



## RESEARCH ARTICLE

**REVISED** A pilot study assessing the similarity between core outcome sets and outcomes included in health technology assessments [version 3; peer review: 2 approved]

Previously titled: A pilot study assessing the similarity between core outcome sets and outcomes included in health technology assessments

Peter Cox <sup>1</sup>, Paula R. Williamson<sup>2</sup>, Susanna Dodd <sup>2</sup>

<sup>1</sup>Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

<sup>2</sup>Department of Health Data Science, University of Liverpool (a member of Liverpool Health Partners), Liverpool, UK

**V3** First published: 25 Oct 2021, 10:1084  
<https://doi.org/10.12688/f1000research.73647.1>  
Second version: 28 Feb 2022, 10:1084  
<https://doi.org/10.12688/f1000research.73647.2>  
Latest published: 03 May 2022, 10:1084  
<https://doi.org/10.12688/f1000research.73647.3>

**Abstract**





**Objective:** Core outcome sets (COS) are an agreed standardised collection of outcomes created with representation from all key stakeholders (such as patients, clinicians, researchers), which should be reported as a minimum for all trials in that corresponding clinical area. There has been little research investigating the use of core outcomes in Health technology assessments (HTAs) and none in non-oncology HTAs. This study aimed to assess the similarity between COS and HTA outcomes.


**Methods:** Ten COS published between 2015 and 2019 were selected, with patient participation taken as a proxy measure for a high quality COS. The INAHTA database was used as a source to identify relevant HTAs, which were accessed through the hyperlinks provided. Outcomes selected for these assessments were categorised as either a specific, partial or no match compared to the COS. An additional cohort of non-oncology HTAs published between 2019 and 2021 were identified from the NICE website and compared against a relevant COS.


**Results:** Six hundred and fifty-one HTAs were matched to the ten COS areas, of which 119 were reviewed. Of a possible 1318 core outcome matches, there were 562 (43%) matches, 413 (31%) specific and 149 (11%) partial. NICE HTA matches against corresponding COS ranged from 44% to 100%, with a total of 78% (73/94) matches, 57 (61%) specific and 16 (17%) partial.


**Conclusion:** Further work is required to promote the awareness and implementation of COS within HTAs. The degree of matching between COS and NICE HTA outcomes is encouraging, demonstrating

**Open Peer Review****Approval Status**  

	1	2
<b>version 3</b> (revision) 03 May 2022		 view
<b>version 2</b> (revision) 28 Feb 2022	 view	  view
<b>version 1</b> 25 Oct 2021	  view	

1. **Catriona McDaid** , University of York, York, UK

**Andrew Mott** , University of York, UK, York, UK

2. **Alison E. Turnbull** , Johns Hopkins University, Baltimore, USA  
Johns Hopkins University, Baltimore, USA

Any reports and responses or comments on the article can be found at the end of the article.

acceptance of COS by HTA producers.

### Keywords

Core outcome set, COS, health technology assessment, HTA, outcome research



This article is included in the [Research on Research, Policy & Culture gateway](#).

**Corresponding author:** Peter Cox ([hlpcox@liverpool.ac.uk](mailto:hlpcox@liverpool.ac.uk))

**Author roles:** **Cox P:** Data Curation, Formal Analysis, Investigation, Project Administration, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; **R. Williamson P:** Conceptualization, Formal Analysis, Methodology, Supervision, Validation, Visualization, Writing – Review & Editing; **Dodd S:** Conceptualization, Formal Analysis, Methodology, Supervision, Validation, Visualization, Writing – Review & Editing

**Competing interests:** Paula Williamson chairs the COMET Initiative Management Group.

**Grant information:** The author(s) declared that no grants were involved in supporting this work.

**Copyright:** © 2022 Cox P *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Cox P, R. Williamson P and Dodd S. **A pilot study assessing the similarity between core outcome sets and outcomes included in health technology assessments [version 3; peer review: 2 approved]** F1000Research 2022, **10**:1084 <https://doi.org/10.12688/f1000research.73647.3>

**First published:** 25 Oct 2021, **10**:1084 <https://doi.org/10.12688/f1000research.73647.1>

**REVISED Amendments from Version 2**

Following the helpful reviewer suggestions, we have included a new Figure 3 and subgroup analysis describing outcome matches for HTAs developed after COS publication.

**Any further responses from the reviewers can be found at the end of the article**

**Introduction**

Clinical trials are performed to evaluate the effects of treatment interventions, with the gold standard being randomised controlled trials.<sup>1</sup> Fundamentally this includes a test and a control treatment, with random assignment of treatment groups and at least one outcome measure.<sup>2</sup> Researchers decide on outcomes which best answer their research question, and consequently differences arise between trials in the same field as the chosen outcome measures are often of particular relevance to each study. The lack of uniformity amongst trial designs is abundant, demonstrated by a review of 8942 oncology trials which revealed that over 25,000 outcomes occurred only once or twice<sup>3</sup>; furthermore the top cited and most accessed Cochrane reviews in 2009 described problems due to inconsistencies in the reported outcomes in included studies.<sup>4</sup> Discrepancies are also observed when examining how outcomes are measured; for example, a survey of 10,000 trials investigating schizophrenia discovered 2194 different measurement scales were employed.<sup>5</sup> This produces a substantial challenge in data analysis, limiting the ability to compare studies, synthesise the available evidence and perform meta-analysis, ultimately leading to avoidable research waste.<sup>6</sup> There are also questions concerning whether trial outcomes are always relevant to patients or clinicians, meaning statistically significant results may have limited clinical bearing, hence not translating into improved clinical care.<sup>7</sup>

In January 2010 the **Core Outcome Measures in Effectiveness Trials (COMET) Initiative** was launched with the aim of bringing together people interested in the development and application of core outcome sets (COS). COS address the issue of inconsistency and outcome reporting bias, while reducing the difficulties facing systematic reviewers due to heterogeneity in outcome measures.<sup>8</sup> COS are an agreed standardised collection of outcomes which should be reported as a minimum for all trials in the corresponding clinical area. The Core Outcome Set-STANDARD for Development (COS-STAD) provides criteria against which to assess the quality of COS.<sup>9</sup> Researchers are not restricted to the COS, but there is an expectation it will always be reported.<sup>8</sup> The importance of COS has become increasingly acknowledged over time, with COMET endorsed by trial funders such as the **National Institute for Health Research (NIHR)**, the **Cochrane Collaboration** and **National Institute for Health and Care Excellence (NICE)**. There is currently ongoing work assessing the representation of COS in US Food and Drug Administration (FDA) and European Medicines Agency (EMA) regulatory guidance.<sup>10</sup>

Health technology assessments (HTAs) are multidisciplinary processes which use explicit methods to determine the value of a health technology at different points in its lifecycle.<sup>11</sup> They are produced to inform decision makers and promote an equitable, efficient and high-quality health system.<sup>11</sup> The technologies are often interventions designed to prevent, diagnose or treat medical conditions. The International Network of Agencies for Health Technology Assessments (INAHTA) is a network of 51 different HTA agencies which support healthcare decision making. This network produces the **international HTA database**, which provides a free, single point of access to information about ongoing and published HTAs.

Recent work has assessed the International Consortium for Health Outcomes Measurement (ICHOM) outcome sets against oncology HTAs, finding that HTAs tend to focus on generic measures which allow comparison across disease areas. However, it should be noted that these outcome sets are developed for routine care, unlike COS which are produced specifically for research so are more relevant for HTAs.<sup>12</sup> Additionally, a review of technology appraisal oncology scopes found that in the majority of cases there was complete overlap with those outcomes in COS; in a small number of exceptions the COS included two additional outcomes to those specified in the scope.<sup>13</sup> Other work has investigated a trial funder and member of INAHTA, the National Institute for Health Research Health Technology Assessment (NIHR HTA), examining their recommendation for applicants to search for a COS to include in their trial. By examining research applications and then surveying applicants it was found that 38% (36/95) searched for a COS through either the COMET database or another method, e.g. review of the literature; where a published COS existed it was included in 29% (7/24) of cases.<sup>14</sup> There is reason to believe that uptake may have increased, as a limitation was that no COS existed for 68 of the 95 studies assessed at the time of submission.<sup>14</sup> However, this study assessed the uptake of COS in trials rather than HTAs, and to our knowledge, ours is the first study to examine the similarity between COS for research and HTA outcomes. There is potential for COS to further improve the quality of HTAs, providing more uniform and patient-centred evidence to inform clinical guidelines and policy makers, such as NICE.

The purpose of this study was to assess the similarity between COS and outcomes included in HTAs.

## Methods

The **COMET database**, a database of studies relevant to the development of COS, was searched to select individual COS for this review. As a surrogate measure to ensure quality, only COS published between 2015 and 2019 where patients were included were selected, maximising the number of COS-STAD standards met.<sup>9</sup> The methodology of producing COS has improved with time, hence this 5-year period was chosen to include the most recent COS, which had a greater likelihood of including outcomes most relevant to all appropriate stakeholders. Additionally, COS developed for common diseases and interventions were selected to ensure sufficient HTAs were available for assessment. COS outcomes were extracted by one author (PC) from a pre-existing COMET database of outcomes from all COS for research published up to 2019.

The **INAHTA** database was searched in February 2021 for relevant HTAs, using the disease name as the search term with the English language filter applied. Where an intervention was specified by the COS this was included as a search term, e.g. obesity AND surgery. Quotation marks were used to refine searches that returned a large number of HTAs or when the results lacked specificity to the desired condition, e.g. searching type 1 diabetes produced HTAs relating to type 2 and gestational diabetes. There were no restrictions by year or country. MeSH terms were employed when there were variations in how search terms were spelt, for example postpartum haemorrhage versus postpartum hemorrhage.

HTAs were accessed using the hyperlinks provided in the INAHTA database. Where the hyperlink did not direct to the HTA but to the publisher's website, this was searched using the title of the HTA. If a hyperlink did not work, the abstract in the INAHTA record was screened for outcomes if available. HTAs which were irrelevant to the COS, inaccessible (e.g. due to a non-functional hyperlink or no hyperlink was present) or which summarised clinical evidence with no outcome measures stated were excluded. Outcome measures were then extracted from the HTA alongside study identifier, year and data source. Data extraction involved creating a table comprising the core outcomes as headings with each HTA listed below. For each HTA, outcomes were matched to the appropriate headings and then colour coded green, yellow or red to indicate specific, partial or no match respectively. For partial matches, bold and italic text indicated an outcome more general or specific than the COS respectively.

Outcomes were categorised as either a specific match, a partial match or no match compared to the COS, following an approach used previously.<sup>15,16</sup> We considered a match between an HTA outcome and a core outcome to exist if they were either specifically or partially related. We defined a specific match as one where both outcomes corresponded to each other exactly, while a partial match was defined as one where the outcomes correspond to each other non-specifically. Taking the example of swollen joint count as a core outcome, if the HTA included swollen joint count as an outcome it would be considered a specific match, whilst disease activity would be considered a partial match. Partial matches were further categorised according to whether the HTA outcome was either more general or more specific than the COS outcome. The outcome matching process was quality checked by a seconder reviewer (SD). The data extracted from each HTA included the HTA identifier, publisher, publication year and outcome measures. Outcomes were ascertained from stated outcomes of interest or from the PICO (population, intervention, comparator, outcomes) statement. Where outcomes were not explicitly stated, the HTA was reviewed to elicit them and on occasion outcomes had to be ascertained from a series of key research questions. Outcomes identified in the scope of an HTA which could not be reported due to a lack of available data were included in the analysis. Considering the research question "what is the expected beneficial effect of atezolizumab on mortality?" as an example, this was deemed a specific match for overall survival and a partial match for cancer-specific survival.

Within the UK, NICE HTAs are of particular importance. It was noted that no NICE HTAs had been registered on the INAHTA database beyond 2011, so in April 2021, the NICE **website** was searched for HTAs corresponding to each of the COS from the original cohort. Non-oncology conditions were selected given previous work has examined uptake across oncology HTAs.<sup>12,13</sup> and where no appropriate NICE HTAs were identified, alternative COS were selected, using the same criteria of publication between 2015 and 2019 with patient involvement. The most recently published reports which matched a COS were included, with outcomes compared to core outcomes from the corresponding COS.

The results are presented through tables, graphs and using descriptive statistics. The analysis was performed using Excel 2019. Comparison between the INAHTA and NICE results was not examined given they cover different time periods and utilised different COS.

## Results

Ten COS were selected to assess against the INAHTA database. The search terms for each are displayed in **Table 1**. The search relating to the relapsing remitting multiple sclerosis and clinically isolated syndrome COS was performed as two

**Table 1. Summary of search terms and number of search results.**

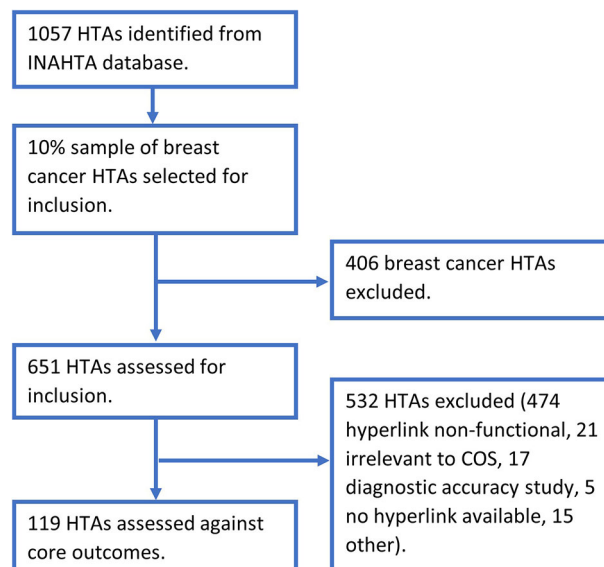
COS	Intervention	Search term	Number of search results
Rheumatoid arthritis <sup>17</sup>	Any	rheumatoid arthritis	135
Relapsing remitting Multiple sclerosis/ Clinically isolated syndrome <sup>†18</sup>	Any	relapsing remitting multiple sclerosis	36
		clinically isolated syndrome	3
Obesity <sup>19</sup>	Metabolic and bariatric surgery	obesity AND surgery	62
Epilepsy <sup>20</sup>	Any	epilepsy	91
Type 1 diabetes <sup>21</sup>	Any	"type 1 diabetes"	67
Acne <sup>22</sup>	Any	acne	7
Coronary artery disease <sup>23</sup>	Any	coronary artery disease	134
Postpartum haemorrhage <sup>24</sup>	Any treatment	"Postpartum Hemorrhage" [mh]	8
Psoriasis <sup>25</sup>	Any	psoriasis	63
Breast cancer <sup>‡26</sup>	Any	"breast cancer"	451

COS, core outcome set.

†The search relating to the relapsing remitting multiple sclerosis/clinically isolated syndrome COS was performed as two individual searches.

‡10% (45) of the breast cancer HTAs were assessed and included in the report.

individual searches to ensure all relevant HTAs were identified. The database search returned a cumulative total of 1057 HTAs, ranging from three for clinically isolated syndrome to 451 for breast cancer. A selection of 10% (45) of the breast cancer results were sampled, with the most recently published HTAs chosen as they were most likely to be accessible, detailed in Figure 1. This selection of HTAs comprised the most recently published HTAs until the 10% sample was reached, encompassing HTAs published between 2015 and 2020. This produced a possible 651 HTAs to be included in the report. Eight of the included COS applied to any intervention, with metabolic and bariatric surgery specified for obesity and treatment specified for postpartum haemorrhage. For each of the ten COS, the majority of the HTAs were not assessed as the hyperlinks listed in the INAHTA database were non-functional, detailed in Table 2. There were several HTAs irrelevant to the COS which were therefore excluded. Other reasons for exclusion included diagnostic accuracy study, HTA where no hyperlink was available, HTA not in English, HTA recommendations, HTA inaccessible on the



**Figure 1. HTA identification and selection flowchart.** COS, core outcome set; HTA, health technology assessment.

**Table 2. Breakdown of database search results for each individual COS.**

Disease name	Number of search results	HTAs assessed	Hyperlink non-functional	Irrelevant to COS	Diagnostic accuracy study	No hyperlink available	Other
RA	135	22	106	4	3	0	0
MS/CIS	39	8	27	1	0	1	2
Obesity	62	10	52	0	0	0	0
Epilepsy	91	15	70	2	3	1	0
T1DM	67	19	41	2	0	0	5
Acne	7	1	5	1	0	0	0
CAD	134	17	104	4	7	0	2
PPH	8	2	2	3	0	0	1
Psoriasis	63	10	47	3	0	1	2
Breast cancer	451 <sup>†</sup>	15	20	1	4	2	3
Total	651	119	474	21	17	5	15

Note: Other: Duplicates, HTA recommendation, HTA not in English, updated version of HTA included, HTA inaccessible on source website. CAD, coronary artery disease; COS, core outcome set; HTA, health technology assessment; MS/CIS, multiple sclerosis/clinically isolated syndrome; PPH, postpartum haemorrhage; RA, rheumatoid arthritis; T1DM, type 1 diabetes mellitus.

<sup>†</sup>10% (45) of the breast cancer HTAs were assessed and included in the report.

source website, duplicates and where an updated version of the HTA had already been assessed. Systematic reviews provided over 90% of studies included, the remainder being trials. To minimise similarity to COS being underestimated by reviews, through outcome reporting being limited to what data exists in the literature, outcomes stated in the methods were included in the analysis, independent of the findings.

Overall, 119 HTAs were assessed against core outcomes with a maximum of 1318 matches possible, shown in [Table 3](#). In total, there were 562 (43%) core outcome matches, 413 (31%) specific and 149 (11%) partial, with the percentage values

**Table 3. The number of outcomes for each COS categorised as specific, partial or no match.**

COS	Number of COs	Specific match (% of total)	Partial match (% of total)		No match (% of total)	Total
			More general	More specific		
Rheumatoid arthritis	8	91 (52)	16 (9)	1 (1)	68 (39)	176
Relapsing remitting Multiple sclerosis/Clinically isolated syndrome	9	28 (39)	1 (1)	2 (3)	41 (57)	72
Obesity	9	34 (38)	4 (4)	0 (0)	52 (58)	90
Epilepsy	8	36 (30)	1 (1)	3 (3)	80 (67)	120
Type 1 diabetes	8	57 (38)	19 (13)	1 (1)	75 (49)	152
Acne	7	3 (43)	0 (0)	1 (14)	3 (43)	7
Coronary artery disease	11	53 (28)	13 (7)	1 (1)	120 (64)	187
Postpartum haemorrhage	12	2 (8)	12 (50)	0 (0)	10 (42)	24
Psoriasis	10	30 (30)	3 (3)	0 (0)	67 (67)	100
Breast cancer	26	79 (20)	71 (18)	0 (0)	240 (62)	390
Total		413	140	9	756	1318

Note: Partial matches are further categorised according to whether the HTA outcome was either more general or more specific than the core outcome.

COS, core outcomes set; COs, core outcomes.

rounded to whole numbers. Rheumatoid arthritis had the greatest percentage of specific matches (52%), postpartum haemorrhage returned the greatest percentage of partial matches (50%) while epilepsy and psoriasis both had the greatest percentage of no matches (67%). Only rheumatoid arthritis, type 1 diabetes, acne and postpartum haemorrhage had more total core outcome matches compared to no matches. The number of HTAs assessed ranged from one for acne to 22 for rheumatoid arthritis. The COS for rheumatoid arthritis had the greatest median value of matches with 69% while epilepsy had the fewest with a median of 25%, detailed in [Table 4](#).

To consider how the scope of the HTA compared to that of the COS, each HTA was assessed with regards to the investigated population and intervention. [Table 5](#) shows the majority of the studies clustered around exact scope matches and the COS being broader, both in terms of the population and intervention.

The distribution of HTAs by year ranged from 1998 to 2020, with 69% of included HTAs published in 2012 and beyond. The most HTAs published in a single year was 13 in 2012. A mean of nine HTAs were published per year in the following years between 2013 and 2020. [Figure 2](#) shows from 2009 onward there has been some fluctuation in the degree of match, though ultimately the trend is stationary, remaining close to the mean of 43%.

A drawback of these results is that many HTAs were produced prior to publication of the COS. To establish a clearer view on the impact of COS, a subgroup analysis consisting of HTAs developed following publication of the COS was

**Table 4. Number of HTA outcomes with partial or specific match to core outcomes.**

COS	Number of COs	Number of specific and partial matches to COs for each individual HTA (% of total)	Median %	% Range
RA	8	8 (100), 8 (100), 5 (63), 5 (63), 0 (0), 8 (100), 8 (100), 8 (100), 0 (0), 8 (100), 4 (50), 3 (38), 7 (88), 2 (25), 2 (25), 6 (75), 8 (100), 8 (100), 3 (38), 7 (88), 0 (0), 0 (0)	69	0, 100
MS/CIS	9	2 (22), 3 (33), 2 (22), 4 (44), 4 (44), 8 (89), 4 (44), 4 (44)	44	22, 89
Obesity	9	6 (67), 5 (56), 2 (22), 4 (44), 6 (67), 2 (22), 1 (11), 4 (44), 5 (56), 3 (33)	44	11, 67
Epilepsy	8	2 (25), 2 (25), 2 (25), 2 (25), 2 (25), 3 (38), 2 (25), 2 (25), 1 (13), 1 (13), 5 (63), 3 (38), 3 (38), 5 (63), 5 (63)	25	13, 65
T1DM	8	5 (63), 2 (25), 1 (13), 3 (38), 4 (50), 8 (100), 6 (75), 3 (38), 5 (63), 6 (75), 4 (50), 3 (38), 3 (38), 4 (50), 5 (63), 6 (75), 4 (50), 2 (25), 3 (38)	50	13, 100
Acne	7	4 (57)	57	n/a
CAD	11	4 (36), 7 (64), 6 (55), 9 (82), 6 (55), 0 (0), 0 (0), 5 (45), 7 (64), 4 (36), 0 (0), 4 (36), 5 (45), 0 (0), 1 (9), 4 (36), 5 (45)	36	0, 82
PPH	12	7 (58), 7 (58)	58	58, 58
Psoriasis	10	4 (40), 4 (40), 5 (50), 3 (30), 4 (40), 1 (10), 3 (30), 2 (20), 2 (20), 5 (50)	35	10, 50
Breast cancer	26	10 (38), 10 (38), 15 (58), 1 (4), 15 (58), 1 (4), 15 (58), 15 (58), 15 (58), 15 (58), 15 (58), 6 (23), 7 (27), 5 (19), 5 (19)	38	4, 58

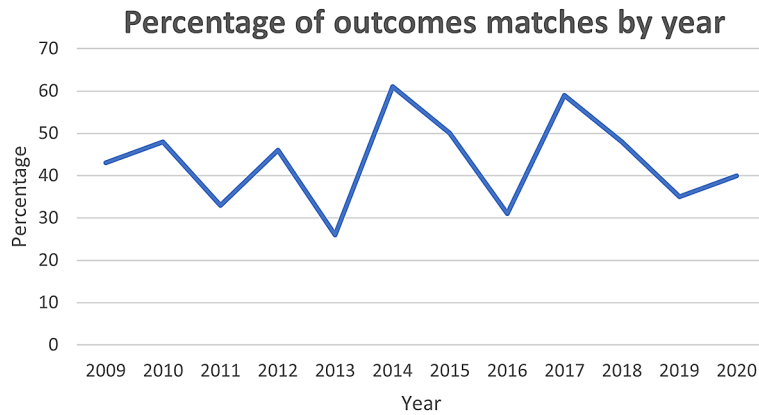
CAD, coronary artery disease; COS, core outcome set; COs, core outcomes; HTA, health technology assessment; MS/CIS, multiple sclerosis/clinically isolated syndrome; PPH, postpartum haemorrhage; RA, rheumatoid arthritis; T1DM, type 1 diabetes mellitus.

**Table 5. The scope of HTAs compared to the COS regards population and intervention.**

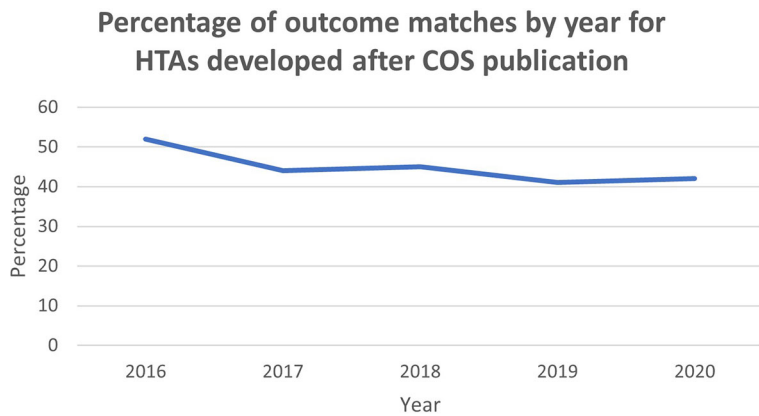
		Intervention			
		COS is narrower	Exact match	COS is broader	Different intervention
Population	COS is narrower	2	10	11	0
	Exact match	1	22	23	0
	COS is broader	2	19	29	0
	Different subgroup of the population	0	0	0	0

COS, core outcome set.





**Figure 2. Percentage of outcome matches by year.**



**Figure 3. Percentage of outcome matches by year for HTAs developed after COS publication.**

undertaken. This comprised 25 HTAs across 8 of the selected COS. From a possible 373 there were 163 outcome matches, encompassing 109 specific matches and 54 partial matches. Accordingly 44% of HTA outcomes were covered by the COS, ranging from 52% in 2016 to 41% in 2019, with a trend of marginal decrease illustrated in Figure 3. This may reflect the fact that, although the clinical trials community are becoming more aware of COS, they are yet to be adopted by the HTA community.

To assess the similarity between COS and recently published NICE HTA outcomes, the NICE website was searched for HTAs corresponding to each of the COS from the original cohort, except breast cancer. NICE HTAs were only identified for four COS, so six alternative COS were selected to include in the analysis. Ten non-oncology HTAs published between 2019 and 2021 were identified from the NICE website and compared against a relevant COS, shown in Table 6. HTA matches ranged from 44% to 100%, with a total of 78% (73/94) outcome matches with NICE HTAs, including 57 (61%) specific matches and 16 (17%) partial matches.

**Discussion**

This review found that HTAs in the INAHTA database included 43% of core outcomes from COS, with specific matches accounting for 31% and partial matches for 11%.

**Similarity between COS and HTA outcomes**

This is a novel piece of work as there are few studies conducted assessing COS uptake, and none investigating their similarity with non-oncology HTA outcomes. Recent work has found that the uptake of COS in randomised controlled trials and systematic reviews varies greatly between different areas of health.<sup>27</sup> Barriers to COS uptake are also noted, with lack of awareness, lack of validated measures or no consensus on measures, and lack of patient involvement being the most common reasons reported.<sup>27</sup> Ideally with time, awareness will continue to grow, leading to the development of more high-quality COS which involve all key stakeholders including patients. The issue regarding a lack of validated



**Table 6. Recently published non-oncology NICE HTAs compared with relevant COS, categorised as specific, partial or no match.**

COS	Number of COs	Specific match (% of total)	Partial match (% of total)		No match (% of total)	HTA matches (% of total)
			More general	More specific		
Type 1 diabetes	8	4 (50)	1 (13)	0 (0)	3 (38)	5 (63)
Type 2 diabetes	18	16 (89)	1 (6)	0 (0)	1 (6)	17 (94)
Rheumatoid arthritis	8	8 (100)	0 (0)	0 (0)	0 (0)	8 (100)
Age-related macular degeneration	7	1 (14)	3 (43)	2 (29)	1 (14)	6 (86)
Multiple sclerosis (Relapsing remitting)	9	3 (33)	0 (0)	1 (11)	5 (56)	4 (44)
Multiple sclerosis (Secondary progressive)	8	4 (50)	0 (0)	0 (0)	4 (50)	4 (50)
Dravet syndrome	5	1 (20)	1 (20)	1 (20)	2 (40)	3 (60)
Psoriasis	10	5 (50)	0 (0)	0 (0)	5 (50)	5 (50)
Coronary artery disease	13	13 (100)	0 (0)	0 (0)	0 (0)	13 (100)
Crohn's disease	8	2 (25)	1 (13)	5 (63)	0 (0)	8 (100)
Total	94	57	7	9	21	73

Note: Partial matches are further categorised according to whether the NICE outcome was either more general or more specific than the core outcome.

COS, core outcome set; COs, core outcomes; HTA, health technology assessment.

measures, however, is a pertinent problem and one which goes hand in hand with the fundamentals of COS. Research to determine the most suitable outcomes to measure must be accompanied by work to determine how best to measure those outcomes. This is particularly important for life impact domains which tend to be patient reported. There are many well established methods to measure outcomes from clinical, resource use and adverse event domains as these are often linked to everyday work of professionals in healthcare, unlike life impact outcomes, the importance of which are becoming increasingly recognised as a result of COS.

A potential reason for reduced similarity between COS and HTA outcomes in this review may be the target populations selected when originally designing the COS. Eight of the ten COS related to any intervention for their given disease. The result is that studies investigating specific interventions, such as bioabsorbable stents for coronary artery disease, are compared against measures designed for all coronary artery disease patients. A study investigating bioabsorbable stents may not have had the resources available to follow up their participants after five years to survey for outcomes such as stroke, depression or functional status, as this would have large time implications and ultimately may be of little relevance to their research question. [Table 5](#) illustrates this point, with the distribution of HTAs around that of "COS is broader" for both population and intervention. Likewise, as many of the COS related to any individual with the disease, studies concerning specific populations would have fallen under their scope. While this could suggest that more COS are needed to cover the different subgroups, the value of COS are the uniformity they provide. If trials concerning subpopulations, such as pregnancy, advanced disease, or hormone responsive disease, recorded the outcomes designed for the greater population, it could reveal currently unknown trends, such as interventions experiencing improved efficacy in specific groups.

There are instances where outcomes identified in the scope of an HTA cannot be included in the final assessment as sufficient available data is lacking. This may reflect appreciation and acknowledgement of the importance of these outcomes by those appraising the value of a technology which needs to be filtered down to those designing and conducting trials. Trial funders and researchers may need to consider the time taken to design, conduct and report research to build a sufficiently detailed literature base. This would suggest value in allowing an extended period of time before further research to allow publication of studies which have incorporated COS into their protocols.

Another potential explanation for low representation of core outcomes could relate to the selection of COS published between 2015 and 2019. This was used as a proxy measure for quality COS, but the majority of the HTAs assessed (71/119) were published before 2015. While this could account for the lack of overlap prior to this year, [Figure 2](#) demonstrates there has been no sustained increase in the percentage of outcomes included in HTAs.

As a 10% subset of the most recent breast cancer HTAs were included in the analysis, it is intriguing to note it is amongst the lowest matching conditions. While reports published from 2016 onwards were solely analysed, it should be noted that the COS included 26 separate core outcomes. This was markedly greater compared to other COS in the analysis and may point to the importance of being selective when developing COS, as trial designers may find vast numbers of outcomes too overwhelming to take them all into consideration. Including too many outcomes may also serve to dilute the relevance of other more significant outcomes.

Recent work has compared ICHOM standard sets to oncology HTAs using a qualitative approach. They found that HTAs favour more generic outcome measures which allow comparison with other disease areas, rather than disease-specific outcomes recommended by the ICHOM standard sets.<sup>12</sup> It has been shown that HTAs consider overall survival data to be most crucial when making decisions on the value of the technology.<sup>28</sup> However, overall survival data are not always mature when submitted for HTA assessment and so intermediate outcomes, such as progression-free survival, may be suitable for HTA agencies. HTA agencies only accept validated intermediate outcomes and the level of validity varies between HTA agencies.<sup>12,28</sup> However, it should be noted the study only focussed on five areas of oncology and that ICHOM standard sets are developed for routine care, whereas COS designed for research are therefore of more relevance for HTA agencies.

#### INAHTA database

Despite the accessibility difficulties, the INAHTA database remains a great resource for identifying HTAs. It can be easily navigated and continues to update details of new HTAs whenever they become available. In response to feedback on the difficulties accessing the HTAs, the managers of the database recognised the hyperlink issues, which arose as the records are self-administered and the producer of the report adds their own copy, noting that the correction of this issue will take time.

A recent study surveying members of the INAHTA found that one of their biggest challenges is including stakeholders in HTAs, particularly where short deadlines offer little time for input.<sup>29</sup> Implementing COS that have been developed with patient input could ensure HTAs included patient representation and reduce the time-consuming process of surveying patients themselves.

#### NICE HTAs

The finding that 78% of core outcomes were covered by NICE HTAs is very encouraging and notably greater than that for HTAs from the INAHTA database. This may be linked to the endorsement of COS by NICE or their rigorous approach of protocol design and stakeholder involvement.

#### Limitations

The main limitation in this review was the large number of hyperlinks (474/651, 73%) in the INAHTA database which did not lead to a report, resulting in a large number of exclusions. This allowed for feedback to the database providers, beginning the process of amending the accessibility of the reports so future research can make use of the resource. On occasion the hyperlink led to the source website from which the HTA could be searched for, but this did not always produce the desired HTA. Additionally, only one database was searched for HTAs to include in the review potentially omitting appropriate HTAs not registered on the INAHTA database. Another limitation is that many of the COS included in the review related to any intervention for their given disease. This may have exaggerated reduced incorporation of core outcomes in the context of very specific HTAs.

Although the HTA outcome data extraction was conducted by a single researcher (PC), the opinion of a second researcher (SD) with extensive experience of outcome extraction and classification was consulted in all cases of doubt or ambiguity. In addition, all outcome matching assessments were double checked by this reviewer (SD).

#### Conclusion

This novel piece of research includes, to our knowledge, the first comparison of non-oncology HTAs and COS for research. This pilot study highlights that further work is needed to promote awareness and implementation of COS within HTAs, with less than half of core outcomes currently measured. The degree of matching between COS and NICE HTA outcomes, albeit from a small sample, suggests this is both practical and achievable with the additional benefit of

increasing patient representation, a particular challenge for HTA agencies. Incorporation of COS will allow HTA findings to better reflect the opinions and preferences of all relevant stakeholders.

## Data availability statement

### Underlying data

The [COMET database](#) was utilised to identify suitable COS. Each of the 10 COS are accessible through the following hyperlinks: [rheumatoid arthritis](#), [relapsing remitting multiple sclerosis/clinically isolated syndrome](#), [obesity](#), [epilepsy](#), [type 1 diabetes](#), [acne](#), [coronary artery disease](#), [postpartum haemorrhage](#), [psoriasis](#), [breast cancer](#). HTAs can be accessed through the [INAHTA database](#). The search terms are described in [Table 1](#).

The 10 COS compared to the NICE HTAs can be accessed through the following hyperlinks: [type 1 diabetes](#), [type 2 diabetes](#), [rheumatoid arthritis](#), [age-related macular degeneration](#), [relapsing remitting multiple sclerosis/clinically isolated syndrome](#), [secondary progressive multiple sclerosis](#), [Dravet syndrome](#), [psoriasis](#), [coronary artery disease](#), [Crohn's disease](#).

The corresponding NICE HTAs can be accessed through their [website](#) or through the following hyperlinks: [type 1 diabetes](#), [type 2 diabetes](#), [rheumatoid arthritis](#), [age-related macular degeneration](#), [relapsing remitting multiple sclerosis/clinically isolated syndrome](#), [secondary progressive multiple sclerosis](#), [Dravet syndrome](#), [psoriasis](#), [coronary artery disease](#), [Crohn's disease](#).

### Extended data

Harvard Dataverse: Data analysis for “A pilot study assessing the uptake of core outcome sets in health technology assessments” <https://doi.org/10.7910/DVN/62NTAA>.

Data are available under the terms of the [Creative Commons Zero “No rights reserved” data waiver](#) (CC0 1.0 Public domain dedication).

## References

- Gartlehner G, Hansen RA, Nissman D, *et al.*: **Criteria for distinguishing effectiveness from efficacy trials in systematic reviews**. 2006.
- Meinert CL: *ClinicalTrials: design, conduct and analysis*. USA: OUP; 2012.
- Hirsch BR, Califf RM, Cheng SK, *et al.*: **Characteristics of oncology clinical trials: insights from a systematic analysis of ClinicalTrials.gov**. *JAMA Intern. Med.* 2013; **173**(11): 972–9. [PubMed Abstract](#) | [Publisher Full Text](#)
- Tovey D: **The impact of cochrane reviews: The Cochrane Collaboration**. 2010. [Publisher Full Text](#).
- Miyar J, Adams CE: **Content and quality of 10 000 controlled trials in schizophrenia over 60 years**. *Schizophr. Bull.* 2013; **39**(1): 226–9. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Chan A-W, Song F, Vickers A, *et al.*: **Increasing value and reducing waste: addressing inaccessible research**. *Lancet.* 2014; **383**(9913): 257–66. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Webbe J, Sinha I, Gale C: **Core outcome sets**. *Archives of Disease in Childhood-Education and Practice*. 2018; **103**(3): 163–6.
- Williamson PR, Altman DG, Blazeby JM, *et al.*: **Developing core outcome sets for clinical trials: issues to consider**. *Trials.* 2012; **13**(1): 1–8. [Publisher Full Text](#).
- Kirkham JJ, Davis K, Altman DG, *et al.*: **Core outcome Set-STAndards for development: the COS-STAD recommendations**. *PLoS Med.* 2017; **14**(11): e1002447. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dodd S, Fish R, Gorst S, *et al.*: **Representation of published core outcome sets for research in regulatory guidance: protocol**. *HRB Open Research.* 2021; **4**(45): 45. [Publisher Full Text](#)
- O'Rourke B, Oortwijn W, Schuller T: **The new definition of health technology assessment: A milestone in international collaboration**. *Int. J. Technol. Assess. Health Care.* 2020; **36**(3): 187–90. [PubMed Abstract](#) | [Publisher Full Text](#)
- Kalf RR, Vreman RA, Delnoij DM, *et al.*: **Bridging the gap: Can International Consortium of Health Outcomes Measurement standard sets align outcomes accepted for regulatory and health technology assessment decision-making of oncology medicines**. *Pharmacol. Res. Perspect.* 2021; **9**(2): e00742. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- National Institute of Health and Care Excellence: **Health-related quality of life task and finish group report. Report 4: Core outcome sets 2020**. [Reference Source](#)
- Hughes KL, Kirkham JJ, Clarke M, *et al.*: **Assessing the impact of a research funder's recommendation to consider core outcome sets**. *PLoS One.* 2019; **14**(9): e0222418. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Dodd S, Harman N, Taske N, *et al.*: **Core outcome sets through the healthcare ecosystem: the case of type 2 diabetes mellitus**. *Trials.* 2020; **21**(1): 1–7. [Publisher Full Text](#)
- Saldanha JJ, Dodd S, Gorst SL, *et al.*: **More than half of systematic reviews have relevant core outcome sets**. *J. Clin. Epidemiol.* 2021; **136**: 168–79. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Radner H, Chatzidionysiou K, Nikiphorou E, *et al.*: **2017 EULAR recommendations for a core data set to support observational research and clinical care in rheumatoid arthritis**. *Ann. Rheum. Dis.* 2018; **77**(4): 476–9. [PubMed Abstract](#) | [Publisher Full Text](#)
- Marrie RA, Miller A, Sormani MP, *et al.*: **Recommendations for observational studies of comorbidity in multiple sclerosis**. *Neurology.* 2016; **86**(15): 1446–53. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

19. Coulman KD, Hopkins J, Brookes ST, *et al.*: **A core outcome set for the benefits and adverse events of bariatric and metabolic surgery: the BARIACT project.** *PLoS Med.* 2016; **13**(11): e1002187.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
20. Noble A, Marson A: **Which outcomes should we measure in adult epilepsy trials? The views of people with epilepsy and informal carers.** *Epilepsy Behav.* 2016; **59**: 105–10.  
[PubMed Abstract](#) | [Publisher Full Text](#)
21. 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**(1): 1–8.  
[Publisher Full Text](#)
22. Layton AM, Eady EA, Thiboutot DM, *et al.*: **Identifying what to measure in acne clinical trials: first steps towards development of a core outcome set.** *J. Invest. Dermatol.* 2017; **137**(8): 1784–6.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
23. 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.  
[Publisher Full Text](#)
24. 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. Gynaecol.* 2019; **126**(1): 83–93.  
[PubMed Abstract](#) | [Publisher Full Text](#)
25. Duffin KC, Merola JF, Christensen R, *et al.*: **Identifying a core domain set to assess psoriasis in clinical trials.** *JAMA Dermatol.* 2018; **154**(10): 1137–44.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
26. Ong WL, Schouwenburg MG, Van Bommel AC, *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.  
[PubMed Abstract](#) | [Publisher Full Text](#)
27. Hughes KL, Clarke M, Williamson PR: **A systematic review finds core outcome set uptake varies widely across different areas of health.** *J. Clin. Epidemiol.* 2020; **129**: 114–23.  
[Publisher Full Text](#)
28. Kleijnen S, Lipska I, Alves TL, *et al.*: **Relative effectiveness assessments of oncology medicines for pricing and reimbursement decisions in European countries.** *Ann. Oncol.* 2016; **27**(9): 1768–75.  
[PubMed Abstract](#) | [Publisher Full Text](#)
29. O'Rourke B, Werkö SS, Merlin T, *et al.*: **The 'top 10' challenges for health technology assessment: INAHTA viewpoint.** *Int. J. Technol. Assess. Health Care.* 2020; **36**(1): 1–4.  
[PubMed Abstract](#) | [Publisher Full Text](#)

# Open Peer Review

Current Peer Review Status:  

---

## Version 3

Reviewer Report 04 May 2022

<https://doi.org/10.5256/f1000research.133578.r136562>

© 2022 Turnbull A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Alison E. Turnbull** 

<sup>1</sup> Division of Pulmonary and Critical Care Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

<sup>2</sup> Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

The authors responded satisfactorily to my suggested revisions. I have no further comments.

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Clinical epidemiology, critical care outcomes, patient-centered care

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

## Version 2

Reviewer Report 19 April 2022

<https://doi.org/10.5256/f1000research.121621.r129869>

© 2022 Turnbull A. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Alison E. Turnbull** 

<sup>1</sup> Division of Pulmonary and Critical Care Medicine, School of Medicine, Johns Hopkins University, Baltimore, MD, USA

<sup>2</sup> Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD, USA

This manuscript describes the similarity between outcomes used to evaluate health technologies and the outcomes included in Core Outcomes Sets (COS). Attempts to evaluate the effects of health technologies - termed health technology assessments (HTAs) - were identified through the International Network of Agencies for Health Technology Assessments (INHATA) and the National Institute for Health and Care Excellence (NICE) website. 90% of identified HTAs were systematic reviews and the remainder were trials. Ten COS that were developed for common diseases in collaboration with patients, and published between 2015 and 2019 were identified from the COMET database. Basic descriptive statistics were used to describe the similarity between outcomes in HTAs and each COS, to compare the populations and interventions targeted by COS and HTAs, and to describe changes in the level of similarity between COS and HTAs over time. The study's primary limitation, which the authors acknowledge, is that the vast majority of HTAs eligible for inclusion in the study did not have functional hyperlinks in INHATA, and it is unclear to what extent the included studies are representative of the entire target population of HTAs. Put another way, the external validity of the HTAs described is unknown.

My primary recommendation regards how the study's results have been interpreted. Conclusions in both the abstract and body of the manuscript address awareness, acceptance, and implementation of COS by the people designing HTAs. However, the temporal relationship between the publication of COS and the design of HTAs included in this study is not straightforward. Fully 60% of included HTAs were published before 2015 (and possibly designed years before publication), while the 10 COS were published between in 2015 and later. In Figure 2, there does not appear to be a temporal trend between 2009 and 2020 in the degree of similarity between COS and HTA outcomes. Thus, another interpretation of these results is that COS developers consistently include 40 - 70% of the outcomes assessed by HTAs when developing their COS and this does not appear to be changing. I recommend performing a sub-group analysis of HTAs developed after the publication of relevant COS to address this chicken-and-egg problem. The results of such an analysis may permit stronger and more substantiated inferences about the awareness, acceptance, and implementation of COS.

**Is the work clearly and accurately presented and does it cite the current literature?**

Yes

**Is the study design appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**

Yes

**If applicable, is the statistical analysis and its interpretation appropriate?**

Yes

**Are all the source data underlying the results available to ensure full reproducibility?**

Yes

**Are the conclusions drawn adequately supported by the results?**

Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Clinical epidemiology, critical care outcomes, patient-centered care

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.**

Author Response 27 Apr 2022

**Peter Cox**, University of Liverpool, Liverpool, UK

Dear Dr Turnbull

Thank you for reviewing our manuscript "A pilot study assessing the similarity between core outcome sets and outcomes included in health technology assessments". Thank you for your recommendation of a subgroup analysis, which we have undertaken. We have included the following paragraph to the results section alongside a new figure describing the percentage of outcome matches by year for HTAs developed after COS publication.

"A drawback of these results is that many HTAs were produced prior to publication of the COS. To establish a clearer view on the impact of COS, a subgroup analysis consisting of HTAs developed following publication of the COS was undertaken. This comprised 25 HTAs across 8 of the selected COS. From a possible 373 there were 163 outcome matches, encompassing 109 specific matches and 54 partial matches. Accordingly 44% of HTA outcomes were covered by the COS, ranging from 52% in 2016 to 41% in 2019, with a trend of marginal decrease illustrated in Figure 3. This may reflect the fact that, although the clinical trials community are becoming more aware of COS, they are yet to be adopted by the HTA community."

**Competing Interests:** Paula Williamson chairs the COMET Initiative Management Group.

Reviewer Report 04 March 2022

<https://doi.org/10.5256/f1000research.121621.r125645>

© 2022 McDaid C. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Catriona McDaid** 

York Trials Unit, Department of Health Sciences, University of York, York, UK

We thank the authors for their detailed response to our comments and associated amendments.



We have no further comments.

**Is the work clearly and accurately presented and does it cite the current literature?**

Partly

**Is the study design appropriate and is the work technically sound?**

Partly

**Are sufficient details of methods and analysis provided to allow replication by others?**

Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**

Partly

**Are all the source data underlying the results available to ensure full reproducibility?**

Partly

**Are the conclusions drawn adequately supported by the results?**

Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Applied health research, systematic reviews, randomised controlled trials, health technology assessment

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

---

Version 1

Reviewer Report 18 January 2022

<https://doi.org/10.5256/f1000research.77311.r118406>

© 2022 McDaid C et al. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Catriona McDaid**

York Trials Unit, Department of Health Sciences, University of York, York, UK

**Andrew Mott**

University of York, UK, York, UK

This peer review has been undertaken jointly with Andrew Mott, NIHR Pre-doctoral Research Fellow at the University of York. We each reviewed independently and then combined our reviews.

The article explores the extent of agreement/match between 10 individual COS published between 2015 and 2019 and HTA reports published on the INHATA database up to February 2021, supplemented by a comparison between COS and a cohort of UK NICE HTA reports. Overall, the study methods and results are comprehensively reported, tables are informative, and the data underlying the results are available. There is no reference to a publicly available protocol, though we have not identified any obvious sources of potential bias. Key limitations are addressed. There are a few areas where the reporting could be made clearer and some aspects of interpretation given further consideration.

- The study is described as a pilot in the title and the conclusions. In the discussion there is also reference to “the feasibility assessment in this pilot study”. However, this is not apparent from the study aim/purpose or the methods so we suggest the term pilot is not used unless further information can be presented to show that it was indeed intended as a pilot – in which case the objective and methods would need amended to reflect this.
- We were not totally convinced by the rationale for excluding oncology studies from the NICE cohort but not the INHATA cohort, when the stated purpose of including the former was to fill the gap identified in what was recorded on the INHATA database. It also looks as though a different process was used – as we interpret it, for the INHATA part of the study the COS were identified first then the reports and vice versa for the NICE cohort. We don't think this will have introduced any important bias though possibly you could end up with the UK reports covering a different time period. You could consider adding some further descriptive information on the NICE cohort on range of publication year as you have for the INHATA cohort so that the similarity of the two datasets is clear.
- Selecting a subset of 10% of the breast cancer HTAs seems reasonable given the volume. Given that the most recent were selected it would be worth picking up in the discussion what effect this might have on the results as these are more likely to have been undertaken after the COS were published than any of the other conditions. This might be of particular interest as it appears to be amongst the lower matching (median 38%) conditions.
- The methods state that the HTAs were taken from the INAHTA database. This database includes a variety of study designs and the term HTA can be defined in different ways. As reviewers we both interpreted differently whether or not trials were included so it would be useful to be explicit about what study designs were included under the term HTA. If all study designs were included some discussion around how the different designs may have affected the use of COS should be added.
- If available, the date of searching the NICE website should be added as the content is regularly changing.
- The methods state that COS were selected from the COMET website based on year of publication (2015-2019) and whether they addressed a common disease or intervention. Is there any further information on how the decision about being common was applied? A study selection flow diagram would also be a helpful addition to the text in the first paragraph of the results.
- Further information in the methods on how outcomes were identified/established/extracted

from reports would be useful. There may not be full agreement between the outcomes of interest identified in the scope/methods of a technology assessment (defining this as evidence synthesis and cost-effectiveness analysis) and the outcomes that are eventually included in the assessment due to the lack of data on the specific outcome; for example an outcome may be defined as of interest in the scope but no relevant data found and included in the synthesis. Also in the scope the outcomes may be more broadly expressed but in the included studies outcomes may be more specific reflecting what they could find. How was this approached in the data extraction? This may seem pedantic/theoretical but is there a possibility that the scope/methods sections could match well to a COS if it is a recent report, but the outcomes reported in the actual included studies may not if the data available are mainly older trials. So, there is another time lag issue here. Some further unpacking of this may be worth considering and whether there are any related recommendations to be made regarding future research on this topic. To ensure study replicability some additional information in the methods on how data were extracted would be helpful.

- As the COSs were chosen initially and the HTAs identified that fit the COS it is possible these are not the most appropriate COS for each individual HTA. This is partially demonstrated in table 5 where almost half the HTAs have a narrower scope than the COS. It would be beneficial to have some discussion about whether these HTAs outcomes are specific to the narrower population or intervention.
- We agree with the authors that a limitation of the study is that the majority of the included reports were published prior the introduction of the COS. In some parts of the paper e.g., the title and some parts of the discussion, we are not sure that the terminology 'COS uptake' is appropriate given that the COS was not necessarily available at the time a report was published. The terminology used for the study purpose and elsewhere "similarity between COS and outcomes included in HTAs" is more accurate and appropriate and this could be applied across the paper.
- The conclusion refers to "improved uptake across NICE HTAs" – we are not convinced that the data presented warrant such a strong conclusion.
- Underlying data – For the INHATA database only the general link to the site is available – is it possible to provide a list of the included reports?

**Is the work clearly and accurately presented and does it cite the current literature?**

Yes

**Is the study design appropriate and is the work technically sound?**

Yes

**Are sufficient details of methods and analysis provided to allow replication by others?**

Partly

**If applicable, is the statistical analysis and its interpretation appropriate?**

Yes

**Are all the source data underlying the results available to ensure full reproducibility?**

Yes

**Are the conclusions drawn adequately supported by the results?**

Partly

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Applied health research, systematic reviews, randomised controlled trials, health technology assessment

**We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.**

Author Response 22 Feb 2022

**Peter Cox**, University of Liverpool, Liverpool, UK

Dear Dr McDaid and Mr Mott,

Thank you for reviewing our manuscript "A pilot study assessing the uptake of core outcome sets in health technology assessments". We hope that the corrections we have made capture your points.

Yours sincerely,

Peter Cox

**The study is described as a pilot in the title and the conclusions. In the discussion there is also reference to "the feasibility assessment in this pilot study". However, this is not apparent from the study aim/purpose or the methods so we suggest the term pilot is not used unless further information can be presented to show that it was indeed intended as a pilot – in which case the objective and methods would need amended to reflect this.**

**Author response:** *Thank you for raising this point. This study was conceptualised as a small scale study investigating the similarity between COS and HTA outcomes, given the presence of few prior studies assessing COS uptake. The results of this study provide backing for a similar study on a larger scale, to assess if these trends identified hold true given a larger sample size. This led to the description as a pilot study.*

*The following line regards the feasibility assessment has been removed, "This was an important finding in terms of the feasibility assessment in this pilot study".*

**We were not totally convinced by the rationale for excluding oncology studies from the NICE cohort but not the INHATA cohort, when the stated purpose of including the former was to fill the gap identified in what was recorded on the INHATA database. It**

also looks as though a different process was used – as we interpret it, for the INHATA part of the study the COS were identified first then the reports and vice versa for the NICE cohort. We don't think this will have introduced any important bias though possibly you could end up with the UK reports covering a different time period. You could consider adding some further descriptive information on the NICE cohort on range of publication year as you have for the INHATA cohort so that the similarity of the two datasets is clear.

**Author response:** *Thank you for this comment. NICE HTAs are of particular importance in the UK and following the discovery recent HTAs were not registered on the INAHTA database beyond 2011, they were searched for and identified from the NICE website to supplement our findings. Given COS vs oncology NICE HTAs has already been assessed in a previous project [Kalf et al.], non-oncology COS were assessed. From the original cohort of 10 COS, relevant NICE HTAs were only identified for 4 COS. Subsequently a further 6 non-oncology COS were selected to be assessed against NICE HTAs, providing the results in table 6.*

*The methods text has been edited to include the following:*

*“Within the UK, NICE HTAs are of particular importance. It was noted that no NICE HTAs had been registered on the INAHTA database beyond 2011, so in April 2021, the NICE website was searched for HTAs corresponding to each of the COS from the original cohort. Non-oncology conditions were selected given previous work has examined uptake across oncology HTAs (12, 13) and where no appropriate NICE HTAs were identified, alternative COS were selected, using the same criteria of publication between 2015 and 2019 with patient involvement. The most recently published reports which matched a COS were included, with outcomes compared to core outcomes from the corresponding COS.*

*“The results are presented through tables, graphs and using descriptive statistics. The analysis was performed using Excel 2019. Comparison between the INAHTA and NICE results was not examined given they cover different time periods and utilised different COS.”*

*The results section has been edited to include the following:*

*“To assess the similarity between COS and recently published NICE HTA outcomes, the NICE website was searched for HTAs corresponding to each of the COS from the original cohort, except breast cancer. NICE HTAs were only identified for four COS, so six alternative COS were selected to include in the analysis. Ten non-oncology HTAs published between 2019 and 2021 were identified from the NICE website and compared against a relevant COS, shown in Table 6.”*

**Selecting a subset of 10% of the breast cancer HTAs seems reasonable given the volume. Given that the most recent were selected it would be worth picking up in the discussion what effect this might have on the results as these are more likely to have been undertaken after the COS were published than any of the other conditions. This might be of particular interest as it appears to be amongst the lower matching (median 38%) conditions.**

**Author response:** *Thank you for this consideration. The following has been added to the*

*discussion stating the dates of the HTAs included and to suggest why this COS had amongst the lowest outcome matching.*

*“As a 10% subset of the most recent breast cancer HTAs were included in the analysis, it is intriguing to note it is amongst the lowest matching conditions. While reports published from 2016 onwards were solely analysed, it should be noted that the COS included 26 separate core outcomes. This was markedly greater compared to other COS in the analysis and may point to the importance of being selective when developing COS, as trial designers may find vast numbers of outcomes too overwhelming to take them all into consideration. Including too many outcomes may also serve to dilute the relevance of other more significant outcomes.”*

**The methods state that the HTAs were taken from the INAHTA database. This database includes a variety of study designs and the term HTA can be defined in different ways. As reviewers we both interpreted differently whether or not trials were included so it would be useful to be explicit about what study designs were included under the term HTA. If all study designs were included some discussion around how the different designs may have affected the use of COS should be added.**

**Author response:** *Thank you for this comment. The following has been added to the results section to clarify the different study designs included in the analysis.*

*“Systematic reviews provided over 90% of studies included, the remainder being trials. To minimise similarity to COS being underestimated by reviews, through outcome reporting being limited to what data exists in the literature, outcomes stated in the methods were included in the analysis, independent of the findings.”*

**If available, the date of searching the NICE website should be added as the content is regularly changing.**

**Author response:** *The NICE website was searched in April 2021, this has been included in the methods section.*

**The methods state that COS were selected from the COMET website based on year of publication (2015-2019) and whether they addressed a common disease or intervention. Is there any further information on how the decision about being common was applied? A study selection flow diagram would also be a helpful addition to the text in the first paragraph of the results.**

**Author response:** *A flow chart has been added to the beginning of the results and the methods section edited to include the following line:*

*“The selected COS were chosen as they were considered to be common conditions; examples of those excluded were: systemic sclerosis, malaria, thyroid eye disease, Hirschsprung's disease and twin-twin transfusion syndrome. This consideration heightened the likelihood sufficient HTAs were available for assessment”*

**Further information in the methods on how outcomes were identified/established/extracted from reports would be useful. There may not be full agreement between the outcomes of interest identified in the scope/methods of a technology assessment (defining this as evidence synthesis and cost-effectiveness analysis) and the outcomes that are eventually included in the assessment due to the lack of data on the specific outcome; for example an outcome may be defined as of interest in the scope but no relevant data found and included in the synthesis. Also in the scope the outcomes may be more broadly expressed but in the included studies outcomes may be more specific reflecting what they could find. How was this approached in the data extraction? This may seem pedantic/theoretical but is there a possibility that the scope/methods sections could match well to a COS if it is a recent report, but the outcomes reported in the actual included studies may not if the data available are mainly older trials. So, there is another time lag issue here. Some further unpacking of this may be worth considering and whether there are any related recommendations to be made regarding future research on this topic. To ensure study replicability some additional information in the methods on how data were extracted would be helpful.**

**Author response:** *Thank you for this consideration. Outcomes which were listed in the PICO or as a list of key research questions or equivalent were included and categorised as matches in our analysis. We have added the following line to the methods to clarify this and the paragraph to the discussion to give thought to the relevance of this:*

*“Outcomes identified in the scope of an HTA which could not be reported due to a lack of available data were included in the analysis.”*

*“There are instances where outcomes identified in the scope of an HTA cannot be included in the final assessment as sufficient available data is lacking. This may reflect appreciation and acknowledgement of the importance of these outcomes by those appraising the value of a technology which needs to be filtered down to those designing and conducting trials. Trial funders and researchers may need to consider the time taken to design, conduct and report research to build a sufficiently detailed literature base. This would suggest value in allowing an extended period of time before further research to allow publication of studies which have incorporated COS into their protocols.”*

**As the COS were chosen initially and the HTAs identified that fit the COS it is possible these are not the most appropriate COS for each individual HTA. This is partially demonstrated in table 5 where almost half the HTAs have a narrower scope than the COS. It would be beneficial to have some discussion about whether these HTAs outcomes are specific to the narrower population or intervention.**

**Author response:** *Thank you for this comment. The following has been included in the discussion addressing the scope of COS and relevance to narrower HTAs.*



*“Likewise [to COS being broader than many of the HTAs], as many of the COS related to any individual with the disease, studies concerning specific populations would have fallen under their scope. While this could suggest that more COS are needed to cover the different subgroups, the value of COS are the uniformity they provide. If trials concerning subpopulations, such as pregnancy, advanced disease, or hormone responsive disease, recorded the outcomes designed for the greater population, it could reveal currently unknown trends, such as interventions experiencing improved efficacy in specific groups.”*

**We agree with the authors that a limitation of the study is that the majority of the included reports were published prior the introduction of the COS. In some parts of the paper e.g., the title and some parts of the discussion, we are not sure that the terminology ‘COS uptake’ is appropriate given that the COS was not necessarily available at the time a report was published. The terminology used for the study purpose and elsewhere “similarity between COS and outcomes included in HTAs” is more accurate and appropriate and this could be applied across the paper.**

**Author response:** *Thank you for this comment and suggestion. To reflect this point COS uptake has been replaced throughout the text with “similarity between COS and HTA outcomes” when referring to our findings.*

*The title has been edited to “A pilot study assessing the similarity between core outcome sets and outcomes included in health technology assessments”*

**The conclusion refers to “improved uptake across NICE HTAs” – we are not convinced that the data presented warrant such a strong conclusion.**

**Author response:** *The conclusion has been edited to say “The degree of matching between COS and NICE HTA outcomes, albeit from a small sample, suggests this is both practical and achievable with the additional benefit of increasing patient representation, a particular challenge for HTA agencies”*

**Underlying data – For the INHATA database only the general link to the site is available – is it possible to provide a list of the included reports?**

**Author response:** *Links to individual HTA reports have been included in the data analysis on the Harvard Dataverse Repository under the extended data.*

**Competing Interests:** Paula Williamson chairs the COMET Initiative Management Group.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact [research@f1000.com](mailto:research@f1000.com)

**F1000Research**