Challenges to Global Standardization of Outcome Measures

Zoe Liao, B.S. Pharmacy[1], Yuri Quintana PhD[2,3]
[1]School of Pharmacy, Northeastern University, Boston, MA, USA; [2]Division of Clinical Informatics, Beth Israel Deaconess Medical Center, Boston, MA, USA; [3]Harvard Medical School, Boston, MA, USA

Abstract

Global standardization of outcome measures for disease states can help researchers and healthcare providers compare healthcare institutions' and populations' health outcomes. Despite the creation of standardized outcome sets, clinical institutions' adoption of these sets is not common. A literature review shows that among the challenges to standardizing outcome measures include the difficulties of achieving consensus in the working groups creating these outcome sets, the tradeoffs made when selecting outcome measurement tools, and the high costs of implementing a new or different set of outcome measures. The duplication of effort to create these standard sets can also limit standardization, which could be minimized through increased transparency of how these standard sets are developed. We propose some approaches to improve how to create and implement standard sets to broaden their usability across institutions.

Introduction

As healthcare spending increases across the world, many nations are reforming their healthcare systems to prioritize measuring the value of care over the volume of care. The "value" of health care is defined as "the outcomes patients may experience relative to the cost of delivering these outcomes."[1] Scaling and continuously collecting these outcome data is now possible given the advancement of technological capabilities.[2] Moreover, the systematic use of information collected from patient-reported outcome measures has improved patient-provider communication and patient satisfaction with health care.[2,3] However, even for the same disease state, there is a wide variation in the outcomes recorded in electronic health records (EHRs), claims databases, patient registries, and prescription databases.[2] There is also a variation of outcomes measured in the design of clinical trials for the same disease.[4] On top of this, the actual measures used to evaluate the desired outcomes are heterogeneous. The current lack of global standardization of outcome measures hinders direct comparisons and meta-analyses of clinical trials. It detracts from global learning and goals of using longitudinal data for comparison of intervention effects.[1]

Achieving consensus on key considerations–which outcomes to include, how to measure them, and when to measure them–is a lengthy and challenging process. In 2010, the Core Outcome Measures in Effectiveness Trials (COMET) Initiative created a free, online database detailing the ongoing studies that aim to apply rigorous consensus methods to develop core outcome sets (COS).[4] Core outcome sets are defined as the minimum sets of outcomes that should be measured and reported in all clinical trials of a specific disease or for application in disease registries or clinical practice.[4]

Among the collaboratives that have been working to create these core outcome sets is the nonprofit International Consortium for Health Outcomes Measurement (ICHOM), which prioritizes the incorporation of patient-reported outcomes measures (PROMs) into COS they call "Standard Sets." These PROMs include symptoms, health-related quality of life, and satisfaction of care.[1,5,6] As of 2018, ICHOM Working Groups, which consist of clinicians, researchers, and patient representatives around the world, have developed Standard Sets that cover 54% of the global disease burden.[4] However, from their start in 2012, the only studies documenting the implementation or feasibility of using ICHOM Standard Sets have been conducted or funded by ICHOM themselves.[1,5] This may be due to competing efforts of other groups developing standard sets for the same diseases and the financial and logistical challenges for institutions to start implementing standard outcome sets. This paper will focus on the problems and potential solutions for different goals and compositions of the groups creating these outcome sets, the tradeoff required when choosing between the utility of an outcome measure and its feasibility for collection, the high financial costs of implementation, and the global applications.

Methods

A literature search was conducted in the PubMed database using health outcome measures, outcome assessments, and global health standards. The disease states searched were focused on cardiovascular disease, oncology, diseases common for the elderly, and mental or behavioral health. This was intended to explore any differences between healthcare areas with better infrastructure for care (e.g., cardiovascular disease, oncology) compared to areas with more fragmented care (e.g., mental health care, elder care). Both COS development and COS implementation studies were included to look for potential relationships between how core outcome sets are developed and how they are implemented.

Results

A literature search conducted on PubMed in June 2020 and results were screened for inclusion. Eight studies describing the standard sets for prostate cancer, dementia, heart failure, and hip and knee osteoarthritis, and behavioral health were included for analysis (Table 1).

Table 1. Summary of studies included in the review

Author (Year)	Disease State	Type of study: COS development, COS implementation, Both
Seligman et al. [1] (2018)	Cardiovascular diseases	Both (general overview and case study)
Meregaglia et al. [2] (2020)	Prostate cancer	Both (scoping review and case study)
Webster et al. [4] (2017)	Dementia	COS development (systematic review and consensus)
Ackerman et al. [5] (2018)	Osteoarthritis	COS implementation (feasibility)
Martin et al. [6] (2015)	Prostate cancer	COS development
McNamara et al. [7] (2015)	Coronary artery disease	COS development
Rajaram et al. [8] (2019)	Breast cancer	COS implementation (cross-sectional comparison)
Wing et al. [12] (1998)	Behavioral Health	COS development

Composition of Groups Creating Standard Sets

The major organizations developing core outcomes sets were the International Consortium for Health Outcomes Measurement (ICHOM) and the Core Outcome Measures in Effectiveness Trials (COMET) Initiative. ICHOM was founded in 2012 by Harvard Business school, The Boston Consulting Group, and The Karolinska Institute, through their funding and funding from various international sponsors.[5-9] Their goal for creating standardized, open-access sets of outcome measures is to include outcomes that matter to patients, as well as outcomes that can be tracked across different health systems and clinical registries. The members they seek in their working groups include both clinicians and non-clinicians around the world. The Core Outcome Measures in Effectiveness Trials (COMET) Initiative is an organization that provides methodological support to groups trying to develop core outcome sets.[2,10,11] The multidisciplinary organization grew from a 2010 meeting of researchers, regulators, and policymakers interested in developing core outcomes sets to improve the standards for data reporting and synthesis in clinical trials. Their publicly available database of ongoing COS development studies aims to promote collaboration among researchers as well as the application of the developed COS.[2]

The composition of the working groups creating these sets can impact the consensus of decisions made.[1] In a review of core outcome sets for prostate cancer, Meregaglia et al.[2] found that there were "notable gaps in reporting the 'stakeholders involved' and 'consensus process' adopted," and that "geographic representativeness of stakeholders was unbalanced in favor of Europe and North America." This scoping review applied the COS-STANdards for Development framework developed by COMET to assess COS development studies' quality systematically. Reviewers found that the ICHOM study protocol for the development of the Standard Set for Prostate Cancer did not report if the scoring process, the definition of consensus, and the criteria for including/adding/dropping outcomes, were determined a priori.[2,6] Furthermore, the Methods section of the same ICHOM study protocol does not disclose the extent of patient involvement or provide any background information on the patient representatives.[6] A similar

critique could be made for the study describing how ICHOM developed an outcome set for coronary artery disease patients.[7]

Discussion

Deciding Between Outcome Measures

A key aspect to consider in the development of COS is the tradeoff when choosing between the comprehensiveness of outcome measures and the feasibility of collecting such outcome measures. For ICHOM's prostate cancer Standard Set, Martin et al. [6] chose the Expanded Prostate Cancer Index Composite 26-question short form (EPIC-26) to measure PROMs instead of the validated EPIC-16 instrument. Although EPIC-16 was designed for easy implementation, ICHOM went with the lengthier EPIC-26 because it included a question on rectal bleeding, which they considered crucial since it could indicate late toxicity from radiation. ICHOM also did not have common instruments like the International Prostate Symptom Score and International Index of Erectile Function as part of their Standard Set due to the overlap with the EPIC-26 domains.

Another example of the difficulty of choosing between measurement tools can be found in developing a COS for dementia disease-modifying clinical trials.[4] Outcomes from ICHOM's Standard Set for dementia were considered for inclusion. The COS developers recommended either the Alzheimer's Disease Assessment Scale-Cognitive Subscale (ADAS-Cog) or the Mini-Mental State Examination (MMSE) for measuring cognition due to the difficulty of choosing between cost and time. The ADAS-Cog is free but can only be administered by a trained tester and takes 45 minutes to administer. The MMSE can be administered by clinical staff with minimal extra training, but it is costly due to copyright. Ultimately, the decision of which measurement instrument to use was left to the COS user due to the dependence of feasibility and practicality on the user's resources and current workflow.

Care priorities can also look different among various socioeconomic and cultural groups. In a cross-sectional comparison study of patient-reported outcome measures among breast cancer survivors in Malaysia versus high-income countries,[9] investigators found that well-being, survival, and physical functioning were the most important PROMs for Malaysians and high-income country patients. However, Malaysian breast cancer survivors were less likely to rate social, emotional, cognitive, and sexual functioning as very important. Instead, they were more likely to prioritize symptoms and complications management. A gap analysis between the ICHOM Heart Failure Standard Set and a global selection of real-world data sources revealed that "data captured in data sources from North America and Europe more closely resembled the Standard Set, whereas data sources in Africa deviated the most."[1] Different countries also have different healthcare systems and technological capabilities. ICHOM considered this when developing their coronary artery disease COS, where they restricted longitudinal outcomes to those that could be captured as administrative data since countries without a single-payer health system could have trouble identifying events outside of the specific acute care episode in their registries or electronic health record databases.[7]

Cost of Implementation

The high costs of implementing COS serve as a barrier to the global standardization of outcomes. A feasibility study of implementing the ICHOM Standard Set for Hip and Knee Osteoarthritis in two hospitals in Australia calculated the costs of implementation and 17 months of data collection to be 94,955 AUD ($65,234 USD). This amount accounted for project coordinator time, IT support, ICHOM implementation support, equipment, and physiotherapist support for recruitment.[5] Costs will vary for institutions and disease states, depending on the available clinical, administrative, and IT support available as well as patient volume. Financial costs may increase as the time for data collection continues, since ICHOM recommends long-term or possibly lifetime data collection to properly assess PROMs,[6] but the cost per patient could decrease over time.[5] Rates of COS implementation would increase with financial support from research funders, trial registries, and policymakers.[2] For example, the UK National Institute of Health Research (NIHR) funded the COS development of the COS for disease modification of dementia so that future NIHR-funded trials will use that COS.[4]

Development Process

Studies outlining the development of the standard sets should provide transparency not only on the composition and backgrounds of the working group members but also on how and when the criteria for including outcome measures

were decided. The Methods section of various ICHOM study protocols do not disclose the extent of patient involvement or provide any background information on the patient representatives, limiting the generalizability of the developed outcome sets.[2,6,7] This is especially important, as patients or patient representatives from different socioeconomic, cultural, and educational backgrounds may have different perspectives on which outcomes are important to them. Furthermore, if the Standard Set developers change the criteria after conducting a Delphi survey,[2] this could introduce bias.

Different outcome sets for the same disease state are being created, with the main reason being cited as the difference in use case. Although the use case for ICHOM Standard Sets is intended for clinical practice, there is a potential for overlap between outcomes assessed for practice and research.[10] Groups developing COS for research purposes can incorporate PROMs from ICHOM Standard Sets. There is a greater chance of collaboration if all COS creators are transparent in their COS development protocols and register their studies on the COMET database to prevent the duplication of effort. This will make it easier and more evident to see if projects can be complementary. Collaboration is ideal for aligning standards across contexts, reducing the spread of limited resources for implementation.

Global Mental Health

In 1993, the United Kingdom sought to have some standardized behavioral health outcomes, and in 1998 the Health of the Nation Outcome Scales (HoNOS) were introduced[12]. HoNOS-1 is a 20-item instrument that covers four key areas of functioning of patients: behavior, impairment, symptoms, and social functioning. A study[13] in 2020 looked for evidence of its value or cost-effectiveness to consumers, clinicians, or administrators. Of the 260 studies reviewed, only one study reported positive outcomes, and none of them attempted to assess the cost of using the Health of the Nation Outcome Scale (HoNOS). The study investigated the effect of routine outcome measurement but concluded that it failed to result in the provision of evidence-based care. To date, the ability of HoNOS to improve the health and social functioning of mentally ill people has not been demonstrated. A very recent study[14] suggested that clinician sensitivity and bias may affect the use of the instrument. ICHOM has developed a set of outcome measures for anxiety[16], depression[17], and addiction[18], but it is too early to know if groups will adopt it. International collaboration for increasing mental health services and research capacity in Africa emphasizes the need for cooperation between institutions and training for the successful use of evidence-based knowledge.[15] Informatics training on terminology and data modelling will be needed to successfully collect and apply outcomes for health care improvement.

Global Application

The creation of standard sets for global use should also consider what outcome measures institutions are already using. For example, clinical practices already using common instruments like the International Prostate Symptom Score and International Index of Erectile Function, which ICHOM did not include in their Standard Set,[6] could face disruption in the longitudinal data collected if they switch to ICHOM's proposed measure of EPIC-26. The advantages and limitations of measurement tools proposed (i.e., high cost, not yet validated, only available in certain languages, etc.) should be described by COS developers. Standardized sets are more likely to be implemented if the measurement tools proposed to take less time and resources to use than measurement tools previously used at the institution.[1,5] ICHOM suggests that future work should make commonly used outcome measures more comparable to transition to a universal standard.[6] In the meantime, it could help potential users if Standard Sets specifically outlined which outcome domains overlap among measurement tools. Data in the same domains can be collected at the same time points to make future comparisons easier.

Since the goal is global standardization, COS developers should keep in mind that different countries will vary in resources and in the volume of changes that will need to be made to current data collection processes to standardize. Countries with the advantage of recent investments in registry infrastructure, such as Ireland, Canada, Australia, and the U.S.,[6] should report on their implementations of standardized outcome sets to improve them for countries with more limited resources to spare. Developers of core outcome sets should describe or suggest ways to collect the requested data, such as through administrative data or electronic health record databases or registries.

Conclusion

The global standardization of outcome measures, for both clinical trials and clinical practice, can allow institutions to learn from each other about which interventions are best for improving patient outcomes and reduce the cost of care

through the elimination of ineffective interventions. Many groups are developing core outcome sets with the goal for international implementation, but many barriers currently stand in the way. These barriers include duplication of effort by different groups due to a lack of transparency in the study protocols for COS development, the limitations and compromises of recommended outcome measures, and the financial challenges of implementing long-term COS data collection recommendations. Solutions that address one challenge can also help minimize other challenges. Collaboration among developers can identify overlapping outcome domains for research and clinical practice, reducing the number of measures institutions would have to implement for the same disease state. Transparency of study protocols would clarify which outcome measures are best suited for the institution looking to implement the COS, based on resources and patient population. Future reports on the implementation and performance of this COS should provide insights into improving COS quality. An increase in the number of successful deployments will hopefully convince hesitant institutions and countries to prioritize adopting a common COS until eventually the goal of standardizing health outcome measurements across the world is achieved. Training on medical informatics processes for data collection and representation will be needed to succefully collect and apply these outcome standards.

References

1. Seligman WH, Salt M, la Torre Rosas AD, Das-Gupta Z. Unlocking the potential of value-based health care by defining global standard sets of outcome measures that matter to patients with cardiovascular diseases. *Eur Heart J.* 2018;5(2):92-95. doi: 10.1093/ehjqcco/qcy056
2. Meregaglia M, Ciani O, Banks H, et al. A scoping review of core outcome sets and their 'mapping' onto real-world data using prostate cancer as a case study. *BMC Med Res Methodol.* 2020;20(1):41-41. doi: 10.1186/s12874-020-00928-w
3. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. *BMJ*. 2015;350:g7818. Published 2015 Feb 10. doi:10.1136/bmj.g7818
4. Webster L, Groskreutz D, Grinbergs-Saull A, 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(6):e0179521-e0179521. doi:10.1371/journal.pone.0179521
5. Ackerman IN, Cavka B, Lippa J, Bucknill A. The feasibility of implementing the ICHOM standard set for hip and knee osteoarthritis: a mixed-methods evaluation in public and private hospital settings. *J Patient Rep Outcomes*. 2018;2:32-32. doi: 10.1186/s41687-018-0062-5
6. Martin NE, Massey L, Stowell C, et al. Defining a standard set of patient-centered outcomes for men with localized prostate cancer. *Eur Urol*. 2015;67(3):460-467. doi:10.1016/j.eururo.2014.08.075
7. McNamara RL, Spatz ES, Kelley TA, 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. Published 2015 May 19. doi:10.1161/JAHA.115.001767
8. ICHOM: International Consortium for Health Outcomes Measurement. ICHOM website. https://www.ichom.org/. Accessed August 27, 2020.
9. Rajaram N, Lim ZY, Song CV, et al. Patient-reported outcome measures among breast cancer survivors: A cross-sectional comparison between Malaysia and high-income countries. *Psychooncology*. 2019;28(1):147-153. doi:10.1002/pon.4924
10. MacLennan S, Williamson PR, Lam TB. Re: Neil E. Martin, Laura Massey, Caleb Stowell, et al. Defining a standard set of patient-centered outcomes for men with localized prostate cancer. *Eur Urol* 2015;67:460–7.
11. COMET Initiative: Core Outcome Measures in Effectiveness Trials. COMET Initiative website. http://www.comet-initiative.org/. Accessed August 27, 2020.
12. Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). Research and development. *Br J Psychiatry*. 1998 Jan;172:11-8. doi: 10.1192/bjp.172.1.11. PMID: 9534825.
13. Bender KG. The meagre outcomes of HoNOS. *Australas Psychiatry.* 2020 Apr;28(2):206-209. doi: 10.1177/1039856219875066. Epub 2019 Sep 30. PMID: 31564114.
14. Williams, J. Clinicians' ratings of the Health of the Nation Outcome Scales (HoNOS) show sensitivity/bias that may affect patients' progress: analyses of routine administrative data. Preprint available at medRxiv 2020.11.19.20234674; doi: https://doi.org/10.1101/2020.11.19.20234674
15. Gureje O, Seedat S, Kola L, Appiah-Poku J, Othieno C, Harris B, Makanjuola V, Price LN, Ayinde OO, Esan O. Partnership for mental health development in Sub-Saharan Africa (PaM-D): a collaborative

initiative for research and capacity building. *Epidemiol Psychiatr Sci*. 2019 Aug;28(4):389-396. doi: 10.1017/S2045796018000707. Epub 2018 Nov 27. PMID: 30479242; PMCID: PMC6536364.
16. International Consortium For Health Outcomes Measurement. ICHOM Standard Set for Depression & Anxiety Available at URL: https://www.ichom.org/portfolio/depression-anxiety/
17. International Consortium For Health Outcomes Measurement. ICHOM Standard Set for Children & Young People with Depression & Anxiety. Available at URL https://www.ichom.org/portfolio/anxiety-depression-ocd-and-ptsd-in-children-and-young-people/
18. International Consortium For Health Outcomes Measurement. ICHOM Standard Set for Addiction. Available at URL: https://www.ichom.org/portfolio/addiction/