MR. MARTIJN ANTONIUS HENDRIKUS OUDE VOSHAAR (Orcid ID : 0000-0003-2438-9973)

DR. JEFFREY R CURTIS (Orcid ID : 0000-0002-8907-8976)

DR. YING YING LEUNG (Orcid ID : 0000-0001-8492-6342)

DR. ENRIQUE ROBERTO SORIANO (Orcid ID : 0000-0003-3143-1084)

DR. WILLIAM TAYLOR (Orcid ID : 0000-0001-6075-8479)

Article type : Original Article

The International Consortium for Health Outcome Measurement (ICHOM)Set of Outcomes that Matter to People Living with Inflammatory Arthritis

Consensus from an international Working Group

Martijn A.H. Oude Voshaar[1], PhD; Zofia Das Gupta[2], PhD; Johannes W.J. Bijlsma[3], PhD & MD; Annelies Boonen[4], PhD & MD; Jeffrey Chau[5], BA; Delphine S. Courvoisier[6], PhD; Jeffrey R. Curtis[7],PhD & MD; Benjamin Ellis[8], Msc & MD; Sofia Ernestam[9], PhD & MD; Laure Gossec[10], PhD & MD; Christine Hale[11], MD; Jennifer Hornjeff[12] ,PhD; Katy Y.Y. Leung[13], PhD & MD; Merav Lidar[14], PhD & MD; Phillip Mease[15], PhD & MD; Kaleb Michaud[16], PhD & MD; Girish M. Mody[17], PhD & MD; Mwidimi Ndosi[18], PhD; Christina H. Opava[19], PhD; Geraldo R.C. Pinheiro[20], PhD & MD; Matthew Salt[2] , BSc; Enrique R. Soriano[21], PhD & MD; William J. Taylor[22], PhD & MD; Maria J.H. Voshaar[1], MSc; Angelique E.A.M. Weel[23], PhD & MD; Maarten de Wit[24], PhD; Nico Wulffraat[25], PhD & MD; Mart A.F.J. van de Laar[1,2], PhD & MD; Harald E. Vonkeman[1,2] PhD & MD;

1. University of Twente, Department of Psychology, Health and Technology, Enschede, The Netherlands
2. International Consortium for Health Outcomes Measurement, London, United Kingdom
3. University Medical Center Utrecht Utrecht, Department of Rheumatology and Clinical Immunology, Utrecht, The Netherlands
4. Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Centre+, Maastricht, Netherlands
5. Hong Kong Psoriatic Arthritis Association, Hong Kong, China
6. University Hospitals of Geneva & University of Geneva, Switzerland
7. University of Alabama at Birmingham, Birmingham AL
8. Imperial College Healthcare NHS Trust
9. Clinical epidemiological unit, Dept. of Medicin Solna, Karolinska Institute
10. Sorbonne Université, Paris France; Pitié Salpêtrière hospital, APHP, Rheumatology department, Paris, France.
11. Lockton Dunning Benefits, Dallas
12. Columbia University Medical Centre
13. Singapore General Hospital, Duke-NUS Medical School, Singapore
14. Sheba Medical Centre, Department of internal medicine, Tel-Hashomer, Israel
15. Providence-St. Joseph Health System; University of Washington, Seattle, WA
16. University of Nebraska Medical Center Omaha, NE & The National Databank for Rheumatic Diseases, Wichita, KS
17. Nelson R Mandela School of Medicine, University of KwaZulu-Natal, South Africa
18. Department of Nursing and Midwifery, University of the West of England, Bristol, UK
19. Karolinska Institutet, Dept. pf Neurobiology, Care Sciences and Society, Division of Physiotherapy
20. Department of Internal Medicine, Division of Rheumatology, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brasil
21. Instituto Universitario Hospital Italiano de Buenos Aires

22. Department of Medicine, University of Otago, Wellington, New Zealand
23. Maasstad Hospital, Department of Rheumatology, Rotterdam, The Netherlands
24. VU University Medical Centre, VU University Medical Centre, Dept. Medical Humanities, Amsterdam Public Health (APH), Amsterdam, Netherlands
25. Wilhelmina Children's Hospital, Utrecht, Netherlands

Financial Support: Arthritis Research UK; Karolinska Institutet ;Maasstad Hospital; Santeon; Sheba Academic Medical Center Hospital; Transparency in Healthcare

Address correspondence and reprint requests to M.A.H. Oude Voshaar , Department of Psychology, Health and Technology, University of Twente, PO BOX 50 000; 7500 KA Enschede, the Netherlands
A.H.Oudevoshaar@utwente.nl

Background

The implementation of value based healthcare (VBHC) in inflammatory arthritis (IA) requires a standardized set of modifiable outcomes and risk-adjustment variables that is feasible to implement worldwide.

Methods

The International Consortium for Health Outcomes Measurement (ICHOM) assembled a multidisciplinary working group, consisting of 24 experts from six continents, including six patient representatives, to develop a Standard Set of outcomes for IA. The process followed a structured approach using a modified Delphi process to reach consensus on 1) conditions covered by the set, 2) outcome domains, 3) outcome measures, 4) risk-adjustment variables. Consensus on decision areas two to four were supported by systematic literature reviews and consultation of experts.

Results

The ICHOM IA Standard Set covers patients with rheumatoid arthritis (RA), axial spondyloarthritis, psoriatic arthritis and juvenile idiopathic arthritis (JIA). We recommend that the following outcomes be collected at least annually: pain, fatigue, activity limitations, overall physical and mental health impact, work/school/housework ability and productivity, disease activity, and serious adverse events. Validated

measures for patient-reported outcomes were endorsed, and linked to common reporting metrics. Age, sex at birth, educational level, smoking status, comorbidities, time since diagnosis, and rheumatoid factor and anti-citrullinated protein antibody lab testing for RA and JIA should be collected as risk-adjustment variables.

Conclusion

We present the ICHOM IA Standard Set of outcomes that enables healthcare providers to implement the value based healthcare framework and enable comparison of outcomes important to patients with IA.

Significance & Innovation

- Standards for measuring and comparing treatment outcomes that matter to patients with inflammatory arthritis that are globally implementable are currently lacking
- We used a modified Delphi procedure and systematic reviews of the literature to develop a Standard Set of Outcomes that matter to patient with inflammatory arthritis
- The patient reported outcome measures we recommend for measuring pain, activity limitations, fatigue and assessment of overall emotional and physical health impact were linked to a common Item response theory based common metric, so that users of the set can select their preferred instrument for measuring these outcomes

Introduction

The inflammatory arthritides are a group of systemic, immune mediated rheumatic conditions, characterized by synovitis or inflammation of periarticular tissues and joint damage. The life time risk of adult onset inflammatory arthritis (IA) has been estimated to be about 6% of the population in the United States (1). The availability of strategies to diagnose the diseases earlier,

and the availability of biological and targeted small molecule therapies, in combination with early, tight controlled treatment strategies have led to relevant improvements in outcomes for many patients over the last decades (2,3). However, these improvements have also resulted in an increased financial burden on healthcare systems (4,5).

The prevalence and case recognition of IA is expected to rise further over the next decade, particularly in less economically developed countries (6). Hence it will be increasingly important to optimize care and allocate available resources efficiently to improve or maintain quality of care. Value Based Healthcare (VBHC) is a framework for improving the quality and efficiency of healthcare, in which improving value for the patient is the central concept (7). Value is defined as the patient outcomes achieved, relative to financial costs for obtaining those outcomes. Within this framework, value can be increased by improving patient outcomes or by delivering the same outcomes at a lower cost. Public reporting of patient outcomes by healthcare providers is proposed as a mechanism that will accelerate identification and adoption of high value care, through shared learning and promoting benchmarking of outcomes that matter to patients.

In order for outcomes to be comparable between different healthcare providers, exact definitions for each relevant outcome are required. The outcomes need to be feasible to collected in a variety of healthcare systems and a set of relevant risk-adjustment variables should be included to ensure risk-adjusted comparisons of outcomes between providers that serve different patient populations. The International Consortium for Health Outcomes Measurement (ICHOM) initiative is working toward the global implementation of VBHC by developing Standard Sets of patient outcomes for a range of medical conditions (8). These standards are intended to be implemented in routine clinical practice and therefore complement earlier core sets and reporting standards intended for clinical research, including the work of the Outcome Measures in Rheumatology (OMERACT) group (10).

The ICHOM process is grounded in a conceptual framework which distinguishes three hierarchically ordered tiers of outcome: health status achieved/retained, the process of recovery and sustainability of health (9). To select the most relevant outcomes, outcome measures and risk adjustment variable for particular conditions, various stakeholders including patients, physicians, policy makers and outcome experts are engaged in a consensus building process that is supported by a systematic evaluation of the available evidence base, including critical evaluation of available instruments and evidence supporting their measurement properties.

To further encourage the adoption and implementation of VBHC in rheumatology, we set out to develop a globally applicable set of outcome measures that reflect outcomes that matter to patients with IA, for providers to track in their clinical practice.

Methods

Working group

A working group of outcome experts (n = 24) was convened by ICHOM. Working group members were carefully selected to ensure representation of relevant professional disciplines, different geographic areas, and the patient perspective. The working group included patient representatives (n = 6), registry leaders, and members with a professional background in: adult and pediatric rheumatology; nursing; epidemiology; psychology; rehabilitation medicine; physiotherapy; and psychometrics. Working group members of all 6 inhabited continents were included. The efforts of the working group were guided and facilitated by a core project team.

Working group process

A modified Delphi approach was used that has been developed by ICHOM and was previously applied by ICHOM to develop standards for a number of other conditions (10,11,20–26,12–19) The process involved reaching consensus in four major decision areas 1) which IA conditions

and treatments to include in the Standard Set, 2) a minimally sufficient set of outcomes relevant for each of the conditions, 3) standardized definitions and time points for assessing these outcomes, and 4) standardized definitions for risk-adjustment variables to ensure fair comparisons between healthcare providers who wish to implement the set. Each decision area was discussed during a dedicated videoconference. A list of potentially relevant items (i.e. conditions/domains/time points/risk-adjustment variables) to be included in the Standard Set, along with supporting evidence (see below), was prepared by the project team preceding each meeting. These items were identified in a series of systematic literature reviews and/or consultation with external experts on the topic under consideration. A summary of the preparatory work was provided to working group members preceding each videoconference. During each meeting, the items were discussed and expanded on or revised based on the input of the working group. Following each videoconference, the project team circulated a summary of the discussions and a survey asking working group members to provide feedback and vote on each item considered for inclusion in the Standard Set. For voting during the final survey, a 9-point Likert scale ranging from 1 – "not recommended" to 9 – "essential to have" was used. Items were included in the Standard Set if rated ≥ 7 by at least 70% of the working group or excluded if rated ≤ 3 by at least 70% of the working group. In other cases, the result was considered inconclusive and the item was discussed again during the following videoconference. The conduct and reporting of the Delphi process followed reporting guidelines for core outcome set development using the Delphi approach (see supplemental file 1) (27).

Pre-selection of relevant patient outcome domains

Preceding the videoconference on 'selection of outcome domains', two separate systematic literature reviews were performed in December 2016 in the PubMed database to identify outcomes relevant to patients for the included conditions, that were modifiable in principle and feasible to implement. In the first search we used the "Cochrane Highly Sensitive Search Strategy

for identifying randomized trials" (28) to identify 25 recent randomized trials in each of the conditions included in the Standard Set. For each randomized trial, we checked the relevant online repositories and conducted a second PubMed search using the name of each trial and trial registration number to find additional publications on the same study. All outcomes assessed in each trial were extracted. Reports on randomized trials in languages other than English were excluded. In the second search we identified reports on qualitative studies in which patients with one of the relevant conditions were asked about the most important outcomes of their disease. We included only papers in which an open-ended question format was used, to prevent investigator-imposed biases on the list of patient-generated outcome domains. All outcome domains considered important by patients were extracted from each paper by two reviewers independently. Disagreements were resolved during a consensus meeting with a third referee present. Previously published core set recommendations for outcome measurement in randomized trials were also consulted, as was the EULAR standardized dataset for observational research (29–33). Finally, two patient advisory group sessions with IA patients from the Netherlands and the USA were organized by the project team to serve as a check on the comprehensiveness of the list of identified patient outcomes. The patient advisory group protocol was exempt from ethical review by the Chesapeake Institutional Review Board.

Pre-selection of Outcome measures

All outcome measures used in any randomized trial identified in the initial systematic review, recommended for inclusion by working group members or previously endorsed by relevant consensus statements, were considered for inclusion. The instruments were reviewed with respect to: outcome domains, evidence supporting psychometric properties, feasibility, licensing fees, and degree to which they were established in the field. To support this process, a systematic literature review was performed in May 2017 to identify papers reporting on the measurement properties of 26 potentially relevant Patient-reported outcome measures (PROMs). The

methodological quality of the 159 identified papers was assessed using the COSMIN checklist (34). The studies that were of high methodological quality were then used to rate the measurement properties of the 26 PROMs, using quality criteria proposed by Terwee et al. and ISOQOL (35,36). Understandability, cost and time to complete were all assessed to determine the feasibility of implementing specific PROMs. The Flesch-Kincaid grade level was calculated for the English language version of each PROM (37), information about fees payable for use of the instrument was retrieved from the copyright owner's website when applicable, and information on time to complete was retrieved from a previous series of reviews (38).

Pre-selection of risk-adjustment variables

A preliminary list of risk-adjustment variables was extracted from published reviews on risk factors and validated risk models. Previously published ICHOM Standard Sets were reviewed for definitions of demographic and socioeconomic variables.

External validation by health professionals and patient experts

After proceeding through all the Delphi rounds, the preliminary Standard Set was made available and sent to various stakeholders for review via www.ichom.org. A patient survey was distributed via national patient organizations and the networks of the project team and working group members. Patients were asked to rate the importance of each outcome using a 9-point Likert scale ranging from "1 = not at all relevant" to "9 = essential" and were given the opportunity to suggest additional outcomes. Health professionals were recruited via the professional networks of the working group members and project team. Health professionals were asked to rate the relevance of each domain, provide feedback on feasibility of implementation of the Standard Set in clinical practice, and rate the appropriateness of the risk-adjustment variables and time points for assessment.

Results

Scope

At the start, it was recognized by the working group that the same treatment goals and longitudinal outcomes (pain, physical function, fatigue) are relevant to most types of IA. The ICHOM IA Standard Set was therefore designed to evaluate treatment outcomes of patients with rheumatoid arthritis (RA), axial spondyloarthritis (axSpA) and psoriatic arthritis (PsA), as well as juvenile idiopathic arthritis (JIA) and applies to all treatments for these conditions, including medication, surgery, and physical and occupational therapies. All working group members voted to include RA, axSpA and PsA, and 82% voted to include JIA. The inclusion of gout (59% voted to include initially), connective tissue diseases (41%) and vasculitis (36%) was also considered, especially as few outcome recommendations are available for the latter two conditions. However, after revisiting this topic at a subsequent meeting it was decided that their inclusion might lead to a generic set of outcomes, which might insufficiently capture the outcomes that matter to patients with individual conditions due to the variety of disease manifestations and disease courses.

Outcome domains

Twenty-four outcome domains were initially identified in the 130 randomized trial reports and 28 qualitative studies that were identified in the systematic literature reviews (references are available on request from the corresponding author). This list was expanded upon and refined several times based on group discussions with working group members. The final consolidated list of outcomes is presented in table 1, together with a summary of both systematic reviews and the final voting results. The list of outcome domains assessed in randomized trials and their rank ordering reflects a preference in trials for clinical measures of disease manifestations and patient-reported outcomes of symptoms and their direct impact on functioning. The list and rank

ordering of patient-generated outcome domains on the other hand somewhat deemphasized the importance of outcome measures that reflect the pathophysiology of the specific disease and included a wider variety of outcomes that reflect the different ways arthritis impacts the daily lives of patients. Besides PROMs of symptoms and basic functioning, the patient-generated list also included more generic outcomes such as overall psychological wellbeing and participation restrictions. To characterize the core symptoms and their direct effect on functioning from the patient perspective, the working group recommends that providers assess pain, fatigue, and activity limitations (i.e. physical function). These were the most frequently used outcome domains in randomized trials and were reported as important by patients in almost all of the reviewed qualitative studies. They were also the most frequently endorsed domains in the individual core sets for randomized trials of the respective conditions. To assess impact of IA on the daily lives of patients more broadly, the working group recommends an assessment of overall emotional and physical health impact, and work/school/housework productivity. Participation restrictions other than work or school productivity were also considered important. However, this domain was eventually excluded because of significant overlap with other included domains and because experience with available measurement instruments is currently limited.

Assessments of inflammatory disease activity and therapeutic response are further recommended as measures of disease control, because the absence of signs and symptoms of disease is the primary treatment target for IA and disease activity was considered the main determinant of overall impact of disease. Finally, adverse events should be recorded as a measure of disutility of care.

Outcome measures

The list of recommended patient-reported outcome measures (PROMs) is presented in table 2. Supplemental file 2 provides an overview of characteristics and ratings assigned to the measurement properties of these PROMs and includes an overview of the criteria used for assigning ratings.

The Working Group agreed that a key property was feasibility in different settings globally. Therefore, instruments with a large number of items, or that cannot be hand-scored were avoided. All endorsed PROMs are available in multiple languages and for each outcome domain, at least one PROM is recommended that was judged to have sufficient evidence supporting its measurement properties, based on the results of the systematic review. On the other hand, some instruments were included that do not (yet) meet psychometric standards of the COSMIN checklist. Several PROMIS measures were included so that experience with innovative Item Response Theory (IRT) based measures can accumulate. The rheumatoid- and psoriatic-arthritis impact of disease scores (RAID and PSAID) and ASAS Health index were recommended because these are patient/ICF -derived composite scores assessing the important domains of impact of RA and PSA, and AxSpA respectively. Supplemental file 3 provides an overview of the various measures that are recommended by the clinical guidelines issued by various national and international rheumatology societies. For each of the ICHOM outcome domains, the endorsed outcome measures are congruent with the various guidelines. Users of the Standard Set may select preferred instruments to assess each outcome from the list of endorsed PROMs presented in table 2. The shortest recommended combination of PROMs to assess all outcome domains is NRS/VAS to measure fatigue, overall emotional and physical health impact, and pain; the Work Productivity and Activity Impairment questionnaire (WPAI) to measure work/school/housework ability and productivity, and the Health Assessment Questionnaire II (HAQ-II) to measure activity limitations. This combination of PROMs is free

to use for all users and totals 19 questions, which most patients should be able to fill out in 5 minutes. The endorsed PROMs of pain, fatigue, overall wellbeing and activity limitations have been linked to a common reporting metric, such that outcomes can be compared between providers that have used different instruments (39). Linked scores for benchmarking outcomes using any of the recommended instruments of the Standard Set can be obtained from www.tihealthcare.nl/en/expertise/common-metrics.

To track disease activity and therapeutic response, it is proposed that major evidence-based guidelines are followed (40,41). Patients and healthcare providers should specify target disease activity levels for individual patients (preferably remission; if not, feasible low disease activity) and assess at each visit whether or not this target was achieved. Disease activity should be monitored using a validated and internationally recognized clinical composite score.

Risk-adjustment variables

Most risk-adjustment variables included in the set (table 3) apply to all patients and care was taken to include risk-adjustment variables that are relevant and applicable in a variety of healthcare systems. Year of birth and sex were included as demographic variables. Educational level was ultimately chosen as the only indirect measure of social economic status (SES). Other SES-related variables were considered important, but difficult to collect, due to restrictions on recording race/ethnicity in some countries, area-based measure of SES possibly not being available for each country, and patients potentially feeling reluctant to report on their income/wealth. For baseline status indicators, we included smoking status, comorbidities, diagnosis, time since diagnosis, and rheumatoid factor and ACPA for RA and JIA. HLA B27 was excluded snice it is not routinely collected in healthcare system. Comorbidities should be assessed using the Rheumatic Disease Comorbidities Index (42), modified to include central

sensitization to pain (e.g. fibromyalgia) and obesity. To avoid misclassification of early symptoms that may or may not reflect those specific to the IA diagnosis of interest we elected to include time since diagnosis rather than time since symptom onset.

Data collection time points

To allow meaningful outcomes comparisons between healthcare providers, we recommend all risk-adjustment variables to be collected at the first assessment. All PROMs and clinical measures should also be collected at the first assessment and annually thereafter. In instances of active disease, we recommend that the patient's disease activity status is recorded again at least every six months, but likely more frequently, at the discretion of the patient and their healthcare provider. Adverse events should be collected at each assessment point after baseline. A reference guide with detailed instructions for implementation and exact definitions for all the data elements can be downloaded from www.ICHOM.org. Finally, we stress that these recommendations are intended only for quality improvement purposes and should not be understood as more than minimally acceptable clinical guidelines in patients with established disease. Especially in patients with early disease, more frequent monitoring may be required.

Open review

Eighty-three healthcare professionals, the majority of which (95%) were clinician/researchers, and 630 people living with IA from the United States, France, Argentina, the Netherlands and Brazil reviewed the Standard Set. All outcomes included in the ICHOM IA set were considered very relevant by patients and healthcare professionals (figure 1). Similar to the results of the systematic reviews for identifying outcome domains, patients considered clinical measures slightly less relevant compared with the patient-reported outcomes. A large majority of patients (81.3%) felt that the set comprehensively covers all the relevant outcome domains of their

disease. The healthcare professionals predominantly shared this view. Only three outcomes were suggested to be missing by more than 1 reviewer: financial impact (n=2), joint damage (n=2) and patient satisfaction (n=3). Psychological wellbeing (12.6%) and participation restrictions (5.4%) were the only outcomes that were reported as missing from the set by >2% of patient reviewers (supplemental file 2). The included risk-adjustment variables were rated very relevant by 91.8% of health professionals.

Discussion

We present a Standard Set of outcomes for IA that healthcare providers worldwide can use in routine clinical care to help quantify the value provided for patients in different centers, countries and healthcare systems. This Standard Set was developed through consensus of an international working group with expertise across a range of disciplines relevant to outcome assessment and care for patients with IA. We used a multiple methods approach, in which extensive patient input as well as published qualitative and quantitative data were used to develop a minimally sufficient set of outcomes that we believe represents outcomes that matter to patients with IA. We also proposed time points for data collection and relevant risk-adjustment variables to enable comparisons between providers with different patient populations. Feasibility of implementation in different healthcare systems was a central priority. Therefore, we included PROMs that are not only widely accepted measures of the respective domains but that are also available in multiple translations, and for each outcome there is at least one PROM free to use. However, for the use of several instruments: PedsQL, SF-6D, FACIT and EQ-5D license fees may apply. PROMIS has also introduced financial charges for the use of some of their products, including their computerized adaptive tests. Using the PROMIS Assessment Centre platform will incur a $5000 USD charge per study per year.

One of the challenges faced with international standardization of patient outcomes data collection is that a variety of well validated and frequently used PROMs are typically available to assess the different patient outcomes. The working group for the ICHOM Depression & Anxiety and Chronic Kidney Disease Standard Sets previously responded to this by endorsing PROMs that can be mapped to the PROMIS metric, using resources provided by the PROMIS PROsetta project (43,44). This way, users of these sets use one PROM for each domain, but results can be scored on the PROMS metric. In the work on the ICHOM IA Standard Set presented here, this is taken one step further again by linking multiple PROMs to an IRT based common reporting metric (45). This makes it easier for new or ongoing data collection initiatives to contribute their data since it allows users of the ICHOM IA set to choose one instrument from a number of alternatives for each domain. Provided one of the endorsed instruments presented in table 2 is collected, outcomes can be compared with other healthcare providers that use the ICHOM Standard Set. For example, outcomes assessed using the VAS fatigue scale can be directly compared with outcomes of a different group of patients assessed using FACIT fatigue. In principle, PROMs could be added to and removed from the list of endorsed instruments, without affecting comparability of the outcomes. The ICHOM list of recommended PROMS overlaps significantly with current clinical guidelines. Moreover, the results of two systematic reviews of various national RA patients registries show that the majority of the PROMs that are currently collected in the reviewed registries are also included in the ICHOM IA Standard Set. The IRT approach also allows each of the ICHOM IA outcomes to be assessed using computerized adaptive tests, which would help achieve optimally precise scores with minimal numbers of questions (46,47)

Since the Standard Set is intended to reflect outcomes that are important to patients, the extensive input from patients is a strength of this work. We included 6 patient representatives in the working group, derived the list of outcomes from published qualitative studies in which

patients reported outcomes that matter to them, organized two patient advisory group sessions with patients that were not included in the working group to review the final list of outcomes to be voted on by working group members, and 630 patients from various countries reviewed the final version of the Standard Set. The inclusion of working group members with diverse geographic and professional backgrounds is also a strength. We do however acknowledge that different results might have been obtained had other working group members been selected. We also realize that it may prove challenging to collect all the requested information for all healthcare providers at all time points. In particular, inflammatory disease activity may prove logistically challenging to track in some healthcare systems, as it requires clinical assessment of joint involvement and, in some cases, laboratory assessments. In such situations, we would encourage users of the set to at least monitor the PROMs. All patient-reported outcomes can be collected using a minimum of 20 items, which could be further reduced using computerized adaptive testing or targeted short forms. We finally acknowledge that the value of the ICHOM IA Standard Set has not yet been proven in practice. ICHOM aims to partner with a number of interested institutions to pilot-test the Standard Set. Furthermore, a steering committee has been established that will periodically review the Standard Set, including lessons learnt from the pilot phase and other applications of the Standard Set. This will include but will not be limited to reviewing: PROMs that are endorsed in the IA Standard Set PROMS; how easy these PROMS are to access and monitor; and outcomes related to personal goals that individual patients identify.

In summary, we propose a Standard Set of outcomes for patients with IA that providers of care for patients with IA can track to facilitate the global reporting of outcome data and shared learning. A detailed reference guide is available at www.ichom.org.

References

1. Crowson CS, Matteson EL, Myasoedova E, Michet CJ, Ernste FC, Warrington KJ, et al. The lifetime risk of adult-onset rheumatoid arthritis and other inflammatory autoimmune rheumatic diseases. *Arthritis Rheum* 2011;63:633–9. Available at: http://doi.wiley.com/10.1002/art.30155. Accessed January 1, 2018.

2. Stoffer MA, Schoels MM, Smolen JS, Aletaha D, Breedveld FC, Burmester G, et al. Evidence for treating rheumatoid arthritis to target: results of a systematic literature search update. *Ann Rheum Dis* 2016;75:16–22. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4717391&tool=pmcentrez&rendertype=abstract. Accessed February 18, 2016.

3. Schoels MM, Braun J, Dougados M, Emery P, Fitzgerald O, Kavanaugh A, et al. Treating axial and peripheral spondyloarthritis, including psoriatic arthritis, to target: results of a systematic literature search to support an international treat-to-target recommendation in spondyloarthritis. *Ann Rheum Dis* 2014;73:238–242. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23740234. Accessed December 28, 2017.

4. Mody GM, Cardiel MH. Challenges in the management of rheumatoid arthritis in developing countries. *Best Pract Res Clin Rheumatol* 2008;22:621–641.

5. Kumar B. Global health inequities in rheumatology. *Rheumatology* 2017;56:4–5.

6. Cross M, Smith E, Hoy D, Carmona L, Wolfe F, Vos T, et al. The global burden of rheumatoid arthritis: estimates from the global burden of disease 2010 study. *Ann Rheum Dis* 2014;73:1316–22. Available at: http://www.ncbi.nlm.nih.gov/pubmed/24550173. Accessed January 28, 2016.

7. Porter M, Teisberg E. Redefining health care: creating value-based competition on results. 2006. Available at: https://books.google.nl/books?hl=nl&lr=&id=Kp5fCkAzzS8C&oi=fnd&pg=PR10&dq=+Red

efining+health+care:+creating+value-based+competition+on+results&ots=VZv7MiemdE&sig=dSkoYaJd-79YCl9hjH2-OzFFkLU. Accessed October 23, 2015.

8. Porter ME, Ph D, Larsson S, Ph D, Lee TH. Standardizing Patient Outcomes Measurement. *N Engl J Med* 2016;374:10–12. Available at: http://www.nejm.org/doi/full/10.1056/NEJMp1511701?query=featured_home.

9. Porter ME. Measuring health outcomes: the outcomes hierarchy. *N Engl J Med* 2010;363:2477–2481. Available at: http://www.academia.edu/download/30781767/2011-1_Porter_What-is-value-in-health-care_Appendix-2.pdf. Accessed April 16, 2018.

10. Werneke M. A proposed set of metrics for standardized outcome reporting in the management of low back pain. *Acta Orthop* 2016;87:88. Available at: http://www.tandfonline.com/doi/full/10.3109/17453674.2015.1120127. Accessed March 6, 2018.

11. Martin NE, Massey L, Stowell C, Bangma C, Briganti A, Bill-Axelson A, et al. Defining a standard set of patient-centered outcomes for men with localized prostate cancer. *Eur Urol* 2015;67:460–7. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0302283814008458. Accessed March 6, 2018.

12. Mahmud I, Kelley T, Stowell C, Haripriya A, Boman A, Kossler I, et al. A Proposed Minimum Standard Set of Outcome Measures for Cataract Surgery. *JAMA Ophthalmol* 2015;133:1247–52. Available at: http://archopht.jamanetwork.com/article.aspx?doi=10.1001/jamaophthalmol.2015.2810. Accessed March 6, 2018.

13. Allori AC, Kelley T, Meara JG, Albert A, Bonanthaya K, Chapman K, et al. A Standard Set of Outcome Measures for the Comprehensive Appraisal of Cleft Care. *Cleft Palate Craniofac J* 2017;54:540–554. Available at: http://www.cpcjournal.org/doi/10.1597/15-292.

Accessed March 6, 2018.

14. Rodrigues IA, Sprinkhuizen SM, Barthelmes D, Blumenkranz M, Cheung G, Haller J, et al. Defining a Minimum Set of Standardized Patient-centered Outcome Measures for Macular Degeneration. *Am J Ophthalmol* 2016;168:1–12. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0002939416301829. Accessed March 6, 2018.

15. Rolfson O, Wissig S, Maasakkers L van, Stowell C, Ackerman I, Ayers D, et al. Defining an International Standard Set of Outcome Measures for Patients With Hip or Knee Osteoarthritis: Consensus of the International Consortium for Health Outcomes Measurement Hip and Knee Osteoarthritis Working Group. *Arthritis Care Res (Hoboken)* 2016;68:1631–1639. Available at: http://doi.wiley.com/10.1002/acr.22868. Accessed March 6, 2018.

16. Mak KS, Bommel ACM van, Stowell C, Abrahm JL, Baker M, Baldotto CS, et al. Defining a standard set of patient-centred outcomes for lung cancer. *Eur Respir J* 2016;48:852–60. Available at: http://erj.ersjournals.com/lookup/doi/10.1183/13993003.02049-2015. Accessed March 6, 2018.

17. Ong WL, Schouwenburg MG, Bommel ACM van, Stowell C, Allison KH, Benn KE, et al. A Standard Set of Value-Based Patient-Centered Outcomes for Breast Cancer: The International Consortium for Health Outcomes Measurement (ICHOM) Initiative. *JAMA Oncol* 2017;3:677–685. Available at: http://oncology.jamanetwork.com/article.aspx?doi=10.1001/jamaoncol.2016.4851. Accessed March 6, 2018.

18. Roos P de, Bloem BR, Kelley TA, Antonini A, Dodel R, Hagell P, et al. A Consensus Set of Outcomes for Parkinson's Disease from the International Consortium for Health Outcomes Measurement. *J Parkinsons Dis* 2017;7:533–543. Available at: http://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/JPD-161055.

Accessed March 6, 2018.

19. Obbarius A, Maasakkers L van, Baer L, Clark DM, Crocker AG, Beurs E de, et al. Standardization of health outcomes assessment for depression and anxiety: recommendations from the ICHOM Depression and Anxiety Working Group. *Qual Life Res* 2017;26:3211–3225. Available at: http://link.springer.com/10.1007/s11136-017-1659-5. Accessed March 6, 2018.

20. Morgans AK, Bommel ACM van, Stowell C, Abrahm JL, Basch E, Bekelman JE, et al. Development of a Standardized Set of Patient-centered Outcomes for Advanced Prostate Cancer: An International Effort for a Unified Approach. *Eur Urol* 2015;68:891–8. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0302283815005102. Accessed March 6, 2018.

21. Foust-Wright C, Wissig S, Stowell C, Olson E, Anderson A, Anger J, et al. Development of a core set of outcome measures for OAB treatment. *Int Urogynecol J* 2017;28:1785–1793. Available at: http://link.springer.com/10.1007/s00192-017-3481-6. Accessed March 6, 2018.

22. Akpan A, Roberts C, Bandeen-Roche K, Batty B, Bausewein C, Bell D, et al. Standard set of health outcome measures for older persons. *BMC Geriatr* 2018;18:36. Available at: http://www.ncbi.nlm.nih.gov/pubmed/29394887. Accessed March 6, 2018.

23. Salinas J, Sprinkhuizen SM, Ackerson T, Bernhardt J, Davie C, George MG, et al. An International Standard Set of Patient-Centered Outcome Measures After Stroke. *Stroke* 2016;47:180–186. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26604251. Accessed March 6, 2018.

24. Kim AH, Roberts C, Feagan BG, Banerjee R, Bemelman W, Bodger K, et al. Developing a Standard Set of Patient-Centred Outcomes for Inflammatory Bowel Disease - an International, Cross-disciplinary Consensus. *J Crohns Colitis* 2017. Available at: http://academic.oup.com/ecco-jcc/advance-article/doi/10.1093/ecco-jcc/jjx161/4694039. Accessed March 6, 2018.

25. Zerillo JA, Schouwenburg MG, Bommel ACM van, Stowell C, Lippa J, Bauer D, et al. An International Collaborative Standardizing a Comprehensive Patient-Centered Outcomes Measurement Set for Colorectal Cancer. *JAMA Oncol* 2017;3:686.

26. McNamara RL, Spatz ES, Kelley TA, Stowell CJ, Beltrame J, Heidenreich P, et al. Standardized Outcome Measurement for Patients With Coronary Artery Disease: Consensus From the International Consortium for Health Outcomes Measurement (ICHOM). *Jaha* 2015;4:e001767-. Available at: http://jaha.ahajournals.org/content/4/5/e001767.abstract?cpetoc.

27. Sinha IP, Smyth RL, Williamson PR. Using the Delphi Technique to Determine Which Outcomes to Measure in Clinical Trials: Recommendations for the Future Based on a Systematic Review of Existing Studies. *PLoS Med* 2011;8:e1000393. Available at: http://www.ncbi.nlm.nih.gov/pubmed/21283604. Accessed April 16, 2018.

28. Higgins J, Green S. Cochrane handbook for systematic reviews of interventions. 2011. Available at: https://books.google.com/books?hl=nl&lr=&id=NKMg9sMM6GUC&oi=fnd&pg=PT13&dq=1+Lefebvre+C,+Manheimer+E,+Glanville+J.+Chapter+6:+Searching+for+studies.+In:+Higgins+J,+Green+S+(editors).+Cochrane+Handbook+for+Systematic+Reviews+of+Interventions.+Version+5.1.0. Accessed December 28, 2017.

29. Chatzidionysiou K, Nikiphorou E, Gossec L, Hyrich K. OP0285 The Eular Task Force for Standardising Minimum Data Collection in Rheumatoid Arthritis Observational Research: Results of A Hierarchical Literature. 2016. Available at: http://ard.bmj.com/content/75/Suppl_2/166.1.abstract. Accessed December 28, 2017.

30. Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid

Arthritis Clinical Trials. *Arthritis Rheum* 1993;36:729–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/8507213. Accessed May 27, 2015.

31. Giannini EH, Ruperto N, Ravelli A, Lovell DJ, Felson DT, Martini A. Preliminary definition of improvement in juvenile arthritis. *Arthritis Rheum* 1997;40:1202–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9214419. Accessed March 16, 2018.

32. Orbai A-M, Wit M de, Mease P, Shea JA, Gossec L, Leung YY, et al. International patient and physician consensus on a psoriatic arthritis core outcome set for clinical trials. *Ann Rheum Dis* 2017;76:673–680. Available at: http://ard.bmj.com/lookup/doi/10.1136/annrheumdis-2016-210242. Accessed March 16, 2018.

33. Heijde D van der, Bellamy N, Calin A, Dougados M, Khan MA, Linden S van der. Preliminary core sets for endpoints in ankylosing spondylitis. Assessments in Ankylosing Spondylitis Working Group. *J Rheumatol* 1997;24:2225–9. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9375888. Accessed March 16, 2018.

34. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. *Qual Life Res* 2010;19:539–549. Available at: http://www.ncbi.nlm.nih.gov/pubmed/20169472. Accessed July 17, 2017.

35. Terwee CB, Bot SDM, Boer MR de, Windt DAWM van der, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42. Available at: https://www.researchgate.net/profile/Caroline_Terwee/publication/6637555_Quality_criteria_were_proposed_for_measurement_properties_of_health_status_questionnaires/links/0c960515006c8237c6000000.pdf.

36. Reeve BB, Wyrwich KW, Wu AW, Velikova G, Terwee CB, Snyder CF, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013;22:1889–905. Available at: http://link.springer.com/10.1007/s11136-012-0344-y. Accessed December 28, 2016.

37. Kincaid J, Jr RF, Rogers R, Chissom B. Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel. 1975. Available at: http://www.dtic.mil/docs/citations/ADA006655. Accessed July 17, 2017.

38. Katz PP. Introduction to special issue: patient outcomes in rheumatology, 2011. *Arthritis Care Res (Hoboken)* 2011;63 Suppl 11:S1-3. Available at: http://doi.wiley.com/10.1002/acr.20585. Accessed December 28, 2017.

39. Oude Voshaar MAH, Vonkeman HE, Courvoisier D, Finkckh A, Gossec L, Leung YY, et al. Towards standardized patient reported physical function outcome reporting: linking ten commonly used questionnaires to a common metric. *Qual Life Res* 2018.

40. Singh JA, Saag KG, Jr SLB, Akl EA, Bannuru RR, Sullivan MC, et al. 2015 American College of Rheumatology Guideline for the Treatment of Rheumatoid Arthritis. 2015.

41. Smolen JS, Aletaha D, Bijlsma JWJ, Breedveld FC, Boumpas D, Burmester G, et al. Treating rheumatoid arthritis to target: recommendations of an international task force. *Ann Rheum Dis* 2010;69:631–7.

42. England BR, Sayles H, Mikuls TR, Johnson DS, Michaud K. Validation of the Rheumatic Disease Comorbidity Index. *Arthritis Care Res (Hoboken)* 2015;67:865–872. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25186344. Accessed February 26, 2018.

43. Obbarius A, Maasakkers L van, Baer L, Clark DM, Crocker AG, Beurs E de, et al. Standardization of health outcomes assessment for depression and anxiety: recommendations

from the ICHOM Depression and Anxiety Working Group. *Qual Life Res* 2017;26:3211–3225. Available at: http://link.springer.com/10.1007/s11136-017-1659-5. Accessed December 29, 2017.

44. Verberne W, Das-Gupta Z, Allegretti A, Al E. Development of an international standard set of value-based patient-centered outcome measures for patients with chronic kidney disease. *submitted*.

45. Oude Voshaar MAH, Vonkeman HE, Boonen A, Courvoisier D, Finckh A, Gossec L, et al. An Item Response Theory based common metric for 10 self-reported Physical Function questionnaires was developed and evaluated.

46. Linden W van der, Glas C. Elements of adaptive testing. 2010.

47. Linden W Van der. Linear models for optimal test design. 2006. Available at: https://books.google.nl/books?hl=nl&lr=&id=uAd1eYJniBQC&oi=fnd&pg=PA1&dq=Linear+Models+for+Optimal+Test+Design&ots=H0eIYOFLx9&sig=ReLSfJc4Z0HA8AC1Omgqa57BIyM. Accessed May 27, 2016.

Table 1 Final list of outcomes considered for inclusion in ICHOM IA set

Outcome	qualitative studies in which outcome was reported as important disease outcome	Number of IA conditions for which outcome is included in clinical trial core set	Number of IA conditions for which outcome was measured in ≥1 clinical trial, but not in core set	Percentage of working group members voted for inclusion
Pain	27 (96%)	3	1	100%
Physical function	26 (93%)	4	n/a	100%
Adverse events	11 (39%)	0	4	95%
Fatigue	23 (82%)	2	2	90%
Work/school ability and productivity	14 (50%)	0	3	90%
Overall physical and mental Health Impact	n/a	4	n/a	86%
Inflammatory disease activity	6 (21%)	4	n/a	84%
Participation restrictions	18 (64%)	0	0	55%
Joint damage	7 (25%)	2	1	45%
Mortality	1 (4%)	0	0	40%
Psychological wellbeing	22 (79%)	0	0	35%
Sleep	10 (36%)	0	2	35%
Coping & Self-Management	16 (57%)	0	0	33%
Financial impact	3 (11%)	0	0	25%
Joint stiffness	15 (54%)	0	4	20%
Joint range of motion	0 (0%)	1	0	15%
Physical appearance	7 (25%)	0	0	10%
Sexual functioning	8 (29%)	0	0	10%
Cognitive functioning	5 (18%)	0	0	5%
Fever	0 (0%)	0	1	0%

IA = Inflammatory Arthritis

Table 2 Overview of measures endorsed for assessing Patient-reported outcomes included in ICHOM LA Standard Set

Outcome	Construct Validity	Reliability	Responsiveness	Flesch Kincaid Grade level **	Endorsed instruments
Pain	++	++	++	7	NRS/VAS
	++	++	+	8	RAND-36 Bodily Pain
	?	?	+	6	PROMIS Short Form v1.0 - Pain Interference 8a
	?	+	?	?	PedsQL aches and pain*
Fatigue	+	++	+	5	BRAF-MD
	+	++	++	3	FACIT-F
	++	+	++	8	NRS/VAS
	?	?	-	5	PROMIS Short Form v1.0 - Fatigue 8a
	?	+	?	?	PedsQL 4.0 Fatigue *
Activity limitations	++	++	++	3	HAQ-DI
	+	++	?	4	HAQ-II
	+	++	++	3	MHAQ
	+	?	+/-	4	PROMIS Short Form v2.0 - Physical Function 10a
	+	++	+	3	MDHAQ
	?	++	++	7	BASFI
	+	++	++	4	C-HAQ
	?	?	?	2	JAMAR
Health impact	?	?	?	8	PROMIS Global Health
	--	++	++	6	EQ-5D*
	+	++	++	9	SF-6D
	?	++	+	10	Raid (for RA)
	?	++	+	12	Psaid (for PsA)
	+	++	?	6	ASAS Health Index
	?	++	++	7	Patient/Parent Global Assessment (NRS or VAS)
Work/school/ housework ability and productivity	+	+	+	8	WPAI

++ = favorable properties according >1 study of good methodological quality; + = favorable properties according to 1 study with good methodological quality; +/- Mixed findings in studies of good methodological quality; ? no studies of good methodological quality found; -- = unfavorable properties according >1 study of good methodological quality; - = unfavorable properties according to 1 study with good methodological quality; *Including EQ-5D-Y for pediatric patients; * Instrument is intended for pediatric populaitons; ** Estimated using the Kincaid-Grade level statistic; NRS = Numerical rating scale; VAS = Visual Analog Scale; SF-36 = Medical outcome study short form 36; PROMIS= Patient-reported outcome measurement information system; PedsQL = Pediatric Quality of Life Inventory; BRAF-MD = Bristol Rheumatoid Arthritis Fatigue Multidimensional Questionnaire; FACIT-F = Functional Assessment of Chronic Illness Therapy-Fatigue; HAQ-DI = Health Assessment Questionnaire Disability Index; MHAQ = Modified Health Assessment Questionnaire Disability Index; C-HAQ = Childhood Health Assessment Questionnaire Disability Index; BASFI = Bath Ankylosing Spondylitis Functional Activity Index; JAMAR = The Juvenile Arthritis Multidimensional Assessment Report; EQ-5D = Euroqol 5 dimensions; SF-6D Medical outcome study short form 6D; PsAID = Psoriatic Arthritis Impact of Disease ; RAID = Rheumatoid Arthritis Impact of Disease; WPAI = Work Productivity and Activity Impairment Questionnaire

Table 3 Case-mix variables

Variable	Definition and response options	Timing**	Data Source
AGE	Year of birth YYYY	Baseline	Patient
SEX	Sex at birth Female/male	Baseline	Patient
SMOKING STATUS	Never / former / current	Baseline	Patient
LEVEL OF EDUCATION	Highest attained education ISED classification (none / primary / secondary/tertiary)	Baseline	Patient
COMORBIDITIES	Present / Absent / Unknown : Chronic lung disease, myocardial infraction, other heart disease, stroke, hypertension, diabetes, fracture, depression, ulcer or stomach problem, cancer, central sensitization to pain, obesity (i.e. body mass index >=30)	Baseline	Clinical
DIAGNOSIS	Physician reported diagnosis RA / SpA/ PsA/JIA	Baseline and annually	Clinical
DISEASE DURATION	Year of diagnosis YYYY	Baseline	Clinical
IMMUNOLOGICAL*	Rheumatoid factor and ACPA positivity (yes/no)	Baseline	Clinical

*only for RA and JIA. ISED = institute for the Study of Education and Human Development; RA = Rheumatoid arthritis; arthritis Spa = Spondyloarthritis; PsA = Psoriatic Arthritis; Juvenile idiopathic Arthritis; ACPA = anti-citrullinated protein antibody=

**baseline = defined as first measurement for this patient

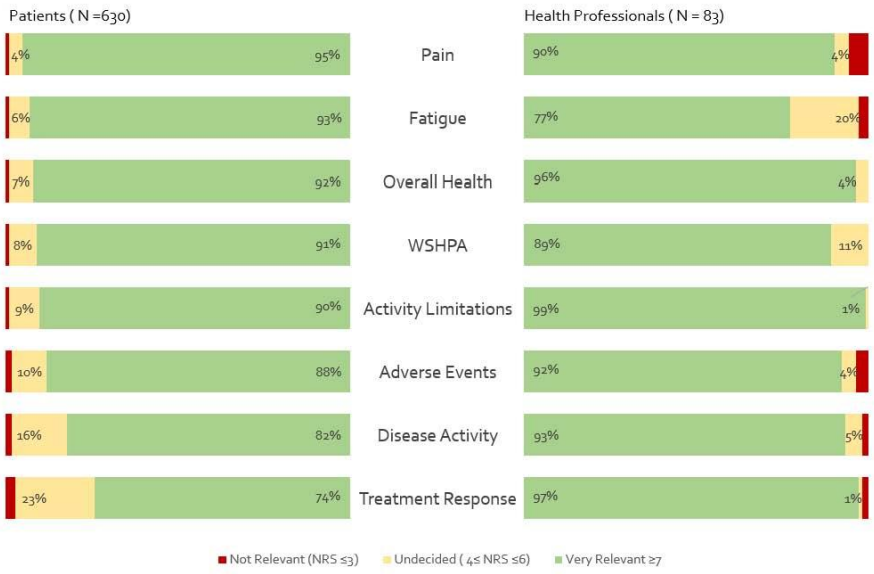

Figure 1 Relevance of outcomes included in ICHOM IA set according to patient and health professional open review

WSPHA = Work/School/Household productivity and ability