

BMJ Open Systematic review of international Delphi surveys for core outcome set development: representation of international patients

Alice Lee ¹, Anna Davies,² Amber E Young ^{3,4}

To cite: Lee A, Davies A, Young AE. Systematic review of international Delphi surveys for core outcome set development: representation of international patients. *BMJ Open* 2020;**10**:e040223. doi:10.1136/bmjopen-2020-040223

► Prepublication history for this paper is available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjopen-2020-040223>).

Received 07 May 2020

Revised 20 August 2020

Accepted 06 October 2020



© Author(s) (or their employer(s)) 2020. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

¹Academic Foundation Doctor, Department of Surgery and Cancer, Imperial College London, London, UK

²Senior Research Fellow, Centre for Academic Child Health, University of Bristol, Bristol, UK

³Consultant Paediatric Anaesthetist and Lead Children's Burns Research Centre, University Hospitals Bristol NHS Foundation Trust, Bristol, UK

⁴Senior Research Fellow, Bristol Centre for Surgical Research, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Correspondence to

Dr Amber E Young;
amber.young1@nhs.net

ABSTRACT

Objectives A core outcome set (COS) describes a minimum set of outcomes to be reported by all clinical trials of one healthcare condition. Delphi surveys are frequently used to achieve consensus on core outcomes. International input is important to achieve global COS uptake. We aimed to investigate participant representation in international Delphi surveys, with reference to the inclusion of patients and participants from low and middle income countries as stakeholders (LMICs).

Design Systematic review.

Data sources EMBASE, Medline, Web of Science, COMET database and hand-searching.

Eligibility criteria Protocols and studies describing Delphi surveys used to develop an international COS for trial reporting, published between 1 January 2017 and 6 June 2019.

Data extraction and synthesis Delphi participants were grouped as patients or healthcare professionals (HCPs). Participants were considered international if their country of origin was different to that of the first or senior author. Data extraction included participant numbers, country of origin, country income group and whether Delphi surveys were translated. We analysed the impact these factors had on outcome prioritisation.

Results Of 90 included studies, 69% (n=62) were completed and 31% (n=28) were protocols. Studies recruited more HCPs than patients (median 60 (IQR 30–113) vs 30 (IQR 14–66) participants, respectively). A higher percentage of HCPs was international compared with patients (57% (IQR 37–78) vs 20% (IQR 0–68)). Only 31% (n=28) studies recruited participants from LMICs. Regarding recruitment from LMICs, patients were under-represented (16% studies; n=8) compared with HCPs (22%; n=28). Few (7%; n=6) studies translated Delphi surveys. Only 3% studies (n=3) analysed Delphi responses by geographical location; all found differences in outcome prioritisation.

Conclusions There is a disproportionately lower inclusion of international patients, compared with HCPs, in COS-development Delphi surveys, particularly within LMICs. Future international Delphi surveys should consider exploring for geographical and income-based differences in outcome prioritisation.

PROSPERO registration number CRD42019138519.

Strengths and limitations of this study

- We conducted a comprehensive systematic search of the literature including three search engines, the Core Outcome Measures in Effectiveness Trials database and hand-searching.
- This is the first review to present stakeholder-level data on the country of origin and income status of Delphi survey participants for core outcome set studies.
- Study reporting on participant demographics was inconsistent and some data were not in a compatible format for data analysis.
- The literature search was time limited (January 2017 to June 2019) and restricted to English language articles.

INTRODUCTION

A core outcome set (COS) is a minimum group of outcomes to be reported in all trials of a specific health condition. Development is undertaken to reduce the heterogeneity of outcome reporting across trials and to enable study results to be compared and combined to inform best medical practice. Core outcomes are identified scientifically by stakeholders as being the most important in determining the effects of an intervention or treatment in one healthcare condition.^{1–3} Consensus in outcome prioritisation among stakeholders in COS development is often achieved using Delphi surveys.^{2,4} A diverse group of stakeholders is recommended to be recruited, including patients, healthcare professionals (HCPs), trialists, regulators, industry representatives, policymakers, researchers and the public.² The Delphi process comprises iterative rounds of surveys in which the importance of outcomes is rated.² After each round, the participants' individual responses, and those of other stakeholders, are fed back in an anonymised manner, so that they can be reconsidered before the next round in an

aim to achieve consensus. The Delphi method is advantageous because it incorporates the views of various stakeholder groups and can be conducted electronically ('e-Delphi') to facilitate international participation.^{1 2 4}

International participation in Delphi surveys is important for the COS to be applicable in global healthcare settings, and because widespread uptake of the COS will facilitate the future synthesis of trial evidence on an international scale.² An increasing number of COS developers are including international participants.⁵⁻⁹³ A recent survey reported that approximately 50% of published COS projects from the last 5 years included participants from two or more countries.⁹² Despite this increase in international stakeholder participation, there is no agreement on how study methodology should be adapted to facilitate such input. The Core Outcome Measures in Effectiveness Trials (COMET) Handbook highlights the logistical and organisational challenges of international COS development projects as well as issues regarding generalisability of small international participant numbers.² The effect of income group and/or geographical location of Delphi participants on core outcome prioritisation has not been systematically investigated, despite international variation in healthcare resources, biomedical beliefs and burden of disease.^{5 73 81}

Using a systematic review of the literature, this study aims to explore participant representation in international Delphi surveys used for COS development. Part of our analysis will explore how COS projects undertaking international Delphi surveys evaluate the impact of participants from countries from different World Bank income groups on prioritising the importance of outcomes.

Methods

This systematic review adheres to a prespecified protocol and the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.⁹⁴ The protocol for this review was registered on PROSPERO (available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42019138519).

Identification of studies

Study eligibility

COS development studies/protocols were included where international participants took part in a Delphi survey to prioritise outcomes for a COS, published between 1 January 2017 and 6 June 2019. The inclusion and exclusion criteria are shown in [table 1](#). This time period was chosen to reflect the 2017 publication of the COMET Handbook² to maximise study homogeneity. Protocol studies were included for methodological detail if equivalent full-length studies were not available, because the review focused on evaluating methodology, not study outcomes. Full-text and protocol studies describing the same international Delphi survey were merged as one reference.

Types of participants

For inclusion in the review, at least one participant in the Delphi survey had to be international, defined as a different country of origin to the first or senior author. Participants were included if they completed at least round 1 of the Delphi survey. Those who registered but did not participate were not included.

Types of interventions

Included studies used an international Delphi survey for prioritisation of outcomes at some point in the consensus process (including modified Delphi surveys and mixed methods approaches). Any study intervention and comparator were acceptable.

Types of outcome

Included studies used an international Delphi survey to prioritise any outcomes for a health condition or intervention to form a COS.

Search strategy

A search string was developed to identify relevant papers. These included key search terms and their synonyms (eg, COS, international, Delphi) and relevant medical subject

Table 1 Study inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
<ol style="list-style-type: none"> 1. Full length, peer-reviewed original studies using the Delphi method with international participants to develop a COS 2. Protocols for studies using international Delphi surveys to develop a COS (may be non-peer-reviewed) 	<ol style="list-style-type: none"> 1. Articles that were not published within the accepted date range: 1 January 2017–6 June 2019 2. Articles not written in English 3. Articles that were not full-text and peer-reviewed (with the exception of protocol studies) 4. Articles that do not report development of a COS for a medical condition or intervention for the purposes of clinical trial reporting 5. Articles that do not report using a Delphi survey 6. Articles that do not report using an 'international' Delphi survey (members of at least one of the following groups should be described as international: patients, patient representatives or healthcare professionals)

COS, core outcome set.

headings. The search string for Ovid MEDLINE is shown in online supplemental appendix 1 and was adapted for different databases (Ovid EMBASE, Web of Science).

Study selection process

Search results were compiled using Mendeley (V.1.19.4). Citations were deduplicated using Mendeley software and manually. Article screening was undertaken by one researcher (AL) against prespecified inclusion and exclusion criteria (table 1) in two stages (title and abstract and full text). For both stages, a second researcher (AD) independently assessed 20% of the screening results. Inter-rater reliability between researchers was assessed with Cohen's kappa.⁹⁵ If discrepancies in article selection could not be resolved, a third researcher (AY) was consulted.

Quality assessment

A risk of bias assessment was not undertaken because the review aimed to assess study methodology and not the effect of study interventions. There is currently no risk of bias assessment tool for COS development or Delphi surveys, and tools for assessing risk of bias in trials or observational studies are not applicable to these reviewed studies.

Data extraction

Data were extracted using a piloted data extraction form (Microsoft Excel) developed for the purpose of the review. Data were extracted under the following domains: study details (year of publication, full text/protocol study and COMET disease category), participant numbers, international status of participants, participant geographical location and income group, the effect of these on prioritisation and whether the Delphi survey was translated.

Number of Delphi participants overall and per stakeholder group

Participants were grouped into two stakeholder groups: patients and representatives (carers and representatives from patient organisations) and HCPs (medical professionals, trialists, regulators, industry representatives, policymakers and researchers). Data were extracted on number of participants (total and per stakeholder group). We recorded number of participants per study based first on the number of participants from the Delphi round which included both patients and HCPs. If both stakeholder groups were included throughout or the study only included either patients or HCPs from the outset, we extracted number of participants from the Delphi round with the largest number of participants.

International status of participants and effect on outcome prioritisation

Participants were categorised as international if their country of origin was not the same as either the first or senior author of the study. Other demographic data included participants' countries of origin and the World Bank world regions and World Bank income groups represented by these countries. All study texts were scrutinised

for any description of analysis of Delphi responses by geographical location or income status, and if so, the outcome of this analysis.

Delphi survey translation

Studies that recruited participants from non-English-speaking countries were scrutinised for any description of survey translation, and if so, details of the translated languages and method of translation.

Data synthesis

Data from individual studies were tabulated by one author (AL). A second researcher (AD) independently extracted data from 20% of included studies. Inter-rater reliability was assessed with Cohen's kappa.⁹⁵ If discrepancies in data extraction could not be resolved, a third researcher (AY) was consulted. Quantitative, non-parametric data were analysed using Microsoft Excel to calculate medians and IQRs for number of participants, percentage of international participants and number of countries of origin, World Bank world regions and World Bank income groups. Categorical data were described narratively, including participant countries of origin, survey translation and the outcomes of analysis of responses by geographical location or income status.

Patient and public involvement

No patients were involved in this review of previously published data.

RESULTS

Identification of studies

The electronic search identified 529 non-duplicate citations, of which 90 were included in the final data set (figure 1). Cohen's kappa demonstrated very good (0.81) and good (0.71) agreement between researchers performing screening at the abstract and full-text stages, respectively.⁹⁶ Of the 90 included studies (online supplemental appendix 3), 69% (n=62) were completed and 31% (n=28) were study protocols. The greatest number of Delphi studies were published from the UK (42%; n=38; online supplemental appendix 3) and the three most frequent COMET disease categories were pregnancy and childbirth (14%, n=13), gastroenterology (9%, n=8) and orthopaedics and trauma (9%, n=8); online supplemental appendix 4.

Number of Delphi participants overall and per stakeholder group

The median number of participants per study was 100 (table 2). Most studies (77%; n=69) included both patients and HCPs. Of the 23% of studies (n=21) with only one stakeholder group, 95% (n=20) recruited only international HCPs. Of all studies, 70% (n=48) recruited international participants in both stakeholder groups, 19% (n=13) recruited into only one stakeholder group (92% of which recruited only HCPs) and in 12% (n=8), the status of stakeholder groups was unclear. Studies

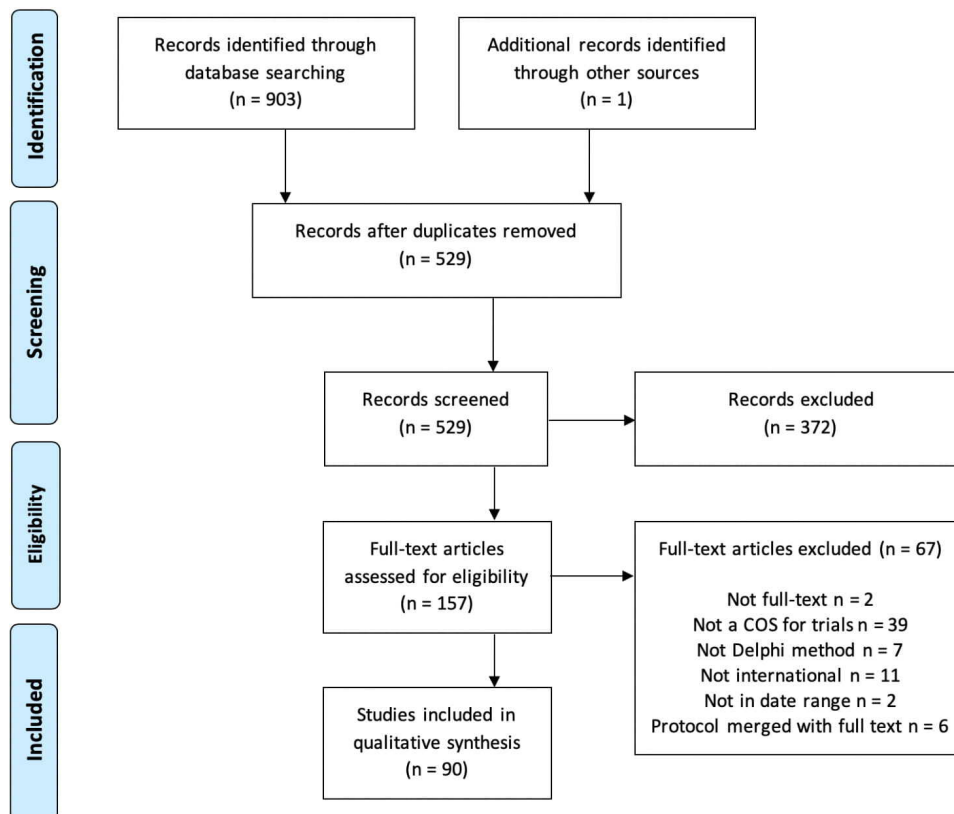


Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram. COS, core outcome set.

recruited two times as many HCPs as patients (median 60 vs 30 participants).

International status of participants and effect on outcome prioritisation

The median percentage of international participants per study was 52%. Studies recruited three times more international HCPs than international patients (57 vs 20%).

The total number of countries represented across the included studies was 95 for HCPs and 46 for patients. Within these studies, the median number of countries represented in each Delphi was 11 for HCPs and 2 for patients.

Participants were recruited from every World Bank world region for both stakeholder groups. HCPs

Table 2 Demographic data for the overall cohort, patients and HCPs

Population	Median statistic (IQR)				
	Number of participants, n	Percentage international participants, %	Number of countries of origin, n	Number of World Bank world regions, n	Number of World Bank income groups, n
Overall cohort	100 (38–166)*	52 (8–83)†	12 (8–21)‡	4 (3–5)‡	2 (1–3)§
HCPs	60 (30–113)¶	57 (37–78)**	11 (8–17)††	4 (3–5)‡	2 (1–3)‡‡‡
Patients	30 (14–66)§§§§	20 (0–68)¶¶¶¶	2 (1–4)††	2 (1–3)‡	1 (1–2)‡‡‡

Median statistics for number of participants, percentage international participants and number of countries of origin, World Bank world regions and World Bank income groups.

*Four protocol studies did not specify the predicted number of participants.

†Data incomplete, unclear or not specified in 54 studies.

‡Data incomplete, unclear or not specified in 44 studies.

§Data incomplete, unclear or not specified in 48 studies.

¶Data incomplete, unclear or not specified in nine studies.

**Data incomplete, unclear or not specified in 57 studies.

††Data incomplete, unclear or not specified in 46 studies.

‡‡Data incomplete, unclear or not specified in 43 studies.

§§Data incomplete, unclear or not specified in 29 studies.

¶¶Data incomplete, unclear or not specified in 45 studies.

HCP, healthcare professional.

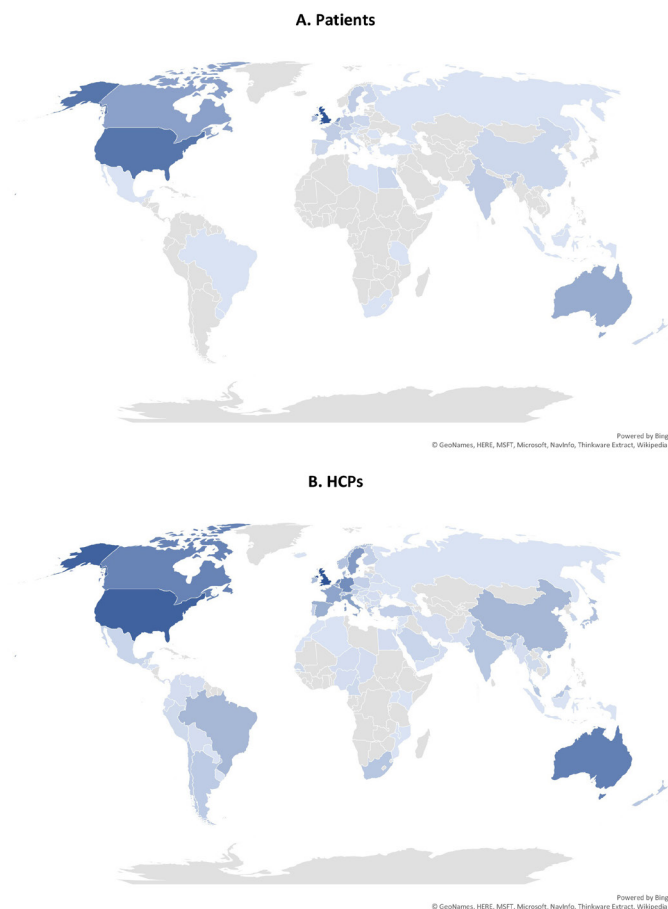


Figure 2 World map showing participant countries of origin for (A) patients and (B) HCPs. The colour gradient represents the percentage of studies which recruited patients and HCPs from each country. For patients, this ranged from 2% (light blue) to 28% (dark blue) and for HCPs this ranged from 1% (light blue) to 39% (dark blue). HCPs, healthcare professionals.

represented two times as many World Bank world regions as patients (four vs two regions). The most frequent countries of participant origin were the USA and UK for HCPs and patients (figure 2). The most frequent World Bank world regions reported for both HCPs and patients were Europe and Central Asia, followed by North America and East Asia and the Pacific (figure 3). Compared with HCPs, fewer studies recruited patients from certain world regions (Sub-Saharan Africa, Latin America and the Caribbean and East Asia and the Pacific) (figure 3). Only 4% studies (n=2) recruited patients from Sub-Saharan Africa when compared with studies recruiting HCPs (13%; n=10).

Most studies recruited participants from high-income countries (48%; n=43), followed by high–middle income (28%; n=25), lower middle income (16%; n=15) and low income countries (9%; n=8; figure 3). HCPs were recruited from two times as many World Bank income groups as patients (two vs one groups). Less than half as many studies recruited patients from low-income (2%, n=1) and lower middle income countries (6%; n=3) when

compared with studies recruiting HCPs (5%, n=4% and 14%, n=11, respectively).

A minority of studies (4%, n=4) analysed, or stated an intention to analyse, the Delphi survey responses by participants' geographical location, either by country^{54 97} or continent.⁷³ Some differences in outcome prioritisation were minor, not affecting the final COS.^{54 97} Park *et al* presented results from round 1 of their Delphi survey on patient-reported outcomes (PROs) for adult myositis.⁵⁴ They found that, unlike participants from the USA and South Korea, Swedish participants rated 'impact on household activity' less favourably, although this outcome was still retained for the second Delphi round. Sautenet *et al* reported consensus outcomes for kidney transplantation among patients, caregivers and HCPs.⁹⁷ They found that patients/caregivers from certain countries ranked 'depression' or 'cognition' as less important and that 'skin cancer' received greater prioritisation in countries with public campaigns for prevention. Since none of the aforementioned outcomes were ranked in the 'top eight' of either patients or HCPs, these differences did not affect the final COS. Van Rijssen *et al*⁷³ reported consensus PROs for pancreatic cancer among European, North American and Asian participants. In this study, the outcomes in the final COS would have been different if the responses were analysed by continent rather than as a whole cohort. In comparison to the whole cohort, European participants reached consensus 'in' on three additional outcomes, American participants reached consensus 'in' for one additional outcome, but did not reach consensus for two outcomes included in the final COS, and Asian participants did not reach consensus on any PROs included in the final COS. A fourth study (protocol) stated an intention to analyse Delphi survey results by 'language and cultural variation'.¹⁷ None of the included studies analysed Delphi responses by income status.

Delphi survey translation

Delphi surveys were translated in a minority of studies that recruited participants from countries in which English was not the first language (16%, n=6/38). This included five studies with both stakeholder groups and one with patients only. Three studies translated the Delphi survey into all languages spoken by the patients^{54 73 83} and a protocol study mentioned translating the Delphi survey 'as required'.⁹⁸ One study translated the Delphi survey from English into Dutch only, despite recruiting patients from many other countries.²⁷ One study translated the Delphi survey from English to Italian but did not list countries of origin for patients.⁹⁹ Some studies (n=14) excluded non-English-speaking participants, despite recruiting patients and representatives from many countries where English is not the first language.^{8 13 18 19 21 22 25 36 55 56 59 60 69 79}

DISCUSSION

The findings of this systematic review demonstrate that most international Delphi surveys for COS development

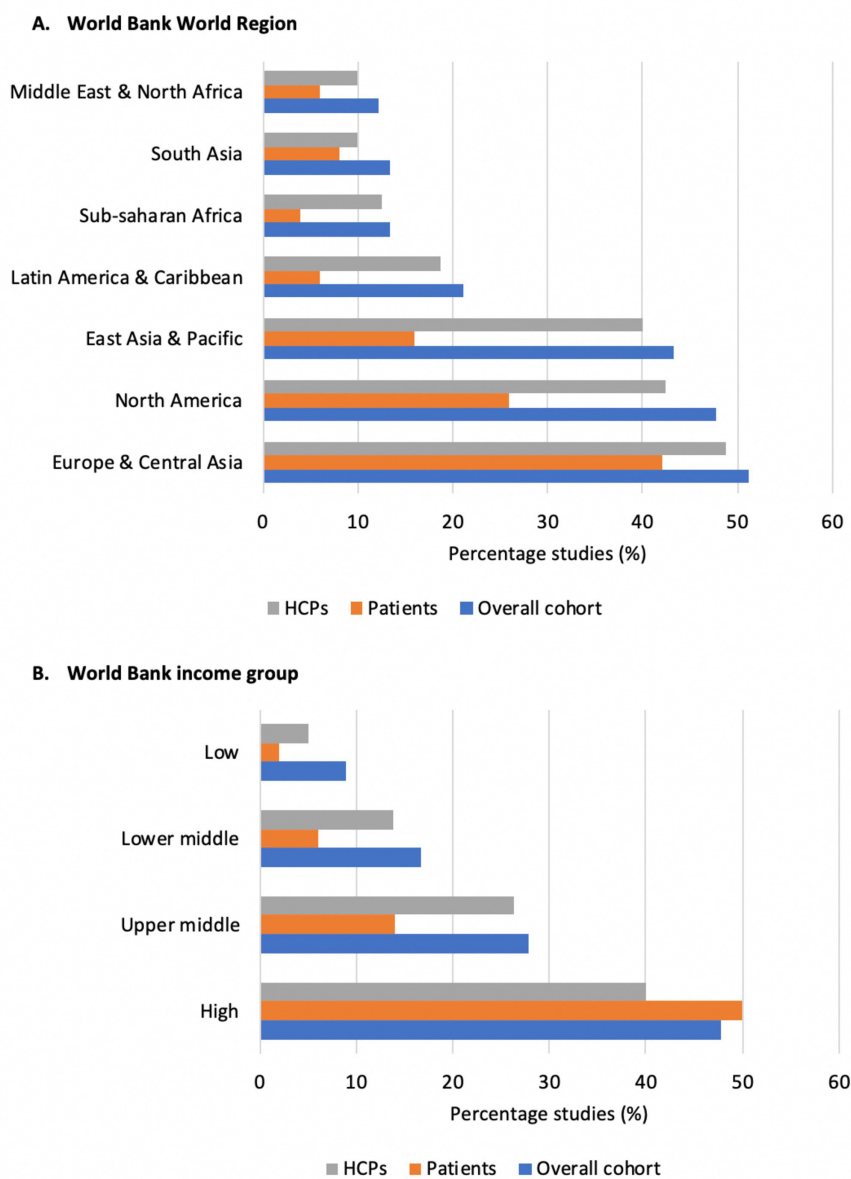


Figure 3 Distribution of (A) World Bank world regions and (B) World Bank income groups of participants by overall cohort and stakeholder group. HCPs, healthcare professionals.

recruited both HCPs and patients. Patients were recruited in fewer numbers and were less likely to be international and especially from LMICs. A minority of studies altered Delphi survey language, despite recruiting participants from several countries with non-English first languages. Importantly, few studies analysed Delphi responses by geographical location of participants, but those that did found differences in outcome prioritisation.

These findings echo those of annual systematic reviews of COS development studies, demonstrating lower overall recruitment of Delphi participants from LMICs.^{5 6} Our review adds new, stakeholder-level information, which has identified a disproportionately lower recruitment of patients (overall and particularly within LMICs). There are various possible explanations for this, including lower English language proficiency, reduced internet access

(required for e-Delphi surveys and online recruitment), differing biomedical beliefs and lack of resources and/or time for voluntary research participation.^{100–102} Development of contextual methods to effectively engage international patients from LMICs is required.

Lack of Delphi survey translation is an important barrier to participation, particularly affecting patients from LMICs. Most included studies did not adapt survey language for participants and many excluded non-English speakers. This could have introduced recruitment bias, particularly within LMICs. Researchers have expressed concern that Delphi survey translation could result in loss of comprehension or meaning.² Some of the included studies have approached this problem by translating outcomes by native-speaking professionals only,⁹⁸ using an online multilingual interface¹⁰³ or by discussing

translated outcomes with patient research partners.⁵⁴ Van Rijssen *et al*⁷³ used a method of forward and backward translation, that is, translated the survey from English to the required participant languages and then back into English. Discrepancies between the forward and backward translations were resolved by consensus between the translators and if necessary, the project team. This is the method recommended by the COMET handbook,² WHO¹⁰⁴ and MAPI Institute¹⁰⁵ for providing equivalent conceptual, cultural and semantic meaning and should be encouraged as the first-line method after consideration of the costs involved.

Increasing the international participation of Delphi surveys for COS development should be encouraged but raises considerations for study design. For example, how to ensure Delphi studies have appropriate and adequate international representation, the criteria to define this (geographically and/or using income status) and how to analyse international data. None of the included studies analysed Delphi responses by income status and few analysed Delphi responses by geographical location of participants. Those that did found differences in prioritisation between participants of different countries of origin^{54 60} and continents of origin.⁷³ Importantly, in the study by Van Rijssen *et al*,⁷³ which reported a core set of PROs for pancreatic cancer, participants from Europe, the USA and Asia did not reach the same consensus on the final COS, and Asian participants did not reach consensus on any PROs included in the final set. There is no clear explanation for these discrepancies. For COSs to have global uptake, it is important that the selected outcomes are representative of all international stakeholders. Performing subanalyses of Delphi responses by participant location and/or income status should be encouraged as a useful indication of applicability across different populations, and significant differences in prioritisation should be explored.

Another important finding of this systematic review was inconsistent study reporting. In some studies, basic demographic information was unclear or not specified including which stakeholder groups were international and from which countries and world regions they were recruited. Many studies only listed the most frequent countries of origin, provided data in non-standard formats, for example, reporting participant numbers for the USA and Canada combined, or only provided geographic detail for the whole cohort instead of per stakeholder group. Some studies provided demographic data on invited participants, rather than those who had completed at least one round of the Delphi survey. As a result, we were unable to use many studies' data in our analyses. Without adequate demographic information, it is difficult to interpret the applicability of COSs across different populations. Current reporting checklists for COS development projects, such as Core Outcome Set–STANDards for Reporting (COS-STAR),¹⁰⁶ could be adapted to reflect international participation.

The findings of this review should be interpreted in the context of its limitations. The HCP group consisted

of various stakeholders with potentially differing views, including healthcare workers, trialists, regulators, industry representatives, policymakers and researchers. Not all studies included each of these constituent groups. For ease of comparison, and because the largest distinction in opinion was likely to be between HCPs and patients, all professionals were combined into one group. This approach has been used previously.¹⁰⁷ We described patients as international if their country of origin was different to the affiliated country of either the first or last author of the study. Some authors had multiple affiliations, which may have reduced the apparent percentage of international participants. We used the World Bank world regions and income groups to categorise demographic information from participants, but some studies did not present their demographic data in a compatible format. Furthermore, we restricted our search to articles available in English and published within a recent time frame.

CONCLUSION

In conclusion, we have demonstrated that Delphi surveys used to develop COSs for clinical trials recruit fewer international patients than HCPs, particularly from LMICs. This could be contributed to by a lack of Delphi survey translation for non-English speakers. Few studies explored any geographical variation in responses; those that did found differences in outcome prioritisation. This review highlights complex issues that need further discussion, for example, how to define adequate international participation and how to analyse international data, including geographical discrepancies in outcome selection. Future studies should consider exploratory analyses of Delphi responses by geographical location or income status to assess applicability of the COS across populations. Study reporting was inconsistent and could be improved with a standardised checklist for international Delphi surveys.

Twitter Alice Lee @AliceEALee

Contributors AY devised the initial review question and provided critical review of the search strategy, data collection process and manuscript. AL designed the search strategy and was the primary researcher involved in screening search results, data extraction and drafting of the manuscript. AD provided critical review of the search strategy and acted as a second researcher for purposes of screening and data extraction, as well as critically reviewing the manuscript.

Funding This research (AL) was funded by The Scar Free Foundation. The Scar Free Foundation is the only medical research charity focused on scarring with the mission to achieve scar free healing within a generation. This study was also supported by the NIHR Biomedical Research Centre at University Hospitals Bristol and Weston NHS Foundation Trust and the University of Bristol. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Map disclaimer The depiction of boundaries on this map does not imply the expression of any opinion whatsoever on the part of BMJ (or any member of its group) concerning the legal status of any country, territory, jurisdiction or area or of its authorities. This map is provided without any warranty of any kind, either express or implied.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in the supplementary material. Requests for data not included in the article or supplementary material can be made to the corresponding author.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

ORCID iDs

Alice Lee <http://orcid.org/0000-0002-8492-837X>

Amber E Young <http://orcid.org/0000-0001-7205-492X>

REFERENCES

- Williamson PR, Altman DG, Blazeby JM, *et al*. Developing core outcome sets for clinical trials: issues to consider. *Trials* 2012;13:1–8.
- Williamson PR, Altman DG, Bagley H, *et al*. The comet Handbook: version 1.0. *Trials* 2017;18:1–50.
- Clarke M, Williamson PR. Core outcome sets and systematic reviews. *Syst Rev* 2016;5:11.
- Hall DA, Smith H, Heffernan E, *et al*. Recruiting and retaining participants in e-Delphi surveys for core outcome set development: evaluating the COMIT'ID study. *PLoS One* 2018;13:e0201378.
- Davis K, Gorst SL, Harman N, *et al*. Choosing important health outcomes for comparative effectiveness research: an updated systematic review and involvement of low and middle income countries. *PLoS One* 2018;13:e0190695.
- Gargon E, Gorst SL, Harman NL, *et al*. Choosing important health outcomes for comparative effectiveness research: 4th annual update to a systematic review of core outcome sets for research. *PLoS One* 2018;13:e0209869:1–15.
- Haller G, Bampoe S, Cook T, *et al*. Systematic review and consensus definitions for the standardised endpoints in perioperative medicine initiative: clinical indicators. *Br J Anaesth* 2019;123:228–37.
- Butler DP, De la Torre A, Borschel GH, *et al*. An international collaborative standardizing patient-centered outcome measures in pediatric facial palsy. *JAMA Facial Plast Surg* 2019;21:351–8.
- Byrne M, O'Connell A, Egan AM, *et al*. A core outcomes set for clinical trials of interventions for young adults with type 1 diabetes: an international, multi-perspective Delphi consensus study. *Trials* 2017;18:602.
- Cho Y, Sautenet B, Rangan G, *et al*. Standardised outcomes in Nephrology—Polycystic kidney disease (SONG-PKD): study protocol for establishing a core outcome set in polycystic kidney disease. *Trials* 2017;18:560.
- Dadouch R, Faheim M, Juando-Prats C, *et al*. Development of a core outcome set for studies on obesity in pregnant patients (COSSOPP): a study protocol. *Trials* 2018;19:665.
- Danese S, Bonovas S, Lopez A, *et al*. Identification of Endpoints for Development of Antifibrosis Drugs for Treatment of Crohn's Disease. *Gastroenterology* 2018;155:76–87.
- Dorman SL, Shelton JA, Stevenson RA, *et al*. Management of medial humeral epicondyle fractures in children: a structured review protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. *Trials* 2018;19:119.
- Dos Santos F, Drymiotou S, Antequera Martin A, *et al*. Development of a core outcome set for trials on induction of labour: an international multistakeholder Delphi study. *BJOG: Int J Obstet Gyn* 2018;125:1673–80.
- Callis Duffin K, Merola JF, Christensen R, *et al*. Identifying a core domain set to assess psoriasis in clinical trials. *JAMA Dermatology* 2018;154:1137–44.
- Duffy JMN, Bhattacharya S, Curtis C, *et al*. A protocol developing, disseminating and implementing a core outcome set for infertility. *Human Reproduction Open* 2018;2018:hoy007.
- Alkhaffaf B, Bruce I, Glennly AM, *et al*. The GASTROS study: standardising outcome reporting in gastric cancer surgery research. *Eur J Cancer* 2017;72:S134.
- Evangelidis N, Tong A, Manns B, *et al*. Developing a set of core outcomes for trials in hemodialysis: an international Delphi survey. *American Journal of Kidney Diseases* 2017;70:464–75.
- Fish R, Sanders C, Adams R, *et al*. A core outcome set for clinical trials of chemoradiotherapy interventions for anal cancer (CORMAC): a patient and health-care professional consensus. *The Lancet Gastroenterology & Hepatology* 2018;3:865–73.
- Foust-Wright C, Wissig S, Stowell C, *et al*. Development of a core set of outcome measures for OAB treatment. *Int Urogynecol J* 2017;28:1785–93.
- Gagnier JJ, Page MJ, Huang H, *et al*. Creation of a core outcome set for clinical trials of people with shoulder pain: a study protocol. *Trials* 2017;18:336.
- Hall DA, Smith H, Hibbert A, *et al*. The COMIT'ID Study: Developing Core Outcome Domains Sets for Clinical Trials of Sound-, Psychology-, and Pharmacology-Based Interventions for Chronic Subjective Tinnitus in Adults. *Trends in Hearing* 2018;22:233121651881438.
- Harman NL, Wilding J, Curry D, *et al*. Selecting core outcomes for randomised effectiveness trials in type 2 diabetes (SCORE-IT): study protocol for the development of a core outcome set. *Trials* 2018;19:427.
- Haywood K, Whitehead L, Nadkarni VM, *et al*. COSCA (core outcome set for cardiac arrest) in adults: an Advisory statement from the International liaison Committee on resuscitation. *Resuscitation* 2018;127:147–63.
- Healy P, Gordijn S, Ganzevoort W, *et al*. Core outcome set for growth restriction: deVeloPping endpoints (COSGROVE). *Trials* 2018;19:451.
- Hodgson CL, Turnbull AE, Iwashyna TJ, *et al*. Core domains in evaluating patient outcomes after acute respiratory failure: international multidisciplinary clinician consultation. *Phys Ther* 2017;97:168–74.
- Horbach SER, van der Horst CMAM, Blei F, *et al*. Development of an international core outcome set for peripheral vascular malformations: the OVAMA project. *Br J Dermatol* 2018;178:473–81.
- Allin BSR, Bradnock T, Kenny S, *et al*. NETS^{1HD} study: development of a Hirschsprung's disease core outcome set. *Arch Dis Child* 2017;102:1143–51.
- Kampstra NA, Grutters JC, van Beek FT, *et al*. First patient-centred set of outcomes for pulmonary sarcoidosis: a multicentre initiative. *BMJ Open Respiratory Research* 2019;6:e000394.
- Kelly A, Tong A, Tymms K, *et al*. Outcome Measures in Rheumatology - Interventions for medication Adherence (OMERACT-Adherence) Core Domain Set for Trials of Interventions for Medication Adherence in Rheumatology: 5 Phase Study Protocol. *Trials* 2018;19:204.
- Kenny KP, Day PF, Sharif MO, *et al*. What are the important outcomes in traumatic dental injuries? an international approach to the development of a core outcome set. *Dental Traumatology* 2018;34:4–11.
- Khalil A, Beune I, Hecher K, *et al*. Consensus definition and essential reporting parameters of selective fetal growth restriction in twin pregnancy: a Delphi procedure. *Ultrasound Obstet Gynecol* 2019;53:47–54.
- Kim AH, Roberts C, Feagan BG, *et al*. Developing a standard set of patient-centred outcomes for inflammatory bowel Disease—an international, Cross-disciplinary consensus. *Journal of Crohn's and Colitis* 2018;12:408–18.
- Knaapen M, Hall NJ, van der Lee JH, *et al*. Establishing a core outcome set for treatment of uncomplicated appendicitis in children: study protocol for an international Delphi survey. *BMJ Open* 2019;9:e028861.
- Leo DG, Leong WY, Gambling T, *et al*. The outcomes of Perthes' disease of the hip: a study protocol for the development of a core outcome set. *Trials* 2018;19:374.
- Ma C, Panaccione R, Fedorak RN, *et al*. Development of a core outcome set for clinical trials in inflammatory bowel disease: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. *BMJ Open* 2017;7:e016146.

- 37 Mackie SL, Twohig H, Neill LM, *et al.* The OMERACT core domain set for outcome measures for clinical trials in polymyalgia rheumatica. *J Rheumatol* 2017;44:1515–21.
- 38 MacLennan S, Williamson PR, Bekema H, *et al.* A core outcome set for localised prostate cancer effectiveness trials. *BJU Int* 2017;120:E64–79.
- 39 Sherratt F, Bagley H, Stones S, *et al.* Improving core outcome set development for children and young people: learning from a case study in acute appendicitis and consultation with an international group of children and young people. *J Evid Based Med* 2019;12:21.
- 40 Manera KE, Tong A, Craig JC, *et al.* Standardized outcomes in Nephrology—Peritoneal dialysis (SONG-PD): study protocol for establishing a core outcome set in PD. *Perit Dial Int* 2017;37:639–47.
- 41 Matvienko-Sikar K, Byrne M, Kelly C, *et al.* Development of an infant feeding core outcome set for childhood obesity interventions: study protocol. *Trials* 2017;18:463.
- 42 Maujean A, Carroll L, Curatolo M, *et al.* A core outcome set for clinical trials in whiplash-associated disorders (WAD): a study protocol. *Trials* 2018;19:635.
- 43 McLlroy DR, Bellomo R, Billings FT, *et al.* Systematic review and consensus definitions for the standardised endpoints in perioperative medicine (step) initiative: renal endpoints. *Br J Anaesth* 2018;121:1013–24.
- 44 McPhee PG, Benner JL, Balemans ACJ, *et al.* Multimorbidity risk assessment in adolescents and adults with cerebral palsy: a protocol for establishing a core outcome set for clinical research and practice. *Trials* 2019;20:176.
- 45 Meher S, Cuthbert A, Kirkham JJ, *et al.* Core outcome sets for prevention and treatment of postpartum haemorrhage: an international Delphi consensus study. *BJOG: Int J Obstet Gy* 2019;126:83–93.
- 46 Milman N, Boonen A, Tugwell P, *et al.* Clinicians' perspective on key domains in ANCA-associated vasculitis: a Delphi exercise. *Scand J Rheumatol* 2017;46:112–7.
- 47 Murugupillai R, Ranganathan SS, Wanigasinghe J, *et al.* Development of outcome criteria to measure effectiveness of antiepileptic therapy in children. *Epilepsy & Behavior* 2018;80:56–60.
- 48 Nabbout R, Auvin S, Chiron C, *et al.* Development and content validation of a preliminary core set of patient- and caregiver-relevant outcomes for inclusion in a potential composite endpoint for Dravet syndrome. *Epilepsy & Behavior* 2018;78:232–42.
- 49 Nijagal MA, Wissig S, Stowell C, *et al.* Standardized outcome measures for pregnancy and childbirth, an ICHOM proposal. *BMC Health Serv Res* 2018;18:953.
- 50 Balakrishnan K, Sidell DR, Bauman NM, *et al.* Outcome measures for pediatric laryngotracheal reconstruction: international consensus statement. *Laryngoscope* 2019;129:244–55.
- 51 Obbarius A, van Maasackers L, Baer L, *et al.* Standardization of health outcomes assessment for depression and anxiety: recommendations from the ICHOM depression and anxiety Working group. *Qual Life Res* 2017;26:3211–25.
- 52 Ong WL, Schouwenburg MG, van Bommel ACM, *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;3:677–85.
- 53 Oude Voshaar MAH, Das Gupta Z, Bijlsma JWJ, *et al.* International Consortium for health outcome measurement set of outcomes that matter to people living with inflammatory arthritis: consensus from an international Working group. *Arthritis Care Res* 2019;71:1556–65.
- 54 Park JK, Mecoli CA, Alexanderson H, *et al.* Advancing the development of patient-reported outcomes for adult myositis at OMERACT 2016: an international Delphi study. *J Rheumatol* 2017;44:1683–7.
- 55 Perry H, Duffy JMN, Reed K, *et al.* Core outcome set for research studies evaluating treatments for twin-twin transfusion syndrome. *Ultrasound Obstet Gynecol* 2019;54:255–61.
- 56 Prins JR, Holvast F, van 't Hooft J, *et al.* Development of a core outcome set for immunomodulation in pregnancy (COSIMPREG): a protocol for a systematic review and Delphi study. *BMJ Open* 2018;8:e021619.
- 57 Rankin A, Cadogan CA, in Ryan C, *et al.* Core outcome set for trials aimed at improving the appropriateness of polypharmacy in older people in primary care. *J Am Geriatr Soc* 2018;66:1206–12.
- 58 Regardt M, Mecoli CA, Park JK, *et al.* OMERACT 2018 modified patient-reported outcome domain core set in the life impact area for adult idiopathic inflammatory myopathies. *J Rheumatol* 2019;46:1351–4.
- 59 Rose L, Agar M, Burry LD, *et al.* Development of core outcome sets for effectiveness trials of interventions to prevent and/or treat delirium (Del-CORs): study protocol. *BMJ Open* 2017;7:e016371.
- 60 Sautenet B, Tong A, Chapman J, *et al.* Developing a core outcome set in trials for adult kidney transplant recipients: an international Delphi survey. *J Evid Based Med* 2017;10:16.
- 61 Bampoe S, Cook T, Fleisher L, *et al.* Clinical indicators for reporting the effectiveness of patient quality and safety-related interventions: a protocol of a systematic review and Delphi consensus process as part of the International standardised endpoints for perioperative medicine initiative (step). *BMJ Open* 2018;8:e023427.
- 62 Schaap T, Bloemenkamp K, Deneux-Tharoux C, *et al.* Defining definitions: a Delphi study to develop a core outcome set for conditions of severe maternal morbidity. *BJOG: Int J Obstet Gy* 2019;126:394–401.
- 63 Schlessinger DI, Iyengar S, Yanes AF, *et al.* Development of a core outcome set for clinical trials in basal cell carcinoma: study protocol for a systematic review of the literature and identification of a core outcome set using a Delphi survey. *Trials* 2017;18:490.
- 64 Shorter GW, Heather N, Bray JW, *et al.* The 'Outcome Reporting in Brief Intervention Trials: Alcohol' (ORBITAL) framework: protocol to determine a core outcome set for efficacy and effectiveness trials of alcohol screening and brief intervention. *Trials* 2017;18:611.
- 65 Smith SM, Wallace E, Salisbury C, *et al.* A core outcome set for multimorbidity research (COSmm). *Ann Fam Med* 2018;16:132–8.
- 66 Somner JEA, Ismail R, Froud R, *et al.* Consensus generation of a minimum set of outcome measures for auditing glaucoma surgery outcomes—a Delphi exercise. *Graefes Arch Clin Exp Ophthalmol* 2018;256:2407–11.
- 67 Steutel NF, Benninga MA, Langendam MW, *et al.* Developing a core outcome set for infant colic for primary, secondary and tertiary care settings: a prospective study. *BMJ Open* 2017;7:e015418.
- 68 Tallouzi MO, Mathers JM, Moore DJ, *et al.* COSUMO: study protocol for the development of a core outcome set for efficacy and effectiveness trials in posterior segment-involving uveitis. *Trials* 2017;18:576.
- 69 Thorlacius L, Ingram JR, Villumsen B, *et al.* A core domain set for hidradenitis suppurativa trial outcomes: an international Delphi process. *Br J Dermatol* 2018;179:642–50.
- 70 Turnbull AE, Sepulveda KA, Dinglas VD, *et al.* Core domains for clinical research in acute respiratory failure survivors: an international modified Delphi consensus study. *Crit Care Med* 2017;45:1001–10.
- 71 Van den Bussche K, Kottner J, Beele H, *et al.* Core outcome domains in incontinence-associated dermatitis research. *J Adv Nurs* 2018;74:1605–17.
- 72 Benstoem C, Moza A, Meybohm P, *et al.* A core outcome set for adult cardiac surgery trials: a consensus study. *PLoS One* 2017;12:e0186772.
- 73 van Rijssen LB, Gerritsen A, Henselmans I, *et al.* Core set of patient-reported outcomes in pancreatic cancer (COPRAC). *Ann Surg* 2019;270:158–64.
- 74 Van Tol R, Melenhorst J, Dirksen C, *et al.* Developing a core outcome set in clinical trials for haemorrhoidal disease: an international Delphi study. *Color Dis* 2017;19:138.
- 75 Verberne WR, Das-Gupta Z, Allegretti AS, *et al.* Development of an international standard set of value-based outcome measures for patients with chronic kidney disease: a report of the International Consortium for health outcomes measurement (ICHOM) CKD Working group. *Am J Kidney Dis* 2019;73:372–84.
- 76 Viau-Lapointe J, D'Souza R, Rose L, *et al.* Development of a core outcome set for research on critically ill obstetric patients: a study protocol. *Obstet Med* 2018;11:132–6.
- 77 Wallace SJ, Worrall L, Rose T, *et al.* Which treatment outcomes are most important to aphasia clinicians and managers? an international e-Delphi consensus study. *Aphasiology* 2017;31:643–73.
- 78 Webbe J, Gale C. Core outcomes in neonatology: a core outcome set based on routinely collected data. *J Evid Based Med* 2017;10:12.
- 79 Wuytack F, Gutke A, Stuge B, *et al.* Protocol for the development of a core outcome set for pelvic girdle pain, including methods for measuring the outcomes: the PGP-COS study. *BMC Med Res Methodol* 2018;18:158.
- 80 Young A, Brookes S, Rumsey N, *et al.* Agreement on what to measure in randomised controlled trials in burn care: study protocol for the development of a core outcome set. *BMJ Open* 2017;7:e017267.
- 81 Zack R, Okunade O, Olson E, *et al.* Improving hypertension outcome measurement in low- and middle-income countries. *Hypertension* 2019;73:990–7.



- 82 Zerillo JA, Schouwenburg MG, van Bommel ACM, *et al.* An international collaborative standardizing a comprehensive patient-centered outcomes measurement set for colorectal cancer. *JAMA Oncol* 2017;3:686–94.
- 83 Beuscart J-B, Knol W, Cullinan S, *et al.* International core outcome set for clinical trials of medication review in multi-morbid older patients with polypharmacy. *BMC Med* 2018;16.
- 84 de Roos P, Bloem BR, Kelley TA, *et al.* A Consensus Set of Outcomes for Parkinson's Disease from the International Consortium for Health Outcomes Measurement. *J Parkinsons Dis* 2017;7:533–43.
- 85 Allori AC, Kelley T, Meara JG, *et al.* A standard set of outcome measures for the comprehensive appraisal of cleft care. *Cleft Palate Craniofac J* 2017;54:540–54.
- 86 Egan AM, Galjaard S, Maresh MJA, *et al.* A core outcome set for studies evaluating the effectiveness of pre-pregnancy care for women with pre-gestational diabetes. *Diabetologia* 2017;60:1190–6.
- 87 Singh JA, Dohm M, Choong PF. Consensus on draft OMERACT core domains for clinical trials of total joint replacement outcome by orthopaedic surgeons: a report from the International consensus on outcome measures in TJR trials (I-COMITT) group. *BMC Musculoskelet Disord* 2017;18:45.
- 88 Spargo M, Ryan C, Downey D, *et al.* Development of a core outcome set for trials investigating the long-term management of bronchiectasis. *Chron Respir Dis* 2019;16:147997231880416.
- 89 Agha RA, Pidgeon TE, Borrelli MR, *et al.* Validated outcomes in the grafting of autologous fat to the breast. *Plast Reconstr Surg* 2018;141:633e–8.
- 90 Hopkins C, Hettige R, Soni-Jaiswal A, *et al.* Chronic rhinosinusitis outcome measures (chrome), developing a core outcome set for trials of interventions in chronic rhinosinusitis. *Rhin* 2018;56:22–32.
- 91 Iorio A, Skinner MW, Clearfield E, *et al.* Core outcome set for gene therapy in haemophilia: results of the coreHEM multistakeholder project. *Haemophilia* 2018;24:e167–72.
- 92 Biggane AM, Brading L, Ravaud P, *et al.* Survey indicated that core outcome set development is increasingly including patients, being conducted internationally and using Delphi surveys. *Trials* 2018;19:113.
- 93 Bogdanet D, Reddin C, Macken E, *et al.* Follow-up at 1 year and beyond of women with gestational diabetes treated with insulin and/or oral glucose-lowering agents: a core outcome set using a Delphi survey. *Diabetologia* 2019;62:2007–16.
- 94 Liberati A, Altman DG, Tetzlaff J, *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009;62:e1–34.
- 95 Cohen J. Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychol Bull* 1968;70:213–20.
- 96 Altman D. *Practical statistics for medical research*. Chapman & Hall, 1991.
- 97 Sautenet B, Tong A, Manera KE, *et al.* Developing consensus-based priority outcome domains for trials in kidney transplantation. *Transplantation* 2017;101:1875–86.
- 98 Knaapen M, Hall NJ, van der Lee JH, *et al.* Protocol for the development of a global core outcome set for reporting treatment of uncomplicated appendicitis in children. *J Evid Based Med* 2019;12:15.
- 99 Smith TO, Hawker GA, Hunter DJ, *et al.* The OMERACT-OARSI core domain set for measurement in clinical trials of hip and/or knee osteoarthritis. *J Rheumatol* 2019;46:981–9.
- 100 Bernier R, Halpin E, Staffa SJ, *et al.* Inclusion of non-English-speaking patients in research: a single institution experience. *Pediatric Anesthesia* 2018;28:415–20.
- 101 Puchalski Ritchie LM, Khan S, Moore JE, *et al.* Low- and middle-income countries face many common barriers to implementation of maternal health evidence products. *J Clin Epidemiol* 2016;76:229–37.
- 102 Cook N, Siddiqi N, Twiddy M, *et al.* Patient and public involvement in health research in low and middle-income countries: a systematic review. *BMJ Open* 2019;9:e026514.
- 103 Beuscart JB, Dalleur O, Boland B, *et al.* Development of a core outcome set for medication review in older patients with multimorbidity and polypharmacy: a study protocol. *Clin Interv Aging* 2017;12:1379–89.
- 104 Process of translation and adaptation of instruments [Internet]. World health organisation 2020.
- 105 MAPI Institute. *Linguistic validation manual for health outcome assessments*. 2nd ed. MAPI Research Trust, 2012.
- 106 Kirkham JJ, Gorst S, Altman DG, *et al.* Core outcome set-standards for reporting: the COS-STAR statement. *J Evid Based Med* 2017;10:7.
- 107 Jones JE, Jones LL, Keeley TJH, *et al.* A review of patient and carer participation and the use of qualitative research in the development of core outcome sets. *PLoS One* 2017;12:e0172937:1–18.