Journal of Clinical Epidemiology 125 (2020) 64−75

Journal of Clinical Epidemiology

REVIEW ARTICLE

Methods used in the selection of instruments for outcomes included in core outcome sets have improved since the publication of the COSMIN/COMET guideline

Sarah L. Gorst[a,*], Cecilia A.C. Prinsen[b], Maximilian Salcher-Konrad[c], Karen Matvienko-Sikar[d], Paula R. Williamson[a], Caroline B. Terwee[b]

[a]*Department of Biostatistics, University of Liverpool, Liverpool, UK*

[b]*Department of Epidemiology and Biostatistics, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam Public Health Research Institute, Amsterdam, the Netherlands*

[c]*Department of Health Policy, London School of Economics and Political Science, Cowdray House, London, UK*

[d]*School of Public Health, University College Cork, Cork, Ireland*

Accepted 20 May 2020; Published online 26 May 2020

Abstract

Objectives: Once a core outcome set (COS) has been defined, it is important to achieve consensus on how these outcomes should be measured. The aims of this systematic review were to gain insight into the methods used to select outcome measurement instruments and to determine whether methods have improved following the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN)/Core Outcome Measures in Effectiveness Trials (COMET) guideline publication.

Study Design and Setting: Eligible articles, which were identified from the annual COMET systematic review, concerned any COS development studies that provided a recommendation on how to measure the outcomes included in the COS. Data were extracted on the methods used to select outcome measurement instruments in accordance with the COSMIN/COMET guideline.

Results: Of the 118 studies included in the review, 48% used more than one source of information when finding outcome measurement instruments, and 74% performed some form of quality assessment of the measurement instruments. Twenty-three studies recommended one single instrument for each core outcome included in the COS. Clinical experts and public representatives were involved in selecting instruments in 62% and 28% of studies, respectively.

Conclusion: Methods used to select outcome measurement instruments have improved since the publication of the COSMIN/COMET guideline. Going forward, COS developers should ensure that recommended outcome measurement instruments have sufficient content validity. In addition, COS developers should recommend one instrument for each core outcome to contribute to the overarching goal of uniformity in outcome reporting.

Keywords: Core outcome set; COS; Development; Outcome measurement instrument; Selection; Methodology

1. Introduction

There is lack of consensus with regard to the selection of outcomes and outcome measurement instruments for clinical trials, which causes inconsistencies in the outcomes reported and difficulties in comparing these outcomes in systematic reviews and meta-analyses [1]. In addition, there is great variability in the quality of outcome measurement instruments used, and it is not always clear if the best instrument is being used for a given outcome. To overcome these issues, standardization of the selection of outcomes and outcome measurement instruments is needed.

Funding Statement: This work was supported by the European Union's Horizon 2020 research and innovation program (CORBEL under grant agreement number 654248).

Conflict of interest: P.R.W. is a member of the Core Outcome Measures in Effectiveness Trials Management Group. C.B.T. is the founder of the COnsensus-based Standards for the selection of health Measurement INstruments initiative. The remaining authors declare no conflict of interest.

* Corresponding author. Tel.: +44 0 151 794 9725.

E-mail address: sgorst@liverpool.ac.uk (S.L. Gorst).

https://doi.org/10.1016/j.jclinepi.2020.05.021

What is new?

Key findings

- Methods used to select core outcome measurement instruments vary across studies, with many studies not meeting the recommended standards.
- Methods used to select outcome measurement instruments have improved since the publication of the COnsensus-based Standards for the selection of health Measurement INstruments/Core Outcome Measures in Effectiveness Trials guideline.

What this adds to what was known?

- This is the first study to assess how the outcome measurement instruments recommended in existing core outcome sets have been selected and whether good practices are being followed.

What is the implication and what should change now?

- Core outcome set developers need to make better use of the guidance available when agreeing on how to measure the outcomes included in core outcome sets.
- Developers need to ensure that outcome measurement instruments are of sufficient quality and especially have sufficient content validity.

The Core Outcome Measures in Effectiveness Trials (COMET) Initiative (www.comet-initiative.org), launched in January 2010, aims to facilitate the development and application of agreed standardized sets of outcomes, also known as ''core outcome sets'' (COSs). A COS is an agreed minimum set of outcomes that should be measured and reported in all clinical trials of a specific disease or trial population (i.e., what to measure) [1]. Once the COS has been defined, it is then important to achieve consensus on how these outcomes should be measured (i.e., how to measure).

The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative (http://www.cosmin.nl/) aims to improve the selection of outcome measurement instruments. (When using the term ''outcome measurement instruments,'' we are referring to any instruments, definitions, tools, procedures, etc., that are used to measure an outcome.) In 2016, COSMIN and COMET published a consensus-based guideline on how to select outcome measurement instruments for outcomes included in a COS [2]. Among a large group of international stakeholders from 14 different countries, including clinicians, clinimetricians/psychometricians, epidemiologists, journal editors, physicians, researchers, and statisticians, consensus was obtained on methods for selecting outcome measurement instruments for outcomes included in a COS. COS developers are guided through the process of instrument selection in four consecutive steps: Step 1, conceptual considerations; Step 2, finding existing outcome measurement instruments; Step 3, quality assessment of outcome measurement instruments; and Step 4, recommendations on the selection of outcome measurement instruments. It is unknown, however, how the outcome measurement instruments recommended in existing COS have actually been selected by COS developers and whether good practices for COS development, as described in the COSMIN/COMET guideline are being followed.

The aims of this systematic review were to (1) gain insight into the methods used by COS developers for selecting core outcome measurement instruments and (2) determine whether the methods have improved following the 2016 publication of the COSMIN/COMET guideline. It is hypothesized that the quality of the methods used to select the core outcome measurement instruments varies considerably, and it is therefore anticipated that there will be considerable room for improvement in COS development with regard to instrument selection. However, there are expected to be improvements in the methods used in studies that have been published from 2017 onward.

2. Methods

2.1. Eligibility criteria

Eligible articles concern COS development studies that provide recommendations on what and how to measure, either done together in one study or done in two separate stages (i.e., two or more studies). We included all COS studies, identified from the original COMET systematic review and annual updates [3−8], that provided a recommendation on the instruments to measure the outcomes included in the COS. COS development studies that only provide a recommendation on what to measure, but do not consider how to measure, and studies that discuss how to measure the outcomes but do not give a recommendation were excluded.

2.2. Literature search

The search strategy for identifying eligible COS development studies has been described elsewhere in detail [3−8]. In brief, a comprehensive search strategy to identify studies that aimed to define COS in any disease area was first developed in 2013 [3] (see Appendix A for full search strategy). Database searches were repeated in 2015 [4], 2016 [5], 2017 [6], 2018 [7], and 2019 [8].

2.3. Data extraction

Four sets of reviewers (C.A.C.P.−S.L.G., S.L.G.−M.S.-K, M.S.-K−C.A.C.P., and S.L.G.−K.M.-S.) independently

extracted data on descriptive information for each identified COS, including the target population, disease area, and the (number of) outcomes. Data have been extracted in accordance with the COSMIN/COMET guideline (see flowchart in Appendix B). Methods used to select instruments for the COS have been extracted, including the approach taken to identify existing instruments, the evaluation of the quality and feasibility of instruments, the number of instruments recommended for use, arguments used for selecting instruments other than quality criteria (measurement properties), recommendations for additional research on instruments, and whether any guidance for instrument selection, including the COSMIN/COMET guideline, was followed (see Appendix C). To ensure consistency in data extraction, the data extraction form was pilot tested for a set of five studies, and the extracted data were compared before extracting data for the remaining studies. Discrepancies in data extraction between pairs of reviewers were sought to be resolved by discussion with the third reviewer, and consensus was reached.

To improve the quality of our data, first authors of the included studies were contacted in person by email to verify the data extracted from their studies, and they were asked to provide additional information that might be missing. In case the email could not be delivered, the last author of that particular study was contacted by email. References of the included papers were also checked to identify any other relevant articles on instrument selection for COS.

3. Results

A total of 163 articles describing 118 COS development studies were included in the review. A flow diagram of the article and abstract selection process is provided in Fig. 1, guided by Preferred Reporting Items for Systematic Reviews and Meta-Analyses [9].

Details on COS development studies (e.g., target population, disease area) can be found in Appendix D. In summary, COSs were developed in a variety of geographical locations, including Asia, Canada, Europe, New Zealand, South Africa, South America, and the United States. All COSs were developed in the English language. The number of core outcomes included in the COS varied between 2 [10,11] and 26 [12]. Following the COMET classifications [13], COS were developed in 24 different disease areas, mostly neurology (n = 19 studies), rheumatology (n = 14), heart and circulation (n = 13), and orthopedics and trauma (n = 11). In 36 studies, the COS were developed for adults; in 17 studies for children and in 20 studies for both adults and children. In 45 studies (38%), the age group for which the COS was developed was not specified. Of the 118 studies, 23 studies reported on core outcomes for different subgroups of patients, such as age groups [14–16], acute or chronic conditions [17], disease severity [18,19], type of study (e.g., prevention trials vs. intervention trials [20,21] or phase I–II vs. phase III clinical trials [15]); for acute vs. long-term treatment [22] or acute treatment vs. prophylaxis [23–25]; for different diseases/conditions [16,26–32]; or for different settings [33,34].

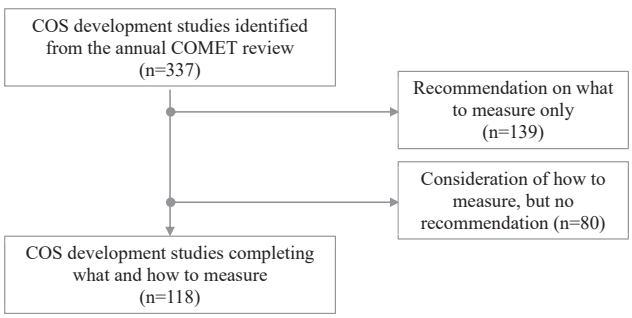

Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of identification of eligible studies from the Core Outcome Measures in Effectiveness Trials database.

3.1. Methods used in the selection of instruments for COS

With regard to COS development, 87 of 118 COS studies used a single process to identify the core outcomes as well the instruments recommended to measure these outcomes; whereas 31 studies used a two-stage process that involved first agreeing on ‘‘what to measure’’ (select core outcomes) before moving onto the ‘‘how to measure’’ (recommend instruments).

3.1.1. Finding existing outcome measurement instruments

It is recommended that COS developers aim for finding all existing outcome measurement instruments. Multiple sources of information can be used to find instruments: (1) performing a systematic review, including a search in MEDLINE (and EMBASE); (2) use existing review(s); (3) reference lists; (4) expert opinion; or (5) other sources of information, such as online databases, book (chapters), or conference proceedings [2]. Of the 118 included studies, 21 studies (18%) used three or more sources of information when finding existing outcome measurement instruments, and 36 studies (30%) used two sources of information. Sixty-one studies (52%) used only one source of information when finding existing outcome measurement instruments, with 39 of 61 studies accessing expert opinion only. Of the 118 studies, 52 studies (44%) performed a systematic review; 19 studies (16%) used an existing review; seven studies (6%) searched reference lists; 49 studies (42%) accessed expert opinion; and 15 studies (13%) used other sources of information, mostly instruments used in clinical trials.

3.1.2. Quality assessment of outcome measurement instruments

COS developers should base their recommendations for outcome measurement instruments on (1) the quality of the existing outcome measurement instruments, that is, their measurement properties (including an evaluation of the quality of the validation studies), and (2) the feasibility aspects of the outcome measurement instruments [2].

3.1.2.1. Measurement properties. It is recommended that evidence on the measurement properties of outcome measurement instruments should be available in the target population. The quality of instruments is determined in studies on measurement properties, which should be of high methodological quality. Of the 118 studies, six studies (5%) [35−40] considered both the results of the measurement properties of the outcome measurement instruments and the quality of studies on these measurement properties. However, in three of these studies [35,36,38], it remains unclear whether a best evidence synthesis was performed. COS developers of 23 of 118 studies (20%) only considered the results of the measurement properties of the included outcome measurement instruments but did not consider the quality of the studies on these measurement properties. In 58 of 118 studies (49%), COS developers referred to quality criteria of outcome measurement instruments; however, there was no mention of any formal assessment of whether the instruments met these criteria. COS developers of 31 of 118 studies (26%) did not take the quality of the outcome measurement instruments into account when making their recommendations.

3.1.2.2. Feasibility. Of the 118 studies, 74 studies (63%) have taken feasibility aspects into consideration in the selection of instruments for the COS, such as availability of the instrument, cost of an instrument, ease of administration, and length of an instrument. In 44 studies (37%), there was no indication that feasibility was taken into consideration in the selection of instruments for the COS.

3.1.3. Recommendations on the selection of outcome measurement instruments

It is advised to recommend only one outcome measurement instrument for each core outcome per subdomain/subpopulation in the COS, as this will serve the ultimate goal of standardization of outcome reporting [2]. Of the 118 studies, only 11 studies (9%) recommended one single instrument for each core outcome included in the COS [12,38,39,41−48]. In seven of these 11 studies, one instrument was selected for each core outcome (range: 4−26) in the COS [12,38,41,43,44,47,48]. In three studies, one instrument was recommended for each core outcome or for each subpopulation (i.e., children and adolescents; range of core outcomes: 4−10) [42,45,46]. In one study, one instrument was recommended for each of the four core outcomes, with two alternative instruments recommended for two of the outcomes because they were free of charge [39]. In 12 of the 118 studies (10%), one single instrument was recommended for each core outcome other than those for which no outcome measurement instrument could be recommended (range of core outcomes with instrument recommended: 1−17) [21,37,49−58]. Twelve of the 118 studies (10%) recommended one instrument for all included core outcomes except one (range included core outcomes: 3−15) [25,35,40,59−67]. Another seven studies (6%) recommended multiple instruments for all core outcomes included in the COS [19,36,68−72]. In 76 of 118 studies (64%), a combination of recommendations was used in the selection of instruments for each core outcome included in the COS (i.e., for some outcomes, one instrument was recommended; for some outcomes, multiple instruments were recommended, either for the entire group or for different subgroups; and for some outcomes, no instrument was recommended).

3.1.4. Consensus procedure used to reach agreement

It is recommended that COS developers use a consensus procedure to get final agreement on the selected instruments included in the COS [2]. In 80 of 118 studies (68%), a consensus procedure was used. In 30 of 80 studies, it was unclear and not specified how consensus was obtained. In 8 of 80 studies, COS developers used a Delphi technique to reach consensus on the selection of core instruments [12,43,48,64,73−76]. Six of the 80 studies were guided by the Outcome Measures in Rheumatology (OMERACT) consensus and validation process, which involved participants voting and then breaking out into groups to review and discuss domains and instruments [39,40,63,72,77,78]. In 36 of 80 studies, COS developers conducted a consensus meeting, including various methods, to reach consensus on the core instruments. Consensus methods used at the meetings included presentations, nominal group techniques, group discussions, consensus workshops, breakout sessions, and voting. In 38 of 118 studies (32%), no consensus procedure was used to agree on the instruments included in the COS, and recommendations were formulated by the COS developers.

3.1.5. Stakeholders involved in the selection of outcome measurement instruments

Of the 118 studies, the following stakeholders were involved in the selection of outcome measurement instruments: clinical experts (n = 73), nonclinical researchers (n = 39), patients and/or public representatives (n = 33), regulatory authorities (n = 20), and industry representatives (n = 17). An additional 43 studies did not provide any details about the stakeholders involved in selecting the outcome measurement instruments. Table 1 displays the different stakeholder combinations across the 118 studies.

Table 1. Number of studies involving each stakeholder group combination

Stakeholder groups	*n* (%)
Clinical experts	19 (16)
Clinical experts, public representatives, and nonclinical research experts	11 (9)
Clinical experts and nonclinical research experts	8 (7)
Clinical experts and public representatives	7 (6)
Clinical experts, public representatives, nonclinical research experts, and industry experts	6 (5)
Clinical experts, public representatives, nonclinical research experts, authorities, and industry experts	4 (3)
Clinical experts, public representatives, nonclinical research experts, and authorities	3 (3)
Clinical experts, nonclinical research experts, and authorities	3 (3)
Clinical experts, authorities, and industry experts	3 (3)
Clinical experts, public representatives, and authorities	2 (2)
Clinical experts and authorities	2 (2)
Clinical experts and other	2 (2)
Clinical experts, nonclinical research experts, authorities, and industry experts	2 (2)
Clinical experts and industry experts	1 (1)
Nonclinical research experts, authorities, and industry experts	1 (1)
Nonclinical research experts	1 (1)
No details provided	43 (36)

3.1.6. Guidance on instrument selection

In 35 of 118 studies (30%), published guidance for instrument selection was used. Most studies (n = 13) used the OMERACT guidance [79], whereas other studies used the Grading of Recommendations, Assessment, and Evaluation approach [16,80]; Initiative on Methods, Measurement, and Pain Assessment in Clinical Trials recommendations [81,82]; World Health Organization–International Classification of Functioning, Disability and Health framework [83,84]; European League Against Rheumatism operating procedures [85]; and International Consortium for Health Outcomes Measurement framework [12,48,64]. Four studies used the COSMIN/COMET guideline [39,40,76,84], and a fifth study [37] used the COSMIN standards for the selection of health status measurement instruments [86]. Eleven studies [11,17,32,36,87–93] referred to other guidance, for example, guidelines from European Crohn's and Colitis Organisation guidelines committee; guidance by Physical Rehabilitation Outcomes Measures, published by the Canadian Physiotherapy Association; previous work by Hudak et al. [94] or Dworkin et al. [95]; other consensus guidelines [96–101]; the Harmonizing Outcome Measures for Eczema (HOME) roadmap [102]; the framework for the selection of clinical trial indices proposed by Tugwell and Bombardier [103]; and guidance from a qualitative evaluation of measures for psychosocial intervention in dementia care. No guidance was used in 43 of 118 studies (37%), whereas in 40 of 118 studies (34%), it remains unclear whether any form of guidance was used.

3.1.7. Recommendations for additional research on instruments

In 55 of 118 studies (47%), recommendations were made for additional validation studies (n = 31) or development of new instruments (n = 24).

3.2. Differences between studies published before and after the publication of the COSMIN/COMET guideline

Of the 118 studies included in this review, 92 studies (78%) were published before the publication of the

COSMIN/COMET guideline, and 26 studies (22%) were published after its publication. Table 2 provides a comparison of the methods used in studies published before and after the publication of the COSMIN/COMET guideline. Studies published after the publication of the COSMIN/COMET guideline were more likely to base their recommendations on the quality of the outcome measurement instruments, with 8 of 26 studies (31%) considering the evaluation of the measurement properties of the instruments and a further 4 of 26 studies (15%) considering both the quality of the measurement properties of the outcome measurement instruments and the quality of studies on measurement properties. In addition, these studies were also more likely to comply with the COSMIN/COMET recommendations on the selection of outcome measurement instruments, with 11 of 26 studies (42%) recommending one single instrument for each core outcome included in the COS, for which a recommendation could be made, and a further 6 of 26 studies (23%) recommending one instrument for all included core outcomes except one. The inclusion of stakeholders across all groups increased in the postguideline studies, with the biggest increase being the inclusion of patients and/or public representatives, which increased from 20% to 58%. There was also a 31% decrease in the number of studies that did not provide any details about stakeholder involvement. Regarding the use of guidance, 12 of 26 studies (46%) published following the COSMIN/COMET guideline used available guidance for instrument selection, with 5 of 26 studies (19%) specifically using the COSMIN/COMET guideline or other COSMIN guidance.

4. Discussion

We identified 118 COS development studies that provided recommendations for how to measure the outcomes included in a COS. Reviewing these studies has enabled us to gain insight into the methods used by COS developers to select outcome measurement instruments. After the publication of the COSMIN/COMET guideline, there has been an improvement in the methods used, specifically in relation to quality assessment, recommendations on the selection of instruments, stakeholder involvement, and the use of published guidance.

4.1. Finding all existing instruments

COS developers should make better use of the literature to inform their instrument selection process. Relying solely on expert opinion to find existing outcome measurement instruments, as 33% of studies did in the current review, may result in only the most commonly used instruments or those that are favored by clinicians, being considered. To assist developers in identifying instruments, COSMIN maintains a database of systematic reviews of outcome measurement instruments [104]. The ‘‘COSMIN guideline for systematic reviews of patient-reported outcome measures’’ [105] can be used for performing a comprehensive literature search or full systematic review to find all available instruments if a good-quality systematic review is not available.

4.2. Quality assessment of instruments

COS developers should take both the quality of the studies on measurement properties and the results of the measurement properties of the outcome measurement instruments into account in their recommendations. This will ensure that the most reliable and valid outcome measurement instruments are selected. However, only 10 of the COS development studies included in this review considered both aspects. To assist in the assessment of the quality of the studies, COSMIN has developed a risk of bias checklist for use in systematic reviews of patient-reported outcome measures (PROMs) to assess the risk of bias of studies on measurement properties [106]. COSMIN has also proposed quality criteria for measurement properties of health status questionnaires to assist COS developers in assessing the quality of identified instruments [106].

In contrast to the assessment of measurement properties, the feasibility of the identified instruments was generally taken into consideration in the selection of outcome measurement instruments. We expect that the number of studies considering these feasibility aspects will continue to remain high, after the publication of the COSMIN/COMET guideline in 2016 [2].

4.3. Generic recommendations on the selection of outcome measurement instruments

COS developers should try to recommend only a single instrument for each individual outcome included in a COS, to contribute to the overarching goal of uniformity in outcome reporting, and to enhance the comparability of clinical trials. Exceptions can, however, be made for subpopulations, such as children and adults, and different instruments may be necessary for the different age groups.

COS stakeholder groups should include more representative stakeholders, including patients, when agreeing on the most appropriate outcome measurement instruments. Most COS stakeholder groups comprised clinical experts, whereas public representatives were involved in less than one-third of studies. This is concerning, as it suggests that the outcome measures recommended may not be those that public representatives deem to be most appropriate. Patients are increasingly being included in selecting the outcomes for inclusion in COS, with 92% of ongoing COS development studies in the COMET database planning to include some degree of patient input [107]. However, patients and/or public representatives were only included in the selection of outcome measurement instruments in 28% of COS studies.

Table 2. Methods used in studies published before and after the publication of the COSMIN/COMET guideline

COSMIN/COMET Guideline Task	Preguideline, *n/N* (%)	Postguideline, *n/N* (%)
Finding existing outcome measurement instruments		
Used three or more sources to find existing outcome measurement instruments	16/92 (17)	5/26 (19)
Used two sources to find existing outcome measurement instruments	30/92 (33)	6/26 (23)
Used one source to find existing outcome measurement instruments	46/92 (50)	15/26 (58)
Quality assessment of outcome measurement instruments		
(a) Measurement properties		
Considered both the quality of the measurement properties of the outcome measurement instruments and the quality of studies on measurement properties	2/92 (2)	4/26 (15)
Considered the evaluation of the measurement properties of the included outcome measurement instruments	15/92 (16)	8/26 (31)
Referred to quality criteria but no formal assessment	52/92 (57)	6/26 (23)
Quality of outcome measurement instruments not taken into account	23/92 (25)	8/26 (31)
(b) Feasibility		
Feasibility aspects taken into consideration	59/92 (64)	15/26 (58)
Recommendations on the selection of outcome measurement instruments		
One single instrument for each core outcome	3/92 (2)	8/26 (31)
One single instrument for each core outcome where a recommendation could be made	9/92 (10)	3/26 (12)
One instrument for all included core outcomes except one	6/92 (7)	6/26 (23)
Multiple instruments for all core outcomes	7/92 (8)	0/26 (0)
Combination of recommendations for each core outcome	67/92 (73)	9/26 (35)
Consensus procedure used to reach agreement		
Used consensus procedure to get final agreement on the selected instruments	62/92 (67)	18/26 (69)
Specified details of the consensus procedure	35/62 (56)	15/18 (83)
Stakeholders involved in the selection of outcome measurement instruments		
Clinical experts	50/92 (54)	23/26(88)
Patients and/or public representatives	18/92 (20)	15/26 (58)
Nonclinical researchers	26/92 (28)	13/26 (50)
Regulatory authorities	15/92 (16)	5/26 (19)
Industry representatives	13/92 (14)	4/26 (15)

(*Continued*)

Table 2. Continued

COSMIN/COMET Guideline Task	Preguideline, *n/N* (%)	Postguideline, *n/N* (%)
No details provided about the stakeholders involved in selecting the outcome measurement instruments	40/92 (43)	3/26 (12)
Guidance on instrument selection		
Used published guidance	23/92 (25)	12/26 (46)
Recommendations for additional research on instruments		
Made recommendations for the development of new instruments or additional validation studies	42/92 (46)	13/26 (50)

It could be argued that it is more difficult to include patients in the selection of outcome measurement instruments than in the selection of the outcomes because the selection of instruments is mostly based on studies on measurement properties, which may be difficult for patients to understand. COS developers may need to address additional issues when the population concerned includes people with cognitive impairment, communication difficulties, or other vulnerabilities, which make participation in such processes challenging. Despite the complexities involved in the selection of outcome measurement instruments, there has been an increase in the inclusion of patients and/or public representatives since the publication of the COSMIN/COMET guideline. Thus, it is certainly possible to involve patients and/or public representatives, including those from vulnerable groups, in this process. However, it should be acknowledged that these groups are likely to need additional support to participate, and so in some instances, a multistage approach, which allows for adequate support, might be necessary to ensure that engagement is meaningful. One potential means of involvement is for patients to judge the face and content validity (relevance, comprehensiveness, and comprehensibility) of the available outcome measurement instruments, which are considered the most important measurement properties [106,108]. However, it should be noted that this may not be necessary if previously published studies have assessed the validity of outcome measurement instruments with the population in question.

4.4. Implications

This review has highlighted that the methods used to select the core outcome measurement instruments vary across studies, with many studies not meeting the recommended standards. However, the majority of included studies were published before the development of the COSMIN/COMET guideline, in 2016, and so developers may have been unaware of methodology for selecting outcome measurement instruments. There have, however, been clear improvements in the methods used to select outcome measurement instruments in studies published since the publication of the COSMIN/COMET guideline. It is unclear whether such improvements are a direct result of the publication of the COSMIN/COMET guideline or whether other variables are responsible for the pre- and post-guideline reporting differences. Other potential variables may include increased COS awareness prompting COS developers to be more thorough in the outcome measurement instrument selection process. In addition, the differences in the rate of inclusion of patients may be attributable to increased public input in health research in general.

Going forward, we hope that COS developers will use the COSMIN/COMET guideline, along with the other resources listed previously, to ensure that recommendations for outcome measurement instruments are developed using rigorous methodology. A recent paper by Ju et al. [109] highlights how to apply the COSMIN/COMET guidance when identifying outcome measurement instruments.

Apart from the COSMIN/COMET guideline, other guidelines can also be used to guide the selection of outcome measurement instruments for COS, for example, the OMERACT Handbook [110] or HOME roadmap [102]. Different guidelines put different emphasis on different steps of the process. For example, in contrast to the COSMIN/COMET guideline and HOME roadmap, the OMERACT process does not require a search to find all available instruments but starts with a selection of instruments that seem to have a good match with the target domain and are considered feasible. Both the COSMIN/COMET guideline and the OMERACT Handbook address the need for good content validity of outcome measurement instruments. Therefore, we suggest that, when resources are limited, COS developers should evaluate the content validity of available instruments, if this has not been done previously, with a small number of patients in their stakeholder group (e.g., ask patients to evaluate the relevance, comprehensiveness, and comprehensibility of all items/tests). The COSMIN methodology for assessing content validity of PROMs can be used for further guidance [108]. Nevertheless, it should be acknowledged that although guidelines are important, some flexibility should be used to best facilitate the participation of patients and/or public representatives.

4.5. Limitations

All the studies included in the current review were identified from the annual COMET systematic review of COS. We did not perform a systematic search for all studies relating to how outcomes should be measured. For example, COS groups may perform systematic reviews of outcome measurement instruments and select their instruments based on these reviews but may not publish a separate paper on the selection process for the outcome measurement instruments; therefore, we cannot be certain that we have identified all relevant studies. However, we did check the references of all included papers, and authors of COS studies were contacted to provide additional information.

5. Conclusions

In conclusion, COS developers need to make better use of the guidance available when agreeing on how to measure the outcomes included in COS. Specifically, developers need to ensure that outcome measurement instruments are of sufficient quality and especially have sufficient content validity. Furthermore, developers should aim to adhere to uniformity by selecting a single outcome measurement instrument for each outcome within a COS.

CRediT authorship contribution statement

Sarah L. Gorst: Investigation, Methodology, Writing - original draft, Writing - review & editing, Project administration. **Cecilia A.C. Prinsen:** Investigation, Methodology, Writing - original draft, Writing - review & editing, Project administration. **Maximilian Salcher-Konrad:** Investigation, Writing - review & editing. **Karen Matvienko-Sikar:** Investigation, Writing - review & editing. **Paula R. Williamson:** Conceptualization, Methodology, Writing - review & editing. **Caroline B. Terwee:** Conceptualization, Methodology, Writing - review & editing.

Acknowledgments

The authors would like to thank the reviewers for their helpful comments on the manuscript.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jclinepi.2020.05.021.

References

[1] Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials 2012;13:132.

[2] Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" - a practical guideline. Trials 2016;17(1):449.

[3] Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M, et al. Choosing important health outcomes for comparative effectiveness research: a systematic review. PLoS One 2014;9:e99111.

[4] Gorst SL, Gargon E, Clarke M, Blazeby JM, Altman DG, Williamson PR. Choosing important health outcomes for comparative effectiveness research: an updated review and user survey. PLoS One 2016;11:e0146444.

[5] Gorst SL, Gargon E, Clarke M, Smith V, Williamson PR. Choosing important health outcomes for comparative effectiveness research: an updated review and identification of gaps. PLoS One 2016;11:e0168403.

[6] Davis K, Gorst SL, Harman N, Smith V, Gargon E, Altman DG, 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.

[7] Gargon E, Gorst SL, Harman NL, Smith V, Matvienko-Sikar K, Williamson PR. 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.

[8] Gargon E, Gorst SL, Williamson PR. Choosing important health outcomes for comparative effectiveness research: 5th annual update to a systematic review of core outcome sets for research. PLoS One 2019;14:e0225980.

[9] PRISMA statement. Available at http://www.prisma-statement.org/. Accessed October 27, 2016.Available at http://www.prisma-statement.org/.

[10] Donovan DM, Bigelow GE, Brigham GS, Carroll KM, Cohen AJ, Gardin JG, et al. Primary outcome indices in illicit drug dependence treatment research: systematic approach to selection and measurement of drug use end-points in clinical trials. Addiction 2012;107(4):694−708.

[11] Keim SM, Spaite DW, Maio RF, Garrison HG, Desmond JS, Gregor MA, et al. Risk adjustment and outcome measures for out-of-hospital respiratory distress. Acad Emerg Med 2004;11(10):1074−81.

[12] Ong WL, Schouwenburg MG, van Bommel ACM, Stowell C, Allison KH, Benn KE, et al. A standard set of value-based patient-centered outcomes for breast cancer: the International consortium for Health Outcomes Measurement (ICHOM) initiative. JAMA Oncol 2017;3(5):677−85.

[13] Dodd S, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. J Clin Epidemiol 2017;96:84−92.

[14] Busse WW, Morgan WJ, Taggart V, Togias A. Asthma outcomes workshop: overview. J Allergy Clin Immunol 2012;129(3 Suppl):S1−8.

[15] Ramsey BW, Boat TF. Outcome measures for clinical trials in cystic fibrosis. Summary of a Cystic Fibrosis Foundation consensus conference. J Pediatr 1994;124(2):177−92.

[16] Sanyal AJ, Brunt EM, Kleiner DE, Kowdley KV, Chalasani N, Lavine JE, et al. Endpoints and clinical trial design for nonalcoholic steatohepatitis. Hepatology 2011;54(1):344−53.

[17] McGrath PJ, Walco GA, Turk DC, Dworkin RH, Brown MT, Davidson K, et al. Core outcome domains and measures for pediatric acute and chronic/recurrent pain clinical trials: PedIMMPACT recommendations. J Pain 2008;9:771−83.

[18] Vellas B, Andrieu S, Sampaio C, Coley N, Wilcock G. Endpoints for trials in Alzheimer's disease: a European task force consensus. Lancet Neurol 2008;7(5):436−50.

[19] Chen RC, Chang P, Vetter RJ, Lukka H, Stokes WA, Sanda MG, et al. Recommended patient-reported core set of symptoms to

measure in prostate cancer treatment trials. J Natl Cancer Inst 2014; 106:dju132.
[20] Abellan van Kan G, Cameron Chumlea W, Gillette-Guyonet S, Houles M, Dupuy C, Rolland Y, et al. Clinical trials on sarcopenia: methodological issues regarding phase 3 trials. Clin Geriatr Med 2011;27(3):471−82.
[21] Balakrishnan K, Sidell DR, Bauman NM, Bellia-Munzon GF, Boesch RP, Bromwich M, et al. Outcome measures for pediatric laryngotracheal reconstruction: international consensus statement. Laryngoscope 2019;129(1):244−55.
[22] Recommendations for guidelines on clinical trials of mucoactive drugs in chronic bronchitis and chronic obstructive pulmonary disease. Task Group on Mucoactive Drugs. Chest 1994;106:1532−7.
[23] Tfelt-Hansen P, Pascual J, Ramadan N, Dahlof C, D'Amico D, Diener HC, et al. Guidelines for controlled trials of drugs in migraine: third edition. A guide for investigators. Cephalalgia 2012;32(1):6−38.
[24] Lipton RB. Guidelines for controlled trials of drugs in cluster headache. Cephalalgia 1995;15(6):452−62.
[25] Bendtsen L, Bigal ME, Cerbo R, Diener HC, Holroyd K, Lampl C, et al. Guidelines for controlled trials of drugs in tension-type headache: second edition. Cephalalgia 2010;30(1):1−16.
[26] Ruperto N, Ravelli A, Murray KJ, Lovell DJ, Andersson-Gare B, Feldman BM, et al. Preliminary core sets of measures for disease activity and damage assessment in juvenile systemic lupus erythematosus and juvenile dermatomyositis. Rheumatology (Oxford) 2003;42(12):1452−9.
[27] Porst H, Vardi Y, Akkus E, Melman A, Park NC, Seftel AD, et al. Standards for clinical trials in male sexual dysfunctions. J Sex Med 2010;7(1 Pt 2):414−44.
[28] Hausenloy DJ, Erik Botker H, Condorelli G, Ferdinandy P, Garcia-Dorado D, Heusch G, et al. Translating cardioprotection for patient benefit: position paper from the Working Group of cellular biology of the Heart of the European society of cardiology. Cardiovasc Res 2013;98(1):7−27.
[29] Vocci F, de With H. Consensus statement on evaluation of outcome of pharmacotherapy for substance abuse/dependence: report from a NIDA/CPDD meeting. Bethesda, MD: National Institute on Drug Abuse Medications Development Division; 1999.
[30] Labs KH, Dormandy JA, Jaeger KA, Stuerzebecher C, Hiatt WR. Trans-atlantic conference on clinical trial guidelines in PAOD (Peripheral arterial occlusive disease) clinical trial methodology. Eur J Vasc Endovasc Surg 1999;18(3):253−65.
[31] Merkies IS, Lauria G. 131st ENMC international workshop: selection of outcome measures for peripheral neuropathy clinical trials 10-12 December 2004, Naarden, the Netherlands. Neuromuscul Disord 2006;16(2):149−56.
[32] Kenny KP, Day PF, Sharif MO, Parashos P, Lauridsen E, Feldens CA, et al. What are the important outcomes in traumatic dental injuries? An international approach to the development of a core outcome set. Dent Traumatol 2018;34(1):4−11.
[33] Feldman LS, Lee L, Fiore J Jr. What outcomes are important in the assessment of Enhanced Recovery After Surgery (ERAS) pathways? Can J Anaesth 2015;62(2):120−30.
[34] Kloppenburg M, Boyesen P, Visser AW, Haugen IK, Boers M, Boonen A, et al. Report from the OMERACT Hand Osteoarthritis Working Group: set of core domains and preliminary set of instruments for use in clinical trials and observational studies. J Rheumatol 2015;42:2190−7.
[35] Haywood KL, Griffin XL, Achten J, Costa ML. Developing a core outcome set for hip fracture trials. Bone Joint J 2014;96-B(8): 1016−23.
[36] Moniz-Cook E, Vernooij-Dassen M, Woods R, Verhey F, Chattat R, De Vugt M, et al. A European consensus on outcome measures for psychosocial intervention research in dementia care. Aging Ment Health 2008;12(1):14−29.
[37] Kwakkel G, Lannin NA, Borschmann K, English C, Ali M, Churilov L, et al. Standardized measurement of sensorimotor recovery in stroke trials: consensus-based core recommendations from the Stroke Recovery and Rehabilitation Roundtable. Int J Stroke 2017; 12(5):451−61.
[38] Pushpanathan E, Setty T, Carvalho B, Sultan P. A systematic review of postoperative pain outcome measurements utilised in regional anesthesia randomized controlled trials. Anesthesiol Res Pract 2018;2018:9050239.
[39] Chiarotto A, Boers M, Deyo RA, Buchbinder R, Corbin TP, Costa LOP, et al. Core outcome measurement instruments for clinical trials in nonspecific low back pain. Pain 2018;159(3):481−95.
[40] Spuls PI, Gerbens LAA, Simpson E, Apfelbacher CJ, Chalmers JR, Thomas KS, et al. Patient-Oriented Eczema Measure (POEM), a core instrument to measure symptoms in clinical trials: a Harmonising Outcome Measures for Eczema (HOME) statement. Br J Dermatol 2017;176:979−84.
[41] Deyo RA, Dworkin SF, Amtmann D, Andersson G, Borenstein D, Carragee E, et al. Report of the NIH Task Force on research standards for chronic low back pain. J Pain 2014;15:569−85.
[42] Mindell JA, Emslie G, Blumer J, Genel M, Glaze D, Ivanenko A, et al. Pharmacologic management of insomnia in children and adolescents: consensus statement. Pediatrics 2006;117(6):e1223−32.
[43] Klokkerud M, Dagfinrud H, Uhlig T, Dager TN, Furunes KA, Klokkeide A, et al. Developing and testing a consensus-based core set of outcome measures for rehabilitation in musculoskeletal diseases. Scand J Rheumatol 2018;47(3):225−34.
[44] Allin BSR, Hall NJ, Ross AR, Marven SS, Kurinczuk JJ, Knight M, et al. Development of a gastroschisis core outcome set. Arch Dis Child Fetal Neonatal Ed 2019;104(1):F76−82.
[45] Allin BSR, Bradnock T, Kenny S, Kurinczuk JJ, Walker G, Knight M, et al. NETS(1HD) study: development of a Hirschsprung's disease core outcome set. Arch Dis Child 2017;102: 1143−51.
[46] Warners MJ, Hindryckx P, Levesque BG, Parker CE, Shackelton LM, Khanna R, et al. Systematic review: disease activity indices in eosinophilic esophagitis. Am J Gastroenterol 2017; 112(11):1658−69.
[47] Dohner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Buchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood 2017;129:424−47.
[48] McNamara RL, Spatz ES, Kelley TA, Stowell CJ, Beltrame J, Heidenreich P, et al. Standardized outcome measurement for patients with coronary artery disease: consensus from the International consortium for Health Outcomes Measurement (ICHOM). J Am Heart Assoc 2015;4(5):e001767.
[49] Marrie RA, Miller A, Sormani MP, Thompson A, Waubant E, Trojano M, et al. Recommendations for observational studies of comorbidity in multiple sclerosis. Neurology 2016;86:1446−53.
[50] Carlson GA, Jensen PS, Findling RL, Meyer RE, Calabrese J, DelBello MP, et al. Methodological issues and controversies in clinical trials with child and adolescent patients with bipolar disorder: report of a consensus conference. J Child Adolesc Psychopharmacol 2003;13(1):13−27.
[51] Chow E, Wu JSY, Hoskin P, Coia LR, Bentzen SM, Blitzer PH. International consensus on palliative radiotherapy endpoints for future clinical trials in bone metastases. Radiother Oncol 2002;64(3): 275−80.
[52] Bellm LA, Cunningham G, Durnell L, Eilers J, Epstein JB, Fleming T, et al. Defining clinically meaningful outcomes in the evaluation of new treatments for oral mucositis: oral mucositis patient provider advisory board. Cancer Invest 2002;20(5−6): 793−800.
[53] Meuleman C, Tomassetti C, D'Hooghe TM. Clinical outcome after laparoscopic radical excision of endometriosis and laparoscopic

segmental bowel resection. Curr Opin Obstet Gynecol 2012;24(4):245−52.
[54] Schellinger PD, Bath PM, Lees KR, Bornstein NM, Uriel E, Eisert W, et al. Assessment of additional endpoints for trials in acute stroke - what, when, where, in who? Int J Stroke 2012;7(3):227−30.
[55] Sharrock AE, Barker T, Yuen HM, Rickard R, Tai N. Management and closure of the open abdomen after damage control laparotomy for trauma. A systematic review and meta-analysis. Injury 2016;47:296−306.
[56] van den Bos W, Muller BG, Ahmed H, Bangma CH, Barret E, Crouzet S, et al. Focal therapy in prostate cancer: international multidisciplinary consensus on trial design. Eur Urol 2014;65(6):1078−83.
[57] Wolters PL, Martin S, Merker VL, Gardner KL, Hingtgen CM, Tonsgard JH, et al. Patient-reported outcomes in neurofibromatosis and schwannomatosis clinical trials. Neurology 2013;81:S6−14.
[58] Wallace SJ, Worrall L, Rose T, Le Dorze G, Breitenstein C, Hilari K, et al. A core outcome set for aphasia treatment research: the ROMA consensus statement. Int J Stroke 2019;14(2):180−5.
[59] Barlow GD, Lamping DL, Davey PG, Nathwani D. Evaluation of outcomes in community-acquired pneumonia: a guide for patients, physicians, and policy-makers. Lancet Infect Dis 2003;3(8):476−88.
[60] Griffiths AM, Otley AR, Hyams J, Quiros AR, Grand RJ, Bousvaros A, et al. A review of activity indices and end points for clinical trials in children with Crohn's disease. Inflamm Bowel Dis 2005;11(2):185−96.
[61] van Riel PL. Provisional guidelines for measuring disease activity in clinical trials on rheumatoid arthritis. Br J Rheumatol 1992;31(12):793−4.
[62] Wilde EA, Whiteneck GG, Bogner J, Bushnik T, Cifu DX, Dikmen S, et al. Recommendations for the use of common outcome measures in traumatic brain injury research. Arch Phys Med Rehabil 2010;91:1650−1660 e17.
[63] Grieve S, Perez R, Birklein F, Brunner F, Bruehl S, Harden RN, et al. Recommendations for a first Core Outcome Measurement set for complex regional PAin syndrome Clinical sTudies (COMPACT). Pain 2017;158(6):1083−90.
[64] Obbarius A, van Maasakkers L, Baer L, Clark DM, Crocker AG, de Beurs E, 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.
[65] Webster L, Groskreutz D, Grinbergs-Saull A, Howard R, O'Brien JT, Mountain G, et al. Core outcome measures for interventions to prevent or slow the progress of dementia for people living with mild to moderate dementia: systematic review and consensus recommendations. PLoS One 2017;12:e0179521.
[66] Haywood K, Whitehead L, Nadkarni VM, Achana F, Beesems S, Bottiger BW, 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.
[67] Hopkins C, Hettige R, Soni-Jaiswal A, Lakhani R, Carrie S, Cervin A, et al. 25 CHronic rhinosinusitis outcome MEasures (CHROME) − developing a core outcome set for trials of interventions in chronic rhinosinusitis. Rhinology 2018;56(1):22−32.
[68] Fitzpatrick R, Chambers J, Burns T, Doll H, Fazel S, Jenkinso C, et al. A systematic review of outcome measures used in forensic mental health research with consensus panel opinion. Health Technol Assess 2010;14:1−94.
[69] Khanna D, Lovell DJ, Giannini E, Clements PJ, Merkel PA, Seibold JR, et al. Development of a provisional core set of response measures for clinical trials of systemic sclerosis. Ann Rheum Dis 2008;67(5):703−9.
[70] Becker LB, Aufderheide TP, Geocadin RG, Callaway CW, Lazar RM, Donnino MW, et al. Primary outcomes for resuscitation science studies: a consensus statement from the American Heart Association. Circulation 2011;124:2158−77.
[71] Pavletic SZ, Martin P, Lee SJ, Mitchell S, Jacobsohn D, Cowen EW, et al. Measuring therapeutic response in chronic graft-versus-host disease: national Institutes of Health Consensus Development Project on Criteria for Clinical Trials in Chronic Graft-versus-Host disease: IV. Response criteria working group report. Biol Blood Marrow Transplant 2006;12(3):252−66.
[72] Merkel PA, Herlyn K, Mahr AD, Neogi T, Seo P, Walsh M, et al. Progress towards a core set of outcome measures in small-vessel vasculitis. Report from OMERACT 9. J Rheumatol 2009;36:2362−8.
[73] Distler O, Behrens F, Pittrow D, Huscher D, Denton CP, Foeldvari I, et al. Defining appropriate outcome measures in pulmonary arterial hypertension related to systemic sclerosis: a Delphi consensus study with cluster analysis. Arthritis Rheum 2008;59:867−75.
[74] Miller FW, Rider LG, Chung YL, Cooper R, Danko K, Farewell V, et al. Proposed preliminary core set measures for disease outcome assessment in adult and juvenile idiopathic inflammatory myopathies. Rheumatology 2001;40(11):1262−73.
[75] Vargus-Adams JN, Martin LK. Measuring what matters in cerebral palsy: a breadth of important domains and outcome measures. Arch Phys Med Rehabil 2009;90:2089−95.
[76] Needham DM, Sepulveda KA, Dinglas VD, Chessare CM, Friedman LA, Bingham CO 3rd, et al. Core outcome measures for clinical research in acute respiratory failure survivors. An international modified Delphi consensus study. Am J Respir Crit Care Med 2017;196:1122−30.
[77] Gladman DD, Mease PJ, Strand V, Healy P, Helliwell PS, Fitzgerald O, et al. Consensus on a core set of domains for psoriatic arthritis. J Rheumatol 2007;34:1167−70.
[78] Nikiphorou E, Mackie SL, Kirwan J, Boers M, Isaacs J, Morgan AW, et al. Achieving consensus on minimum data items (including core outcome domains) for a longitudinal observational cohort study in rheumatoid arthritis. Rheumatology (Oxford) 2017;56(4):550−5.
[79] Boers M, Kirwan JR, Gossec L, Conaghan PG, D'Agostino MA, Bingham CO III, et al. How to choose core outcome measurement sets for clinical trials: OMERACT 11 approves filter 2.0. J Rheumatol 2014;41:1025−30.
[80] Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines: 1. Introduction-GRADE evidence profiles and summary of findings tables. J Clin Epidemiol 2011;64:383−94.
[81] Vincent K, Kennedy S, Stratton P. Pain scoring in endometriosis: entry criteria and outcome measures for clinical trials. Report from the Art and Science of Endometriosis meeting. Fertil Steril 2010;93:62−7.
[82] Rief W, Burton C, Frostholm L, Henningsen P, Kleinstauber M, Kop WJ, et al. Core outcome domains for clinical trials on somatic symptom disorder, bodily distress disorder, and functional somatic syndromes: European network on somatic symptom disorders recommendations. Psychosom Med 2017;79(9):1008−15.
[83] Goldhahn J, Beaton D, Ladd A, Macdermid J, Hoang-Kim A, Distal Radius Working Group of the International Society for Fracture R. Recommendation for measuring clinical outcome in distal radius fractures: a core set of domains for standardized reporting in clinical practice and research. Arch Orthop Trauma Surg 2014;134(2):197−205.
[84] Wallace SJ, Worrall L, Rose T, Le Dorze G, Breitenstein C, Hilari K, et al. A core outcome set for aphasia treatment research: the ROMA consensus statement. Int J Stroke 2018. 1747493018806200.
[85] Hellmich B, Flossmann O, Gross WL, Bacon P, Cohen-Tervaert JW, Guillevin L, et al. EULAR recommendations for conducting clinical studies and/or clinical trials in systemic vasculitis: focus on anti-neutrophil cytoplasm antibody-associated vasculitis. Ann Rheum Dis 2007;66(5):605−17.
[86] Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the

methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539−49.
[87] Ruemmele FM, Hyams JS, Otley A, Griffiths A, Kolho KL, Dias JA, et al. Outcome measures for clinical trials in paediatric IBD: an evidence-based, expert-driven practical statement paper of the paediatric ECCO committee. Gut 2015;64(3):438−46.
[88] Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. Spine 2000;25(24):3100−3.
[89] Micke O, Seegenschmiedt MH, German Working Group on Radiotherapy in Germany. Consensus guidelines for radiation therapy of benign diseases: a multicenter approach in Germany. Int J Radiat Oncol Biol Phys 2002;152(2):496−513.
[90] Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. Arthritis Rheum 1993;36:729−40.
[91] Ruiz CE, Hahn RT, Berrebi A, Borer JS, Cutlip DE, Fontana G, et al. Clinical trial principles and endpoint definitions for paravalvular leaks in surgical prosthesis: an expert statement. J Am Coll Cardiol 2017;69:2067−87.
[92] Stoner MC, Calligaro KD, Chaer RA, Dietzek AM, Farber A, Guzman RJ, et al. Reporting standards of the Society for Vascular Surgery for endovascular treatment of chronic lower extremity peripheral artery disease. J Vasc Surg 2016;64(1):e1−21.
[93] Williams MR, Nayshtut M, Hoefnagel A, McKeown A, Carlson DW, Cravero J, et al. Efficacy outcome measures for pediatric procedural sedation clinical trials: an ACTTION systematic review. Anesth Analg 2018;126:956−67.
[94] Hudak PL, Wright JG. The characteristics of patient satisfaction measures. Spine 2000;25(24):3167−77.
[95] Dworkin RH, Turk DC, Farrar JT, Haythornthwaite JA, Jensen MP, Katz NP, et al. Core outcome measures for chronic pain clinical trials: IMMPACT recommendations. Pain 2005;113(1−2):9−19.
[96] Seegenschmiedt MH, Muller RP. Development of guidelines in radiation oncology. Strahlenther Onkol 1997;173:345−51.
[97] Muller RP, Seegenschmiedt MH, editors. Leitlinien in der Radioonkologie. Tubingen: DEGRO: Eigenverlag; 1999.
[98] Hillner BE, Ingle JN, Berenson JR, Janjan NA, Albain KS, Lipton A, et al. American Society of Clinical Oncology guideline on the role of bisphosphonates in breast cancer. J Clin Oncol 2000;18:1378−91.
[99] Cook DL, Guyatt GH, Laupacis A, Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. Chest 1989;102:305S−11S.
[100] Sackett DL. Rules of evidence and clinical recommendations on the use of anti-thrombotic agents. Chest 1989;102:2S−4S.
[101] Bowman GP, Levine MN, Mohide EA, al e. The practice guidelines development cycle: a conceptual tool for practice guidelines and implementation. J Clin Oncol 1995;13:502−12.
[102] Schmitt J, Apfelbacher C, Spuls PI, Thomas KS, Simpson EL, Furue M, et al. The Harmonizing Outcome Measures for Eczema (HOME) roadmap: a methodological framework to develop core sets of outcome measurements in dermatology. J Invest Dermatol 2015;135:24−30.
[103] Tugwell P, Bombardier C. A methodologic framework for developing and selecting endpoints in clinical trials. J Rheumatol 1982;9:758−62.
[104] COSMIN database of systematic reviews of outcome measurement instruments. Available at https://database.cosmin.nl/. Accessed June 6, 2019.
[105] Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1147−57.
[106] Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1171−9.
[107] COMET initiative database search. Available at http://www.comet-initiative.org/studies/search. Accessed April 23, 2019.
[108] Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res 2018;27:1159−70.
[109] Ju A, Chow BY, Ralph AF, Howell M, Josephson MA, Ahn C, et al. Patient-reported outcome measures for life participation in kidney transplantation: a systematic review. Am J Transplant 2019;19:2306−17.
[110] Beaton DE, Maxwell LJ, Shea BJ, Wells GA, Boers M, Grosskleg S, et al. Instrument selection using the OMERACT filter 2.1: the OMERACT methodology. J Rheumatol 2019;46(8):1028−35.