Two different invitation approaches for consecutive rounds of a Delphi survey led to comparable final outcome

Boel, A.; Navarro-Compan, V.; Landewe, R.; Heijde, D. van der

Journal of Clinical Epidemiology 129 (2021) 31−39

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Two different invitation approaches for consecutive rounds of a Delphi survey led to comparable final outcome

Anne Boel[a,*], Victoria Navarro-Compán[b], Robert Landewé[c,d], Désirée van der Heijde[a]

[a]Department of Rheumatology, Leiden University Medical Center, Leiden, The Netherlands
[b]Department of Rheumatology, University Hospital La Paz, Madrid, Spain
[c]Department of Rheumatology, Zuyderland Medical Center Heerlen, Heerlen, The Netherlands
[d]Department of Clinical Immunology & Rheumatology, Amsterdam University Medical Center, Amsterdam, The Netherlands

Accepted 24 September 2020; Published online 28 September 2020

Abstract

Objectives: There are two different approaches to involve participants in consecutive rounds of a Delphi survey: (1) invitation to every round independent of response to the previous round (''all-rounds'') and (2) invitation only when responded to the previous round (''respondents-only''). This study aimed to investigate the effect of invitation approach on the response rate and final outcome of a Delphi survey.

Study Design and Setting: Both experts ($N = 188$) and patients ($N = 188$) took part in a Delphi survey to update the core outcome set (COS) for axial spondyloarthritis. A study with 1:1 allocation to two experimental groups (ie, ''all-rounds'' [$N = 187$] and ''respondents-only'' [$N = 189$]) was built-in.

Results: The overall response rate was lower in the ''respondents-only group'' (46%) compared to the ''all-rounds group'' (61%). All domains that were selected for inclusion in the COS by the ''respondents-only group'' were also selected by the ''all-rounds group.'' Additionally, the four most important domains were identical between groups after the final round, with only minor differences in the other domains.

Conclusion: Inviting panel members who missed a round to a subsequent round will lead to a better representation of opinions of the originally invited panel and reduces the chance of false consensus, while it does not influence the final outcome of the Delphi.

Keywords: Delphi survey; Delphi technique; Core outcome set; Spondyloarthritis; Consensus building

1. Introduction

The Delphi technique is a structured forecasting method based on the presumption that combining the opinion of a group of experts will result in a more accurate prediction of the truth than relying on the opinion of a most knowledgeable single individual [1]. Responses can be altered between rounds, based on the aggregated information of peers from the previous round [2]. An additional benefit of the Delphi is that participants tend to perceive ownership of the results due to their participation in the process. In turn, this perceived ownership improves the acceptance of the findings among those who participated [3]. As the participants are a reflective sample of the end users, their involvement in the development stage increases implementation in the field [4]. The Delphi process ends when (the predefined level of) consensus is achieved, or when the prespecified number of rounds has been completed [5].

Anne Boel helped in data curation, formal analysis, validation, visualization, writing original draft, and writing review & editing; Victoria Navarro-Compán helped in conceptualization, methodology, project administration, and writing review & editing; Robert Landewé helped in methodology and writing review & editing; and Désirée van der Heijde helped in conceptualization, methodology, project administration, funding acquisition, supervision, and writing review & editing.

Funding: The Assessment of SpondyloArthritis international Society funded the fellow and Victoria Navarro-Compán for the project to update the core outcome set.

Declaration of interest: None declared.

* Corresponding author. Department of Rheumatology, Leiden University Medical Centre, P.O. Box 9600, 2300 RC Leiden, The Netherlands. Tel.: +31-71-526-3423; fax: +3171 526 6752.

E-mail address: a.h.e.m.boel@lumc.nl (A. Boel).

https://doi.org/10.1016/j.jclinepi.2020.09.034

What is new?

- The comparison of two different ways of inviting participants to a 3-round Delphi survey using randomization.

What this adds to what is known?

- Invitation approach did not influence the final outcomes of the Delphi survey.
- An identical first-round survey completed by two randomly selected independent panels resulted in a similar outcome.

What is the implication, what should change now?

- Inviting persons for all rounds irrespective of a response to the previous round will lead to a better representation of the opinions of the originally invited panel, while it does not influence the final outcome of the Delphi.

Common applications of the Delphi technique in healthcare settings are the selection of outcomes for a core set and the identification of research priorities [1,6]. Even though the Delphi technique is often used, there is hardly any guidance on the methodology underlying the Delphi technique [5−7], which results in large variability in its execution. Research on methodological guidance of the Delphi technique in the development of core outcome sets (COS) is slowly increasing [8−11], but a lot of the methodology remains unclear to this day. Guidance on which participants to invite to consecutive rounds has not yet been described in existing literature. There are two options: (1) invite only participants that have completed the previous round for the consecutive round. This approach ensures participants provide their own authentic opinion in the first round and are challenged to rethink their own response in light of the responses of others each round. Hence, this approach increases engagement in the decision-making process and the final outcome of the Delphi will be an accurate representation of the opinions of those who participated. (2) Invite every participant for all consecutive rounds irrespective of whether they have responded or not. This approach decreases the chance of nonrandom loss of opinions which could lead to false consensus [6], as it considers the opinion of every participant who completed one or more rounds, and the final outcome may therefore be a better representation of the opinions of the entire panel that was invited to partake. Scientific evidence to guide Delphi researchers on whether panel members who miss a round can be included in a subsequent round is sparse. Yet, if the results are consistent with the conventional approach of excluding these experts from subsequent rounds, the final outcome may be a better reflection of the opinions of the originally invited panel and false consensus caused by drop-out of those with a different opinion may be reduced.

The objective of this study is to investigate two different approaches of inviting participants to consecutive rounds in a 3-round Delphi survey and their effect on the final result of the Delphi and the (overall) response rate.

2. Materials and methods

2.1. Design and population

Two stakeholder groups were invited by email to partake in two separate Delphi surveys, as part of a larger project to update the Assessment of SpondyloArthritis international Society (ASAS)/Outcome Measures in Rheumatology (OMERACT) COS for axial spondyloarthritis (axSpA) [12]. One group consisted of patients with axSpA and the other group consisted of axSpA experts, including rheumatologists, other healthcare professionals, methodologists, and other stakeholders. The axSpA experts were all ASAS members, who were informed they would be invited to partake in the Delphi survey to update the current COS in an annual meeting before commencement of the project. The patients with axSpA were recruited through three national patient societies (Spondylitis Association of America, National Ankylosing Spondylitis Society [UK], and Canadian Spondylitis Association) and eligible to partake if they had a diagnosis of axSpA from their rheumatologist. Patients were contacted by their associations via email, in which the study was explained and patients were asked to participate. Additionally, information was placed on the websites of each of the organizations. Patients could either send an email directly to the researcher in charge of sending the Delphi invitations (A.B.) or their respective organization if they were interested in partaking in the Delphi survey. Recruitment ceased once the group of patients was equal in size to the group of experts ($N = 188$). This was an opportunistic sample and no sample size calculations were performed upfront. The main objective of this Delphi survey was to define which are the most relevant disease domains (outcomes) for all stakeholders to be included in the updated COS. We did not specifically ask consent for the experiment, because knowledge about the assignment would have biased the results.

For each separate stakeholder group, the invited participants were randomly allocated 1:1 to two experimental groups to ensure an even distribution of stakeholders in each experimental group. Experimental group 1 was labeled as ‘‘respondents-only group,’’ and experimental group 2 as ‘‘all-rounds group.’’ Two randomization sequences (one for patients, one for experts) were created, using a computer-generated schedule (developed by a member of the data management team of the rheumatology department in the Leiden University Medical Center). Randomization was performed by a researcher (A.B.) after

the recruitment of patients was complete. Only the researcher in charge of sending the Delphi and collecting the data (A.B.) was aware of the group allocation of each participant.

Even though the results of the Delphi were analyzed separately for each of the stakeholder groups (ie, patients and experts) to update the COS for axSpA, the stakeholder groups were not the focus of this study. The aim of this experiment was solely to compare the responses of the ''all-rounds group'' with those of the ''respondents-only group'' to investigate whether there is an effect of invitation procedure on the final result of the Delphi survey. Therefore, all data in this manuscript focused only on the differences between the ''respondents-only'' and ''all-rounds'' groups.

The participants in the ''respondents-only group'' received an invitation for the second round only if they completed the first round, and only received an invitation for the third round if they completed the second round. The participants in the ''all-rounds group'' received an invitation for each round irrespective of response to any of the previous rounds. In each round, the participants received summarized information of the previous round, including aggregated scores from their respective stakeholder group. Those participants who partook in the previous round received their individual score of the previous round as well. Participants in the ''all-rounds group'' 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 3.

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 an invitation for the second and third round was conditional on responding to the first round.

Each round was open for 2−3 and a single reminder was sent after 1 wk to those who did not yet complete the round. Data were collected online using SurveyMonkey between November 2, 2018 and December 30, 2018.

2.2. Survey questionnaires

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. In this manuscript, the results for the survey on symptom modification will be described in detail and the results on disease modification only in the appendix.

The survey on symptom modification contained 11 candidate domains for the core set in the first round, and participants had the opportunity to suggest additional domains in this round. This led to the addition of one more domain from round 2 onwards, which brought the total number of domains to 12. It was decided upfront that the survey on disease modification would contain the same domains as the survey on symptom modification.

To identify importance of each of the domains for the core set, each participant was asked to provide one score per domain using a nine-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 [13]. The criteria to include a domain in the next round of the Delphi were that at least 50% of the participants scored the domain as critical and 15% or less scored the domain as not important. If a domain was scored as critical by $\geq$80% in the specific stakeholder group, the domain was considered selected for the core set, and not offered for voting in subsequent rounds within that stakeholder group. The aggregated scores per domain were analyzed separately for each of the stakeholder groups.

A domain was considered for inclusion in the update of the core set if $\geq$70% of participants scored the domain as critical, and $\leq$15% scored not important after the final round, which was in line with the guidelines provided in the OMERACT handbook for the development of COS [13]. This was the first step in the update of the COS for axSpA and explains the availability of the data per round. More detailed information about the selection of domains is beyond the scope of this article. In addition to the nine-point Likert scoring grade, all participants were asked to select six domains from the list of possible domains in each Delphi round. The chosen domains were those with the highest priority to be included in the COS; there was no further ranking within the six chosen domains.

2.3. Outcomes

The main outcomes of this experiment were as follows: (1) the response rates after each round and the final overall response rate and (2) the finally selected domains for the core set at the end of round 3 in the ''all-rounds group'' vs. the ''respondents-only group.'' Secondary aspects of interest were differences between groups regarding the following: (1) the choice of the ''top six'' domains and (2) changes in the ''top six'' domains across rounds. Additionally, the design of this study enabled us to study if the results of the Delphi survey are similar when randomly selecting two independent samples. As the experiment started after round 1, we had the ability to compare domains between two panels (''respondents-only'' and ''all-rounds'') which completed an identical survey (ie, round 1). For the purpose of this study, we used descriptive statistics to describe the data, using mean (standard deviation [SD]) scores; statistical testing of between-group differences was not performed. For the 9-point Likert scale scores per domain, means and SDs were used to describe the data. For the current analysis, the last available scores were used

to compare the "respondents-only" and "all-rounds" groups if a domain was selected before the final round (eg, if a domain was selected after round 2, the last available mean is the mean of round 2). The top six domains were presented as percentages, and the change between rounds in proportion of participants showing change. Similar to the mean scores, if a domain was selected before the final round the percentage of critical votes from the last available round were used to compare the groups.

3. Results

A total of 376 participants were invited by email to partake in this Delphi survey. They were randomized into a "respondents-only group" (n = 187; 93 in the patient survey and 94 in the axSpA expert survey) and "all-rounds group" (n = 189; 95 in the patient survey and 94 in the axSpA expert survey).

3.1. Response rates

The overall response rate after three rounds was lower in the "respondents-only group" compared to the "all-rounds group" (46% [86/187] vs. 61% [116/189]). The response rate in the "respondents-only group" increased per each additional round (from 65% [122/187] to 91% [86/95]), while the response rate in the "all-rounds group" varied only slightly between rounds (from 56% [110/189] to 61% [116/189]) (Table 1). The retention rate was similar for patients and experts (Appendix Table A.1) and there was no difference between participants and nonparticipants in the continent of residence (Appendix Table A.2).

3.2. Domains selected after final round

There was no difference in mean (SD) scores between the "respondents-only" and "all-rounds" groups for any of the domains (Fig. 1).

Fig. 2 depicts the percentage of participants that voted "critical" or "not important" per domain for each of the experimental groups. The vertical lines at 15% and 70% represent the cut-offs as described in Section 2.2. This figure shows that both groups selected the same domains, apart from extra-musculoskeletal manifestations, which was only selected by the "all-rounds group."

The corresponding results from the disease modification survey can be found in Appendix Figs. A.2 and A.3.

Table 1. Response rates per group for each round of the Delphi survey

	"Respondents-only group" (N = 187)	"All-rounds group" (N = 189)
Round 1	Invited: 187 Completed: 122 Response rate: 65%	Invited: 189 Completed: 110 Response rate: 58%
Round 2	Invited: 122 Completed: 95 [response rate: 78%] Overall response rate: 51%	Invited: 189 Completed: 105 Response rate: 56%
Round 3	Invited: 95 Completed: 86 [response rate: 91%] Overall response rate: 46%	Invited: 189 Completed: 116 Response rate: 61%

3.3. Percentage of respondents regarding the top six domains

For each of the domains in the Delphi survey, Table 2 shows the percentage of participants that voted a domain as one of their six most important domains after the final round. The domains were ranked in descending order based on selection by the "respondents-only group" and matched with the same domain in the "all-rounds group." Domains in *italic* represent the top six of the "respondents-only group" and in bold of the "all-rounds group."

After the final round, the four outcomes with highest voting rates were the same in both groups, with only small differences between groups. The domain "disease activity" was voted in the top six of the "respondents-only group" (55%) but not in the top six of the "all-rounds group" (49%), where it was replaced by the domain "overall functioning and health" (57%). The differences were small; the maximum difference between groups was 10% for the domains "overall functioning and health" and "peripheral manifestations".

The corresponding results from the disease modification survey can be found in Appendix Table A.5.

3.4. Changes in the top six domains across rounds

Per individual we determined the number of domains that changed in their top six ranking across rounds. This was done from round 1 to round 2 and from round 2 to round 3 (Fig. 3). For the "all-rounds group" various combinations of completion were possible; these can be found in Appendix Table A.3. Those participants that responded to rounds 1 and 2 but not to round 3 (N = 9) were only included in the change between the rounds they completed, and the same applied to those who completed rounds 2 and 3 but not round 1 (N = 13). There were 12 participants in the "all-rounds group" who only missed round 2. For these 12 participants the change between rounds 1 and 3 was calculated. Finally, those participants that only responded to a single round were excluded from this analysis (N = 38).

Fig. 3 shows that between rounds 1 and 2, 49% of the "respondents-only group" changed at least one domain, whereas this was 38% in the "all-rounds group." In both groups hardly anyone did not change a single domain (1% in the "respondents-only" vs. 2% in the "all-rounds" group). A larger proportion of participants in the "all-rounds group" changed two domains (42%) or more than two domains (18%), compared to the "respondents-only

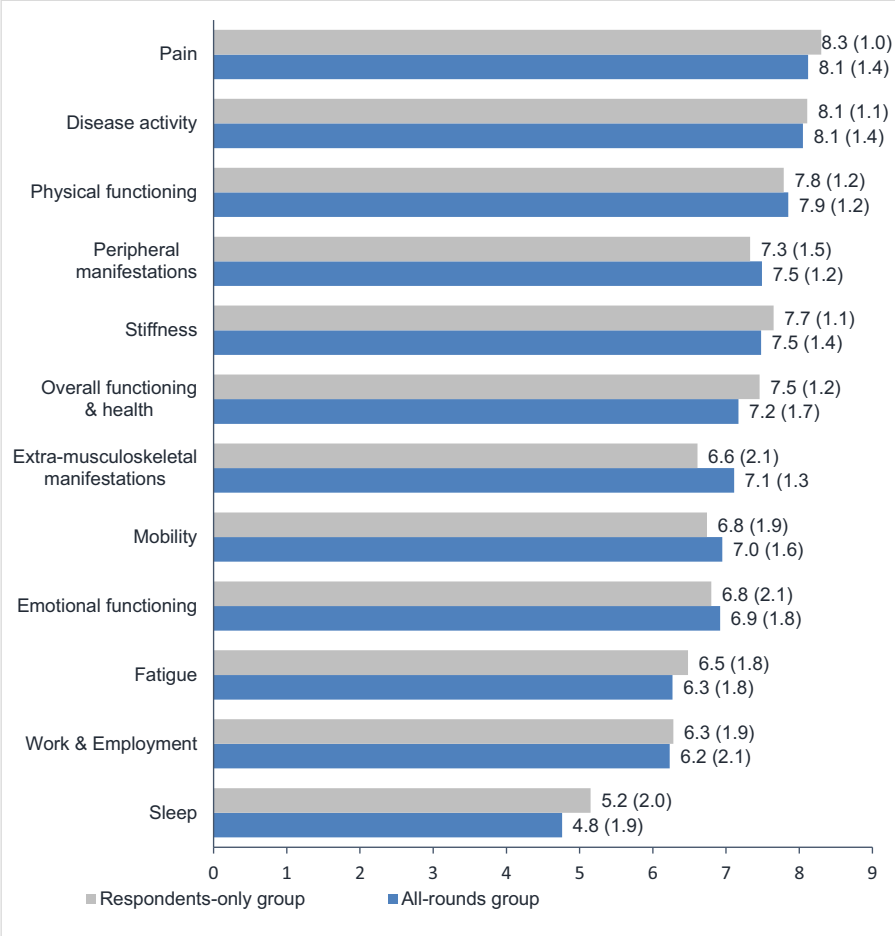

Fig. 1. Mean score (standard deviation) per domain for the ''respondents-only group'' (in gray) and ''all-rounds group'' (in blue) from the round when the domain was selected (ie, the last available scores). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

group'' (37% and 13% respectively). Between rounds 2 and 3, 45% of the participants in the ''respondents-only group'' changed one domain, whereas this was only 31% in the ''all-rounds group.'' Contrary, in the ''all-rounds group'' 22% changed more than two domains, whereas this was only 7% in the ''respondents-only group.'' The proportion of participants changing no domains or two domains was similar between groups.

The corresponding results from the disease modification survey can be found in Appendix Fig. A4.

3.5. Comparison of round 1 results

Mean (SD) scores were very similar between the ''all-rounds'' and ''respondents-only'' groups for all of the domains after the first round of the Delphi survey (Fig. 4). Additionally, the proportions of ''not important votes'' and ''critical votes'' were similar between the ''all-rounds'' and ''respondents-only'' groups after the first round (Appendix Fig. A1). Furthermore, there were no differences in the top six domains after the first round (Appendix Table A.4).

The corresponding results from the disease modification survey can be found in Appendix Table A.6 and Fig. A.5 and A.6.

4. Discussion

This study showed no differences between the ''respondents-only'' and ''all-rounds'' groups in mean (SD) scores, nor in the percentage of critical votes for any of the domains after the final round. These results showed that the same domains were selected by the ''all-rounds group'' and the ''respondents-only group.'' Invitation approach had no impact on which domains were selected after the final round. Additionally, differences between experimental groups regarding the domains in the top six were small and hardly influenced the order of importance, suggesting that the invitation approach does not influence the outcome of a Delphi exercise at all. Thus, it may be preferential to invite panel members who missed a round to a subsequent round, as this approach is less sensitive to the nonrandom loss of opinions that could lead to false consensus [6].

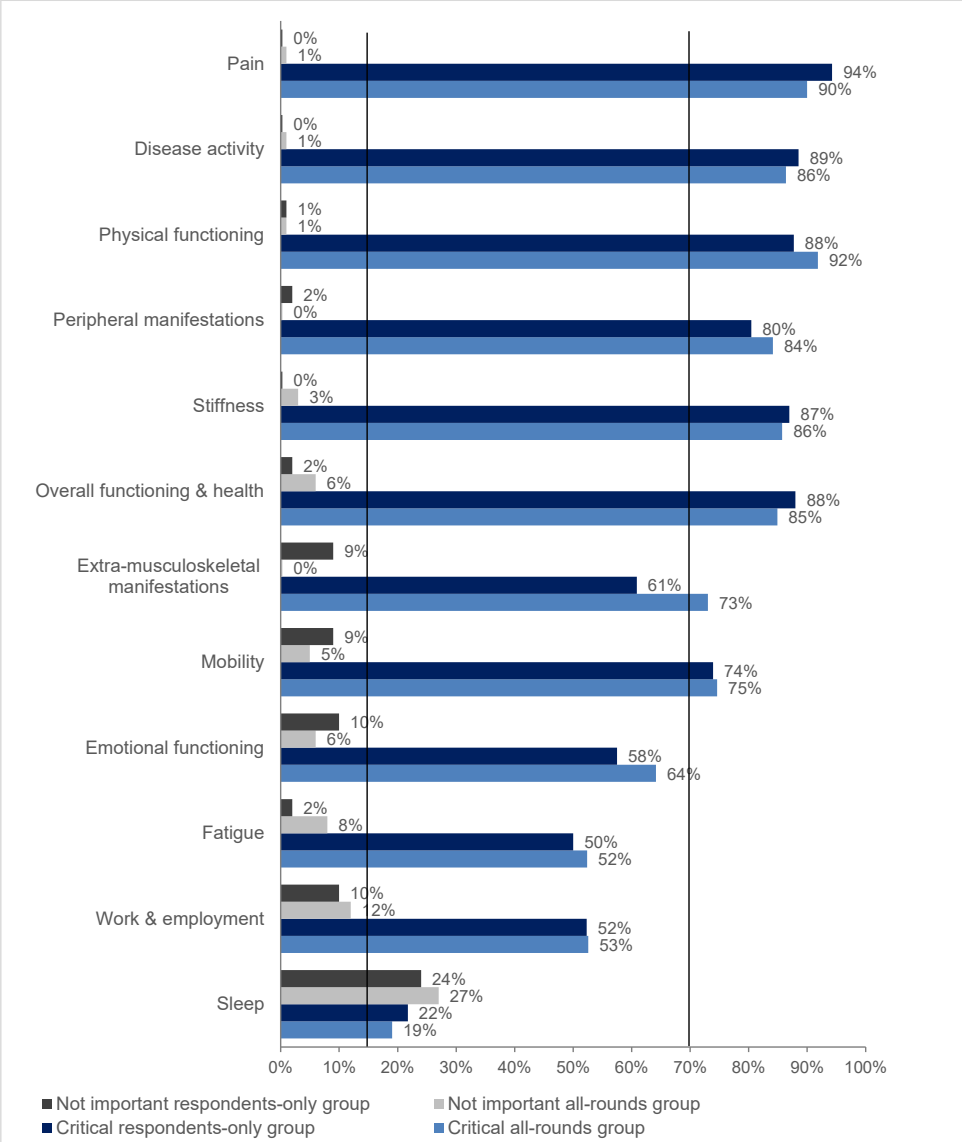

Fig. 2. Percentage of patients that voted not important (in gray) and critical (in blue) for the ''respondents-only group'' (dark color) and ''all-rounds group'' (light color) from the round when the domain was selected (ie, the last available scores); the vertical lines at 15% and 70% represent the cut-offs for inclusion of domains. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

However, we found a difference in the overall response rate, which was higher in the ''all-rounds group'' (61%) compared to the ''respondents-only group'' (46%). This was expected, because the numbers of persons invited for rounds 2 and 3 were larger in the ''all-rounds group'' than in the ''respondents-only group.''

The proportion of participants that changed one domain between rounds was larger in the ''respondents-only group''. Contrary, the proportion of participants that changed more than two domains between rounds was larger in the ''all-rounds group''. From all change categories, the no change category had the smallest proportion of participants across all rounds and in both groups. These results indicate that information of peers from previous rounds is taken into account when evaluating the domains which are deemed most important.

This study may have a few limitations. Since the data used were from a true Delphi experiment, there were no complete data on mean scores and percentages of critical votes for all domains for every round, as selected and excluded domains were not offered in the next round. We attempted to solve this by using the last available data for each of the domains (ie, from the round when the domain was selected). Nonetheless, this may have influenced the comparison between the ''all-rounds'' and ''respondents-only'' groups.

In the current study it was decided upfront to do a three-round Delphi survey, which appears sufficient to achieve

Table 2. Most important domains after round 3 for the "respondents-only group" and "all-rounds group" ranked in descending order, based on the selection by the "respondents-only group" and matched with the same domain in the "all-rounds group", and the difference in percentage of votes between the groups per domain

"Respondents-only group" ($N = 86$)		"All-rounds group" ($N = 116$)		Difference between groups	
Pain	*95%*	**Pain**	**91%**	Pain	4%
Stiffness	*62%*	**Stiffness**	**62%**	Stiffness	0%
Physical functioning	*62%*	**Physical functioning**	**61%**	Physical functioning	1%
Mobility	*59%*	**Mobility**	**54%**	Mobility	5%
Disease activity	*55%*	Disease activity	49%	Disease activity	6%
Fatigue	*50%*	**Fatigue**	**55%**	Fatigue	−5%
Overall functioning and health	47%	**Overall functioning and health**	**57%**	Overall functioning and health	−10%
Extra-musculoskeletal manifestations	44%	Extra-musculoskeletal manifestations	45%	Extra-musculoskeletal manifestations	−1%
Peripheral manifestations	34%	Peripheral manifestations	44%	Peripheral manifestations	−10%
Sleep	30%	Sleep	26%	Sleep	4%
Work and employment	21%	Work and employment	23%	Work and employment	−2%
Emotional functioning	16%	Emotional functioning	18%	Emotional functioning	−2%

Domains in *italic* represent the top 6 in the "respondents-only group" and in bold in the "all-rounds group."

consensus [14–16]. Furthermore, determining the number of rounds upfront may actually be preferential to continuing until consensus is reached, as attrition rates increase with each additional round and those with a very different opinion may drop-out, causing false consensus [6]. Due to this decision, we cannot be sure whether these results

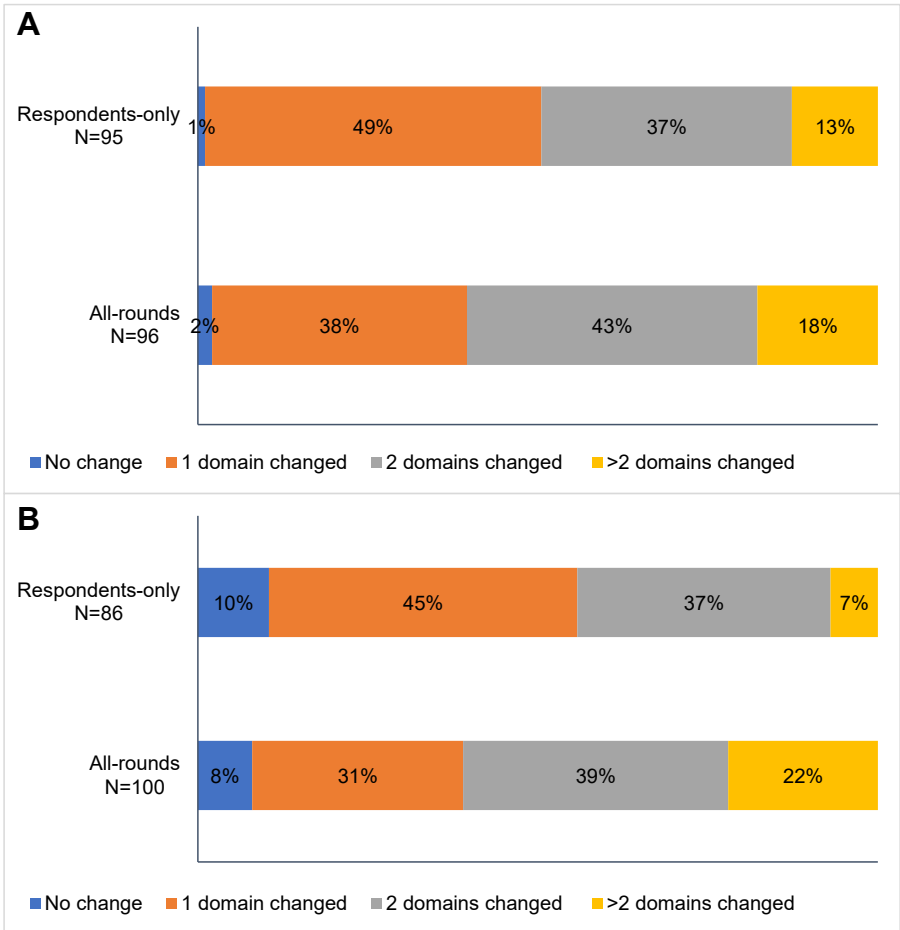

Fig. 3. Changes in the "top six" domains per group, from round 1 to round 2 (A) and round 2 to round 3 (B), presented as the proportion of participants over four change categories (no change, 1 domain changed, 2 domains changed, and >2 domains changed).

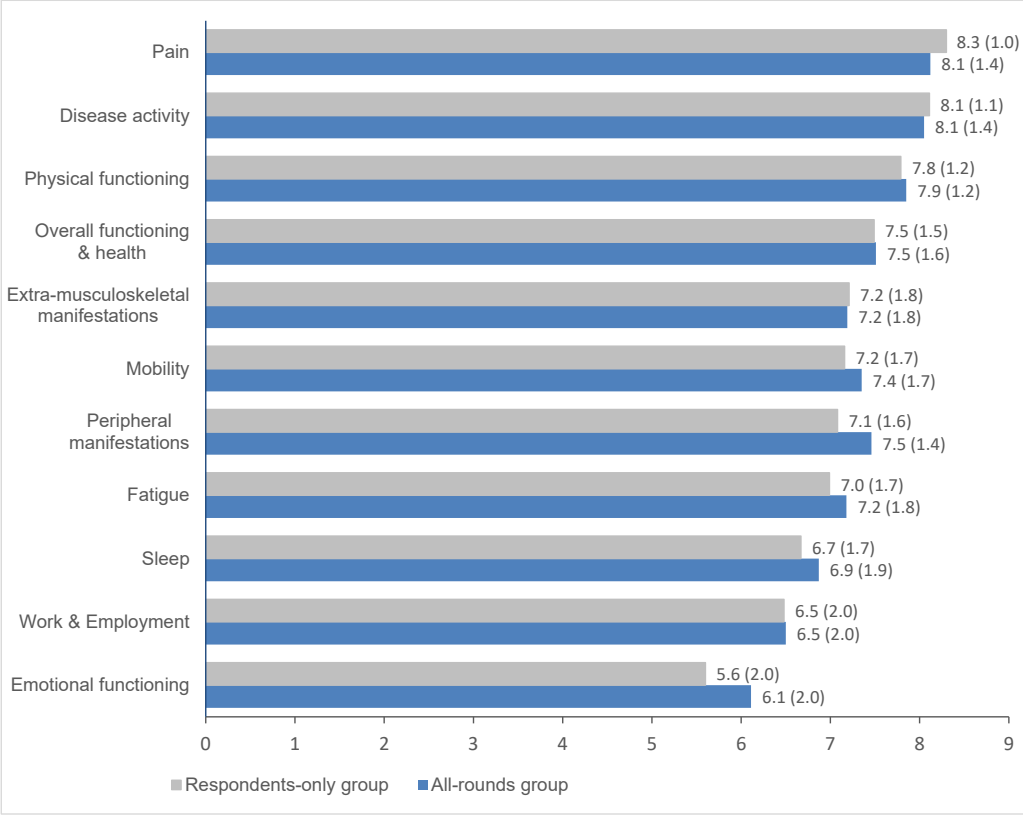

Fig. 4. Mean score (standard deviation) per domain for the ''respondents-only group'' (in gray) and ''all-rounds group'' (in blue) after round 1. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

can be extrapolated to Delphi surveys consisting of more than three rounds. Using an electronically distributed Delphi ensured involvement of international experts and patients as no travel is required, anonymity is guaranteed, and no public speaking is required, which increases patient participation [17]. Although all patients included in this study were native English speakers, as this ensured a good understanding of the content, they did represent three different countries. Furthermore, the inclusion of patients, rheumatologists, other healthcare professionals, policy makers, and representatives of pharmaceutical companies resulted in a sample that is reflective of the population who will use the updated COS.

The strength of this study was the random selection of two independent samples, namely the ''respondents-only'' and ''all-rounds'' groups. As the experiment started after round 1, we had the ability to compare results of the two panels which completed an identical first round survey. Little has been published on the agreement between multiple independent panels going through an identical survey. Previous research showed high correlations between endorsement frequencies in a replication of a Delphi survey, even with a gap of several years between replications and use of a different expert panel [18]. In more recent work groups had been randomized to assess the effect of feedback provided between rounds, and showed high agreement in selected items between different randomization groups [10]. Here we add information to this topic by showing that the results after the first round were similar between two randomly selected independent panels which completed an identical survey.

5. Conclusion

This study showed that the content of the outcome of this three-round Delphi survey was similar regardless of using data from all persons invited to the first round, or only of those persons who participated in all rounds. We therefore conclude that invitation approach does not seem to influence the final results of a Delphi survey.

Acknowledgments

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 Assessment of SpondyloArthritis international Society members for their contribution to the project, with a special thank you to the representatives of the pharmaceutical companies and EMA.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2020.09.034.

References

[1] Keeney S, McKenna H, Hasson F. The Delphi technique in nursing and health research. John Wiley & Sons; 2010.

[2] Walker A, Selfe J. The Delphi method: a useful tool for the allied health researcher. Br J Ther Rehabil 1996;3(12):677−81.

[3] McKenna HPJ. The Delphi technique: a worthwhile research approach for nursing? J Adv Nurs 1994;19(6):1221−5.

[4] Jorm AF. Using the Delphi expert consensus method in mental health research. Aust N Z J Psychiatry 2015;49(10):887−97.

[5] Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs 2000;32(4):1008−15.

[6] Humphrey-Murto S, de Wit M. The Delphi method-more research please. J Clin Epidemiol 2019;106:136−9.

[7] 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(1):e1000393.

[8] Brookes ST, Macefield RC, Williamson PR, McNair AG, Potter S, Blencowe NS, et al. Three nested randomized controlled trials of peer-only or multiple stakeholder group feedback within Delphi surveys during core outcome and information set development. Trials 2016;17(1):409.

[9] De Meyer D, Kottner J, Beele H, Schmitt J, Lange T, Van Hecke A, et al. Delphi procedure in core outcome set development: rating scale and consensus criteria determined outcome selection. J Clin Epidemiol 2019;111:23−31.

[10] MacLennan S, Kirkham J, Lam TBL, Williamson PR. A randomized trial comparing three Delphi feedback strategies found no evidence of a difference in a setting with high initial agreement. J Clin Epidemiol 2018;93:1−8.

[11] Turnbull AE, Dinglas VD, Friedman LA, Chessare CM, Sepúlveda KA, Bingham CO, et al. A survey of Delphi panelists after core outcome set development revealed positive feedback and methods to facilitate panel member participation. J Clin Epidemiol 2018;102:99−106.

[12] 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(4):1003−5.

[13] Boers M, Kirwan J, Tugwell P, Beaton D, Bingham CO, Conaghan PG. The OMERACT handbook. 2015. Available at https://www.omeract.org/omeract_publications.php.2017. Accessed August 7, 2019.

[14] Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. PLoS One 2011;6(6):e20476.

[15] Rowe G, Wright G. Expert opinions in forecasting: the role of the Delphi technique. In: Armstrong JS, editor. Principles of forecasting: a handbook for researchers and practitioners. Boston, MA: Springer US; 2001:125−44.

[16] Belton I, MacDonald A, Wright G, Hamlin I. Improving the practical application of the Delphi method in group-based judgment: A six-step prescription for a well-founded and defensible process. Technol Forecast Soc Change 2019;147:72−82.

[17] 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. Patient 2020;13(1):11−21.

[18] Ross AM, Kelly CM, Jorm AF. Re-development of mental health first aid guidelines for suicidal ideation and behaviour: a Delphi study. BMC Psychiatry 2014;14:241.