Title: Domains to be considered for the core outcome set of axial spondyloarthritis: results from a 3-round Delphi survey.

Authors: Anne Boel (ORCID 0000-0003-2016-1744)[1], Victoria Navarro-Compán[2], Annelies Boonen (ORCID 0000-0003-0682-9533)[3,4], Philip Mease (ORCID 0000-0002-6620-0457)[5], Uta Kiltz (ORCID 0000-0001-5668-4497)[6], Maxime Dougados[7,8], Robert Landewé (ORCID 0000-0002-0577-6620)[9,10], Désirée van der Heijde (ORCID 0000-0002-5781-158X)[1]

Key indexing terms: ankylosing spondylitis, outcomes, spondyloarthropathies

Affiliations:

[1] Rheumatology Department, Leiden University Medical Centre, Leiden, the Netherlands
[2] Rheumatology Department, University Hospital La Paz, IdiPaz, Madrid, Spain
[3] Department of Internal Medicine, Division of Rheumatology, Maastricht University Medical Centre, Maastricht, The Netherlands.
[4] Care and Public Health Research Institute (CAPHRI), Maastricht University, Maastricht, The Netherlands.
[5] Rheumatology, Swedish Medical Center/Providence St. Joseph Health and University of Washington, Seattle, Washington, USA
[6] Rheumatology, Rheumazentrum Ruhrgebiet, Ruhr University Bochum, Herne, Germany.
[7] Rheumatology Department, - Hôpital Cochin. Assistance Publique - Hôpitaux de Paris
[8] INSERM (U1153): Clinical Epidemiology and Biostatistics, PRES Sorbonne Paris-Cité, Paris, France
[9] Rheumatology Department, Zuyderland Medical Center Heerlen, Heerlen, The Netherlands.
[10] Rheumatology, Amsterdam Rheumatology Center, AMC, Amsterdam, The Netherlands.

Funding The Assessment of Spondyloarthritis international Society (ASAS) funded Anne Boel and Victoria Navarro-Compán to work on the project to update the core outcome set.

Conflicts of interest None declared

Correspondence to:
Anne Boel, Department of Rheumatology, Leiden University Medical Centre, *P.O. Box 9600 , 2300 RC* Leiden, The Netherlands. a.h.e.m.boel@lumc.nl (+3171 526 3423)

Running head: Domains core set axSpA

ABSTRACT

Background: Advances in the field of axial spondyloarthritis (axSpA) and the methodology to develop core sets made the ASAS group decide to update the ASAS-OMERACT core set. An important aspect was to ensure it will be applicable to the entire spectrum of axSpA. The first step was to define the most relevant disease domains.

Methods: A 3-round Delphi survey was conducted to gather opinions of 188 patients and 188 axSpA experts to define the most relevant disease domains to be included in the core set. The Delphi survey evaluated two separate research settings: 1) studies assessing symptom modifying therapies; 2) studies evaluating disease modifying therapies. Importance of domains was rated on a 1–9 Likert scale. A domain was considered for inclusion if for both stakeholder groups ≥70% of participants scored the domain as critical (7-9) and ≤15% scored it as not important (1-3) after three rounds.

Results: A total of 132 (70%) patients and 135 (72%) experts completed at least 1 round. After three rounds, 7 domains (pain, physical function, stiffness, disease activity, mobility, overall functioning & health, peripheral manifestations) were selected for the symptom modifying therapies setting. For the disease modifying therapies setting, 6 domains (physical function, disease activity, mobility, structural damage, extra-musculoskeletal manifestations, peripheral manifestations) were selected. All domains selected by experts were also selected by patients. Patients selected all offered domains except 'emotional function'.

Conclusion: This study provides the domains selected by patients and axSpA experts that should be considered for the core set for axSpA.

Introduction

The Assessment of SpondyloArthritis international Society (ASAS) collaborated with Outcome Measures in Rheumatology (OMERACT) to develop a core outcome set for ankylosing spondylitis (AS) in 1999(1). The core set has been well implemented in the field in the past 20 years(2). Nevertheless, since the development of the original core set it has become apparent that AS belongs to a broader disease spectrum, axial spondyloarthritis (axSpA), which consists of two subtypes: radiographic axSpA (also known as AS), and non-radiographic axSpA(3). Furthermore, major advances have occurred in the outcome instruments in the field of axial spondyloarthritis, such as the use of magnetic resonance(4), the development of the Ankylosing Spondylitis Disease Activity Score (ASDAS)(5), validated enthesitis scores(6), the ASAS-health index(7) and the ASAS-flare definition(8).

In addition, the methodology to develop core outcome sets has improved. Despite there is no gold standard on the development or update of a core set, during the last years OMERACT and Core Outcome Measures in Effectiveness Trials (COMET) have intensively worked to provide specific guidance on how a core set should be developed, e.g. OMERACT handbook(9) and OMERACT Filter 2.0(10), COMET handbook(11) and Core Outcome Set-Standards for Development(12). All these advances made the ASAS group decide it was necessary to update the original ASAS-OMERACT core set for AS, to ensure the core set will be applicable to the entire spectrum of axSpA and developed according to the current recommended methodology.

An important step in the process to update the core outcome set was to define which disease domains (outcomes) are relevant. In order to establish these, a three-round Delphi survey was employed to gather opinions from relevant stakeholders. The results of this three-round Delphi survey formed the basis of the proposal for a final core set according to the new format of the OMERACT onion(13). Subsequently, the proposal was presented to OMERACT to seek endorsement for the proposed core domain set. A detailed description of the entire process that led to the selection of domains for the updated core set has been published separately (*Navarro-Compán et al., submitted*). In this paper we describe the methods used to compose and execute the Delphi survey and its results. The aim of this study was to select the domains that should be considered for inclusion in the core set for axSpA.

Materials and Methods

Preparation of the Delphi survey

The original core set(14) was developed for three different scenarios: 1) therapies which improve the symptoms and clinical features of inflammatory manifestations of the disease (so-called SMART: symptom modifying antirheumatic therapy; this includes physical therapy); 2) therapies that change the course of disease through decreasing inflammatory manifestations (thereby improving function) and prevention of or decreasing structural damage (so-called DMARDs: disease modifying antirheumatic drugs); 3) clinical record keeping in daily practice, to facilitate uniform clinical record keeping to enable research from clinical records and to monitor patient care in a standardised way.

The core set update focused on the first two scenarios only. Thus, the Delphi survey consisted of two separate sections, one focused on the outcomes to be included in the core set for studies assessing symptom modifying therapies; the other on the outcomes to be included in the core set for studies evaluating disease modifying therapies.

A list of candidate domains to include in the Delphi survey was computed using three sources: 1) the current core set for AS(14); 2) all domains assessed in studies evaluating pharmacological and non-pharmacological interventions identified in the systematic literature review (SLR) that assessed the implementation of the original core set(2). To ensure the most recent studies were included too, the search strategy from the SLR was used to identify studies published thereafter (i.e. between 2011 and 2018); 3) information collected on the qualitative studies and patient focus group interviews conducted as part of the development of the ASAS/WHO Comprehensive and Brief Core sets of the International Classification of Functioning, Disability and Health (ICF) for AS(15, 16). All aspects of health identified in this process were considered when defining candidate domains for the core set for axSpA.

After eliminating duplicates, the list of candidate domains was grouped and finalised by three of the authors (DvdH, VNC, AB) and later agreed on by the steering committee. The first round of the Delphi survey contained 11 candidate domains for symptom modification therapies and 12 candidate domains for disease modification therapy (the same domains with one additional domain representing structural damage). For this first round, participants had the opportunity to suggest additional domains.

Participants

The invited participants were divided in two main stakeholder groups: one group consisted of patients with axSpA and the other group consisted of a variety of expert stakeholders (all ASAS members, including rheumatologists, other health care professionals, methodologists, and researchers, as well as representatives from pharmaceutical industry and drug regulatory agencies) labelled as axSpA experts. The ASAS members were informed they would be invited to partake in the Delphi survey to update the current core set in an annual meeting prior to commencement of the project. Representatives of pharmaceutical industry and drug regulatory agencies were informed of the project via email and invited to partake prior to commencement of the project.
Patients were recruited through three national patient societies (SAA (Spondylitis Association of America), NASS (National Ankylosing Spondylitis Society), and CSA (Canadian Spondylitis Association)) and eligible to partake if they were ≥18 years and had a diagnosis of axSpA from their rheumatologist. Information regarding the Delphi survey and its purpose was posted on the websites

of each of the organizations and patients were contacted by their associations via email inviting them to partake. Recruitment ceased once the group of patients was equal in size to the group of experts (N=188). Ethical approval and consent to participate in the Delphi survey was not required based on the Dutch Medical Research Involving Human Subjects Act (WMO).

Content of the Delphi survey

An explanatory text was provided at the beginning of the survey in each round, which contained information on the purpose of the Delphi and relevant information to fully understand the content and scoring system. This information was adapted per stakeholder group, using lay wording and more extensive explanations for the patients.

The main objective of this Delphi survey was to select the most relevant disease domains to be included in the core set for axSpA. Nevertheless, simultaneously this Delphi survey was also used to investigate the effect of invitation approach on the response rate and final outcome of a Delphi survey. The methods and results of this experiment are published separately(17). In summary, the participants were not aware of the experiment and received identical information regarding the Delphi survey. All participants knew from the start that this was a three-round Delphi, but did not know that for half of the participants an invitation for the second and third round was conditional on responding to the first round. The experiment on the two different ways of inviting participants showed no effect on the final results of the Delphi survey and this is published separately(17). In any case, for the purpose of the selection of the domains, it was predefined that the information from all participants irrespective of invitation approach would be used. This is the first time that the results of the Delphi for the two different stakeholders, which will be used for the core set, are published.

The Delphi survey was split according to the two established scenarios (i.e. SMART and DMARD) and grouped by domain: i.e. participants were invited to vote on the relevance of a specific domain with regards to symptom modifying therapies first and immediately thereafter they were asked to vote on the same domain with regards to disease modifying therapies. This order was maintained for all domains except structural damage, which was only offered for voting in the disease modifying therapies section of the survey. A definition was provided for each domain in all rounds, including a brief explanation and examples (supplementary table S.1).

In each round, the participants received summarized information of the previous round, including aggregated scores from their respective stakeholder group and their individual score. Participants who responded for the first time to the invitation for the second round received only aggregated scores of the first round, and the same procedure applied to round three.

Each round was open for 2-3 weeks and a single reminder was sent after one week to those who did not yet complete the round. Data were collected online using SurveyMonkey between November 2nd and December 30th 2018.

Domain selection

To identify the importance of each of the domains for the core set, each participant was asked to provide one score per domain using a 9-point Likert scoring system. Domains were graded in accordance to their level of importance. Following the OMERACT handbook, a score of 1-3 signified an outcome as not important; 4-6 as important but not critical; and 7-9 as critical(9). The aggregated scores per domain were analysed separately for each of the stakeholder groups. If a domain was scored as critical by ≥80% of the participants in a stakeholder group, the domain was selected for consideration in the core set, and not offered for voting in subsequent rounds for this stakeholder group. If a domain did not achieve this score, the predefined criteria to include a domain in the next

round of the Delphi per stakeholder group were: at least 50% of the participants scored the domain as critical; and: 20% or less scored the domain as not important.

Finally, a domain was considered for inclusion in the core set if for both stakeholder groups (experts and patients) ≥70% of participants scored the domain as critical and ≤15% scored it as not important after three rounds, which is in line with the guidelines provided in the OMERACT handbook(9).

Statistical analysis

For the purpose of this study, we used descriptive statistics to present the data. To determine which domains fulfilled the criteria to be considered for inclusion the proportion of participants voting critical, important but not critical and not important were calculated.

Results

In total, 376 participants were invited to partake: 188 patients and 188 axSpA experts. Patients represented three countries in two continents, axSpA experts represented 41 countries from five continents (Supplementary table S.2). The axSpA experts who completed at least 1 round consisted of 123 rheumatologists -of whom 10 were also methodologists and 2 were also patient representatives-, 4 physiotherapists, 4 representatives of pharmaceutical companies, 2 radiologists, and 2 researchers.

Participants

The overall response rate was 49% for the patients, and 58% for the axSpA experts after the final round. In addition, round 1 and 2 were completed by 63% and 52% of patients and 60% and 55% of axSpA experts respectively.

Content of the Delphi survey

In round 1, stiffness was mentioned by multiple axSpA experts and was therefore added to the list of domains from round 2 onwards to the survey for patients and axSpA experts to both settings (i.e. symptom and disease modifying therapies). Supplementary table S.3 provides an overview of the domains that were offered for voting in each round for each of the stakeholder groups.

Domain selection

Tables 1 and 2 present the proportion of critical votes per domain after the final round -split by stakeholder group- for the symptom and disease modifying therapies scenario respectively. Domains voted critical by ≥70% and not important by ≤15% are depicted in **bold**. Supplementary tables S.4 and S.5 present additional information on the proportions of critical, important but not critical, and not important votes per domain per round.

For the symptom modifying therapies scenario, seven domains were voted critical by ≥70% of patients and axSpA experts after three rounds, these were: disease activity, pain, overall functioning & health, physical function, mobility, peripheral manifestations, and stiffness (table 1). An additional 4 domains were voted critical by ≥70% of patients only, in fact, the domain emotional function was the only domain voted critical by <70% of patients. There were no domains voted critical by ≥70% of axSpA experts only.

For the disease modifying therapies scenario, six domains were selected by ≥70% of patients and axSpA experts after the final round, these were: disease activity, physical function, mobility, peripheral manifestations, extra-musculoskeletal manifestations, and structural damage (table 2). An additional 6 domains were voted critical by ≥70% of patients only, who thereby selected all domains except emotional function. Identical to the symptom modifying therapies scenario, there were no domains voted critical by experts only.

The domains that were voted critical by ≥70% and voted not important by <15% in both stakeholder groups are presented in figure 1 (1A symptom, 1B disease modifying therapies scenario).

***Table 1** Proportion of critical votes per domain after the final round for the symptom modifying therapies scenario, split by stakeholder group. Domains voted critical by ≥70% and not important by ≤15% of participants are printed in **bold**.

***Table 2** Proportion of critical votes per domain after the final round for the disease modifying therapies scenario, split by stakeholder group. Domains voted critical by ≥70% and not important by ≤15% of participants are printed in **bold**.

***Figure 1** Domains selected after three rounds by patients (dark) and axSpA experts (light) in the setting assessing symptom modifying therapies (A) and disease modifying therpies (B), including the percentage of critical votes.

Discussion

This three-round Delphi survey was an important step in the process to update the core outcome set and aimed to determine which domains should be considered for inclusion according to patients and axSpA experts. Only one additional domain (stiffness) was added from round 2 onwards, indicating that the candidate domains identified in the preparatory steps were a good representation of the domains of interest in the field.

In our study, patients selected more domains to be included in the core set compared with the axSpA experts. Especially, domains such as fatigue, sleep and work & employment were deemed very important by patients, but less so by experts. These domains have a major impact on the daily life of the patient, but are not necessarily specific to the disease, which could explain the difference in importance between patients and experts. In general, axSpA experts deemed the more objectively measurable domains such as structural damage and mobility most critical to be measured in settings investigating disease modifying therapies, whereas the more subjective domains such as pain, stiffness, and overall functioning & health were limited to the settings investigating symptom modifying treatments.

The domain with the highest percentage of critical votes in the axSpA experts group was disease activity, which held true for both settings indicating that this domain is most important to measure in all trials investigating therapies for axSpA according to experts, whereas this was pain according to patients. A noticeable difference was the domain pain in the disease modifying therapies setting, which was voted critical by 95% of patients, yet only 65% of the experts deemed this domain important enough to be measured in all trials investigating DMARDs.

A large panel of international axSpA experts and patients were invited to partake in this study. The use of an electronically distributed Delphi ensured no travel was required and anonymity was guaranteed. Furthermore, no public speaking was required, which is known to increase patient participation(18). Despite these measures, the majority of the axSpA experts who responded to at least one round were from Europe and America. The same limitation applies to the patients, as invitations were restricted to native English speakers to ensure understanding of the survey and its components. Nevertheless, all stakeholder groups who will benefit from an updated core outcome set were included in its development, which we hope will increase uptake. Finally, OMERACT and COMET methodology were followed as closely as possible.

Conclusion

This Delphi survey study identified 7 domains that should be considered for the core set evaluating the efficacy of symptom modifying therapies, and 6 domains that should be considered for the core set investigating disease modifying therapies according to patients and axSpA experts. The results from this study will be used to compose the core outcome set for axSpA, in which a distinction will be made for the domains mandatory for studies assessing symptom modifying therapies and studies evaluating disease modifying therapies. After finalising the core outcome set, the next step for ASAS will be to identify appropriate instruments to measure the chosen domains.

Acknowledgements We would like to thank all patients from the Spondylitis Association of America, Canadian Spondylitis Association and National Axial Spondyloarthritis Society for their participation in the Delphi survey. We would also like to thank all ASAS members as well as the representatives of the pharmaceutical companies and drug regulatory agencies for their contribution to the project.

References

1. van der Heijde D, van der Linden S, Dougados M, Bellamy N, Russell AS, Edmonds J. Ankylosing spondylitis: Plenary discussion and results of voting on selection of domains and some specific instruments. J Rheumatol 1999;26:1003-5.
2. Bautista-Molano W, Navarro-Compán V, Landewé RB, Boers M, Kirkham JJ, van der Heijde D. How well are the ASAS/OMERACT core outcome sets for ankylosing spondylitis implemented in randomized clinical trials? A systematic literature review. Clin Rheumatol 2014;33:1313-22.
3. Rudwaleit M, Khan MA, Sieper J. The challenge of diagnosis and classification in early ankylosing spondylitis: Do we need new criteria? Arthritis Rheum 2005;52:1000-8.
4. Rudwaleit M, Jurik AG, Hermann K-GA, Landewé R, van der Heijde D, Baraliakos X, et al. Defining active sacroiliitis on magnetic resonance imaging (MRI) for classification of axial spondyloarthritis: A consensual approach by the ASAS/OMERACT MRI group. Ann Rheum Dis 2009;68:1520-7.
5. Lukas C, Landewé R, Sieper J, Dougados M, Davis J, Braun J, et al. Development of an asas-endorsed disease activity score (asdas) in patients with ankylosing spondylitis. Ann Rheum Dis 2009;68:18-24.
6. Heuft-Dorenbosch L, Spoorenberg A, van Tubergen A, Landewé R, van ver Tempel H, Mielants H, et al. Assessment of enthesitis in ankylosing spondylitis. Ann Rheum Dis 2003;62:127-32.
7. Kiltz U, van der Heijde D, Boonen A, Cieza A, Stucki G, Khan MA, et al. Development of a health index in patients with ankylosing spondylitis (ASAS HI): Final result of a global initiative based on the ICF guided by ASAS. Ann Rheum Dis 2015;74:830-5.
8. Gossec L, Portier A, Landewé R, Etcheto A, Navarro-Compán V, Kroon F, et al. Preliminary definitions of 'flare' in axial spondyloarthritis, based on pain, BASDAI and ASDAS-CRP: An ASAS initiative. Ann Rheum Dis 2016;75:991-6.
9. Boers M, Kirwan JR, Tugwell P. OMERACT handbook. 2018
10. Boers M, Kirwan JR, Gossec L, Conaghan PG, D'Agostino MA, Bingham CO, et al. How to choose core outcome measurement sets for clinical trials: OMERACT 11 approves filter 2.0. J Rheumatol 2014;41:1025-30.
11. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET handbook: Version 1.0. Trials 2017;18:280.
12. Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, Tunis S, et al. Core outcome set-standards for development: The COS-stad recommendations. PLoS Med 2017;14:e1002447.
13. Maxwell LJ, Beaton DE, Shea BJ, Wells GA, Boers M, Grosskleg S, et al. Core domain set selection according to OMERACT filter 2.1: The OMERACT methodology. J Rheumatol 2019;46:1014-20.
14. van der Heijde D, van der Linden S, Bellamy N, Calin A, Dougados M, Khan MA. Which domains should be included in a core set for endpoints in ankylosing spondylitis? Introduction to the ankylosing spondylitis module of OMERACT IV. J Rheumatol 1999;26:945-7.
15. Boonen A, Braun J, van der Horst Bruinsma IE, Huang F, Maksymowych W, Kostanjsek N, et al. ASAS/WHO ICF core sets for ankylosing spondylitis (AS): How to classify the impact of AS on functioning and health. Ann Rheum Dis 2010;69:102-7.
16. Boonen A, van Berkel M, Kirchberger I, Cieza A, Stucki G, van der Heijde D. Aspects relevant for functioning in patients with ankylosing spondylitis according to the health professionals: A Delphi study with the ICF as reference. Rheumatology (Oxford) 2009;48:997-1002.
17. Boel A, Navarro-Compán V, Landewé R, van der Heijde D. Two different invitation approaches for consecutive rounds of a Delphi survey led to comparable final outcome. J Clin Epidemiol 2020;129:31-9.
18. Khodyakov D, Grant S, Denger B, Kinnett K, Martin A, Peay H, et al. Practical considerations in using online modified-Delphi approaches to engage patients and other stakeholders in clinical practice guideline development. The patient 2020;13:11-21.

Table 1 Proportion of critical votes per domain after the final round for the symptom modifying therapies scenario, split by stakeholder group. Domains voted critical by ≥70% and not important by ≤15% of participants are printed in **bold**.

	Patients			axSpA experts		
	N	Count	%	N	Count	%
Symptom modifying therapies scenario						
Disease activity	97	85	**88**	113	110	**97**
Pain	119	115	**97**	113	98	**87**
Fatigue	119	99	**83**	109	56	51
Sleep	119	96	**81**	109	22	20
Overall functioning & health	119	96	**81**	103	89	**86**
Physical function	119	109	**92**	113	98	**87**
Emotional function	93	57	61	103	13	13
Work & Employment	93	72	**77**	109	34	31
Mobility	119	104	**87**	109	81	**74**
Peripheral manifestations	119	98	**82**	109	90	**83**
Extra-musculoskeletal manifestations	119	99	**83**	109	74	68
Stiffness	97	87	**90**	109	94	**86**

Table 2 Proportion of critical votes per domain after the final round for the disease modifying therapies scenario, split by stakeholder group. Domains voted critical by ≥70% and not important by ≤15% of participants are printed in **bold**.

	Patients			axSpA experts		
	N	Count	%	N	Count	%
Disease modifying therapies scenario						
Disease activity	119	106	**89**	113	99	**88**
Pain	119	113	**95**	109	71	65
Fatigue	97	87	**90**	113	40	35
Sleep	93	72	**77**	113	18	16
Overall functioning & health	119	102	**86**	109	73	67
Physical function	119	109	**92**	103	90	**87**
Emotional function	93	52	56	113	12	11
Work & Employment	93	68	**73**	109	31	28
Mobility	119	105	**88**	109	88	**81**
Peripheral manifestations	119	98	**82**	109	78	**72**
Extra-musculoskeletal manifestations	119	102	**86**	109	77	**71**
Structural damage	119	102	**86**	113	95	**84**
Stiffness	97	87	**90**	109	53	49

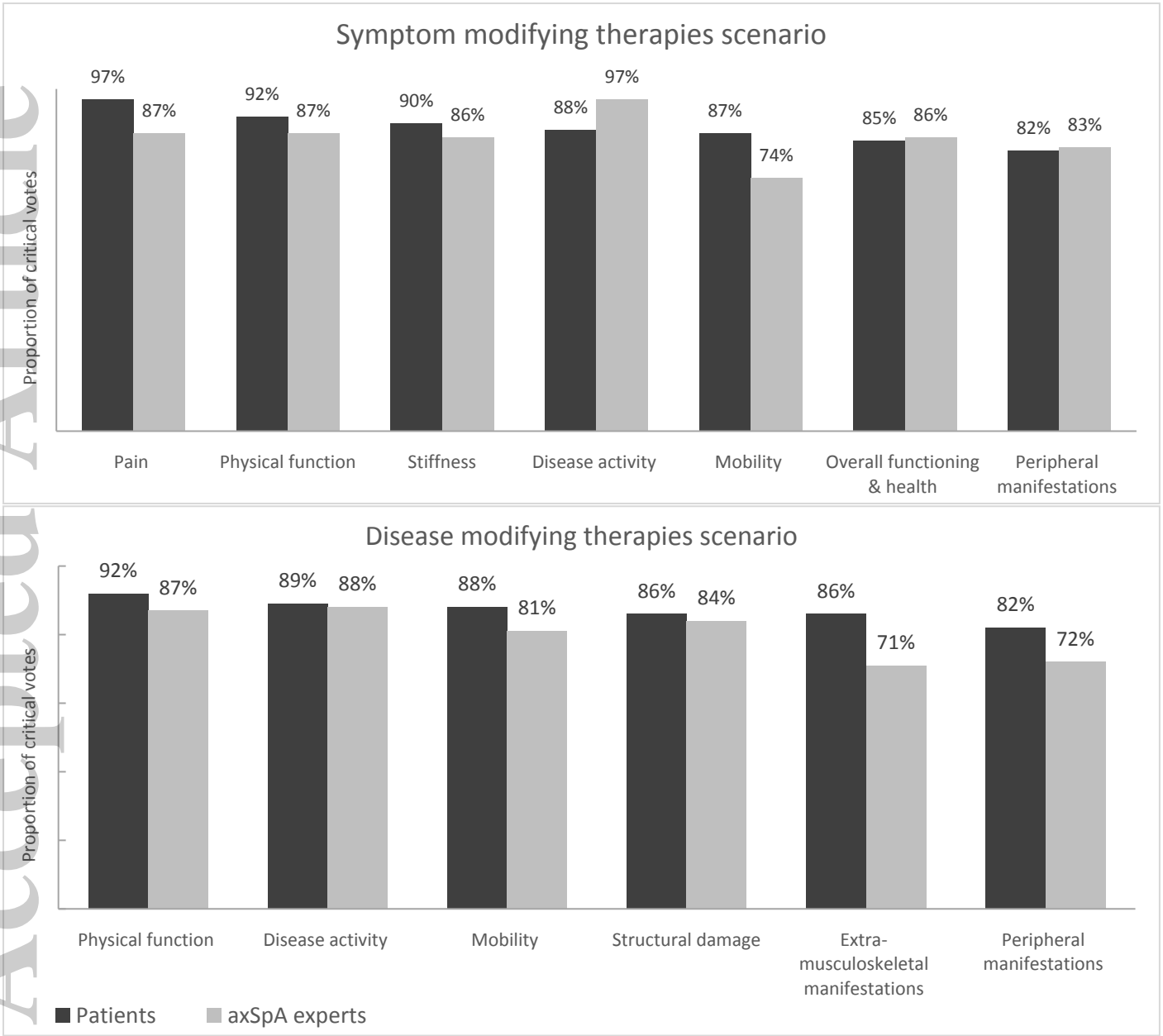

Figure 1 Domains selected after three rounds by patients (dark) and axSpA experts (light) in the setting assessing symptom modifying therapies (A) and disease modifying therpies (B), including the percentage of critical votes.