluate ethnic differences in: (a) ratings of pain unpleasantness and intensity evoked by noxious heat stimuli delivered to their affected knees, and (b) a ligand positron emission tomography imaging measure of change in opioid receptor binding potential in brain areas involved in pain inhibition that is evoked by exposure to the heat stimuli. Additional analyses will assess whether the anticipated ethnic group differences in change in opioid receptor binding potential partially mediates the expected group difference in pain unpleasantness ratings after controlling for demographic, socioeconomic, and disease severity measures. Advances in understanding central nervous system mechanisms that contribute to individual differences in pain responses and pain inhibition may lead to improved understanding of symptoms arising from arthritis in the single joint as well as better interventions and assessments of outcomes.

RESEARCH AGENDA

Breakout groups (Moderated by Drs. Bingham and Mease). After the introductory presentations, 2 smaller breakout groups were assembled to facilitate additional discussion regarding single joint assessment. These groups were asked to identify areas to focus future research, including expanding and evaluating additions, outcome domains and their applications. Unanimous consensus was obtained that the development of outcome measures for interventions in single joints was needed. The group agreed that continued focus on the knee was appropriate with the recognition that, for other joints, similar exercises would ultimately be required.

There was some disagreement within the breakouts regarding the focus for the single joint group. While most individuals acknowledged the unique characteristics of different diseases and the potential need for discrete domains in a study of inflammatory arthritis versus OA versus crystalline arthritis, a smaller group of individuals proposed developing a new evaluation construct in which the joint could be assessed in terms of its component parts

(synovium, cartilage, bone, tendons, and entheses). Others felt that a generic outcome core set for single joints may lack discriminant validity across a range of diseases given that pathological processes and therapies are not uniform. Regarding gout, a domain set for acute versus chronic disease with overlapping and unique domains was raised as an example of the need to potentially develop a central core set with other domains to be used for a particular disease. Many participants recognized the need for additional work in both inflammatory arthritis and OA. However, most acknowledged the need to initially focus the efforts of the single joint group on inflammatory arthritis as this represents the greatest unmet need for development of validated outcomes for ongoing and upcoming interventional studies, since in OA, outcomes have been well established and validated.

Most participants agreed that patient-reported outcomes were important to evaluate further, particularly responsiveness of the outcome tool to changes resulting from effective therapies. Both groups discussed imaging and its role in joint assessment. It was pointed out that while imaging has advanced the ability to examine the different components of the joint organ and may be able to demonstrate change over time at a group level, at the individual level this may not be the case when correlated with clinical and patient reported outcomes. Additional work was needed to better define construct validity, sensitivity to change, and reliability. There was widespread agreement that many confounders were important to consider when evaluating a response in a single joint, especially the co-occurrence of 2 processes in the same joint (e.g., OA and RA, RA and FM).

The discussions from the group helped to develop a larger domain list for single joint assessment. These included: pain, tenderness, swelling, range of motion, physical function, patient global assessment of target joint, physician global assessment of target joint, imaging assessment, evaluation of cartilage, bone, and synovial compartments (potentially dependent on the particular therapy being investigated), soluble biomarkers, and histology. While the data thus far presented from the literature reviews and from the IA steroid and IA gene therapy study raise questions as to the reliability of the clinical examination, many participants continued to view some physical assessment of the joint as relevant and necessary in determining the face validity of an outcome measure.

A summary of activities of the SIG and the discordant areas in the breakouts were presented back to the larger OMERACT plenary. In final group voting, a majority of participants (54%) agreed that the focus of the single joint assessment group should cover multiple forms of arthritis. There were differing opinions regarding focus on particular therapeutic interventions with 28% recommending a focus on IA therapies, 7% evaluating the effects of systemic therapies on single joints, but a near majority (45%) felt that both IA and systemic therapies were appropriate to evaluate for single joint responses. The lack of clear consensus from the larger group on these questions was reflective of the differing opinions in the breakout groups.

Reviewers of this article have raised important additional issues that will need to be considered in future discussions of the working group. Occasionally, IA therapy may have effects beyond the injected joint, both positive and negative. For example, if a large quantity of IA corticosteroid is administered, it may have a systemic effect of reducing joint disease activity in non-injected joints as well as causing adverse effects such as elevation of serum glucose levels. Novel IA therapies similarly may have systemic effects that are not predictable. These potential systemic effects need to be taken into account when one designs and recommends outcome measures for single joint assessment. A further point is that change in disease activity in a single joint may variably impact pain referral patterns that may, in turn, variably affect patient reported outcomes. This needs to be evaluated in future studies of measures of single joints.

In conclusion, based on the discussions at OMERACT 9, it is clear that developing outcome measures for the evaluation of single joints remains an important but difficult endeavor, with few validated outcomes that are appropriate across diseases and therapeutic modalities. The results presented from an IA steroid study in knee OA using US, clinical examination, and PRO as outcomes, and a study of an IA gene therapy with clinical examination, PRO, and MRI both confirm the relative lack of sensitivity of clinical examination in detecting changes over time using current tools. Additional formative research will be required to determine the specific domains that pass the OMERACT filter of truth, discrimination, and feasibility in different contexts. In developing outcomes for single joint assessment, it is important to consider concomitant inflammatory and structural disease, the need for standardization of imaging outcome definitions, and to recognize the lack of a gold standard to establish the "truth" of an effective intervention. Ongoing clinical studies will provide additional important correlative information about clinical, patient reported, and imaging outcomes. It is anticipated that these additional projects will allow the group to move forward with the development of a suggested master set of domains for single joint assessment potentially with different subsets.

ACKNOWLEDGMENT

We thank Targeted Genetics and Kristin Seymour for support of group communication. OMERACT Single Joint Assessment Working Group: Pervin Anklesaria, Scott Baumgartner, Clifton O. Bingham III, Laurence A. Bradley, Barry Bresnihan, Philip G. Conaghan, Hani El-Gabalawy, Jon T. Giles, Alison E. Heald, Gurjit S. Kaeley, Helen I. Keen, Jean-Francis Maillefert, Philip J. Mease, Roland W. Moskowitz, H. Ralph Schumacher, Linda Sloma, Lee Simon, Richard Wakefield, and Thasia G. Woodworth

REFERENCES

1. van Riel PL, van Gestel AM, van de Putte LB. Development and validation of response criteria in rheumatoid arthritis: steps towards

an international consensus on prognostic markers. Br J Rheumatol 1996;35:4-7.
2. Felson DT, Anderson JJ, Boers M, et al. American College of Rheumatology. Preliminary definition of improvement in rheumatoid arthritis. Arthritis Rheum 1995;38:727-35.
3. Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for outcome measures in rheumatology. J Rheumatol 1998;25:198-9.
4. Giles JT, Mease P, Boers M, et al. Assessing single joints in arthritis clinical trials. J Rheumatol 2007;34:641-7.
5. Mease PJ, Hobbs K, Chalmers A, et al. Local delivery of a recombinant adeno-associated vector containing a tumor necrosis factor-alpha antagonist gene in inflammatory arthritis: a phase 1 dose-escalation safety and tolerability study. Ann Rheum Dis 2008;Aug 4 E-pub.
6. Schumacher HR, Taylor W, Edwards L, et al. Outcome domains for studies of acute and chronic gout. J Rheumatol 2009;36: [in press].
7. Grainger R, Taylor W, Dalbeth N, et al. Progress in measurement instruments for acute and chronic gout studies. J Rheumatol 2009;36: [in press].
8. Bradley LA, Kersh BC, DeBerry JJ, Deutsch G, Alarcon GA, McLain DA. Lessons from fibromyalgia: abnormal pain sensitivity in knee osteoarthritis. Novartis Foundation Symp 2004;260:258-70.
9. Edwards RR, Doleys DM, Fillingim RB, Lowery D. Ethnic differences in pain tolerance: clinical implications in a chronic pain population. Psychosom Med 2001;63:316-23.
10. Campbell CM, France CR, Robinson ME, Logan HC, Geffken GR, Fillingim RB. Ethnic differences in diffuse noxious inhibitory controls. J Pain 2008;9:759-66.