OMERACT Instrument Selection

Topic: Synthesis of evidence

This document provides readers with a guide to various resources on synthesizing evidence using OMERACT Instrument Selection methodology.

A. Guidance on synthesis of evidence

A.1. Instrument selection overview whiteboard:

https://omeract.org/instrument-selection/ [see 7:45]

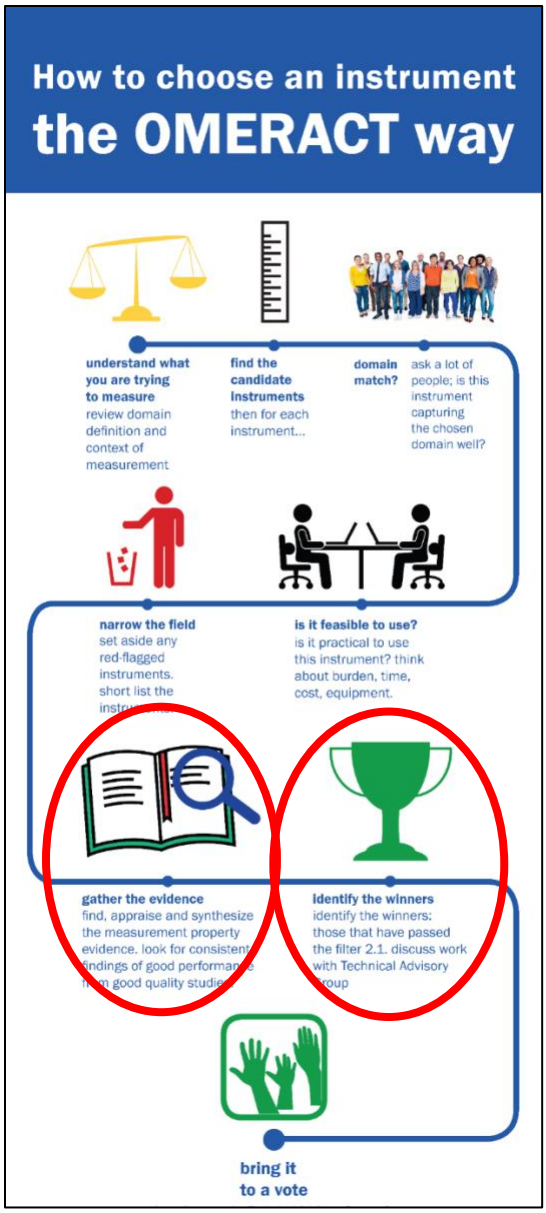
A.2. Synthesis of evidence video:

https://omeract.org/instrument-selection/

A.3. Instrument selection detailed discussion video:

https://omeract.org/instrument-selection/ [see 23:18]

B. OMERACT Way

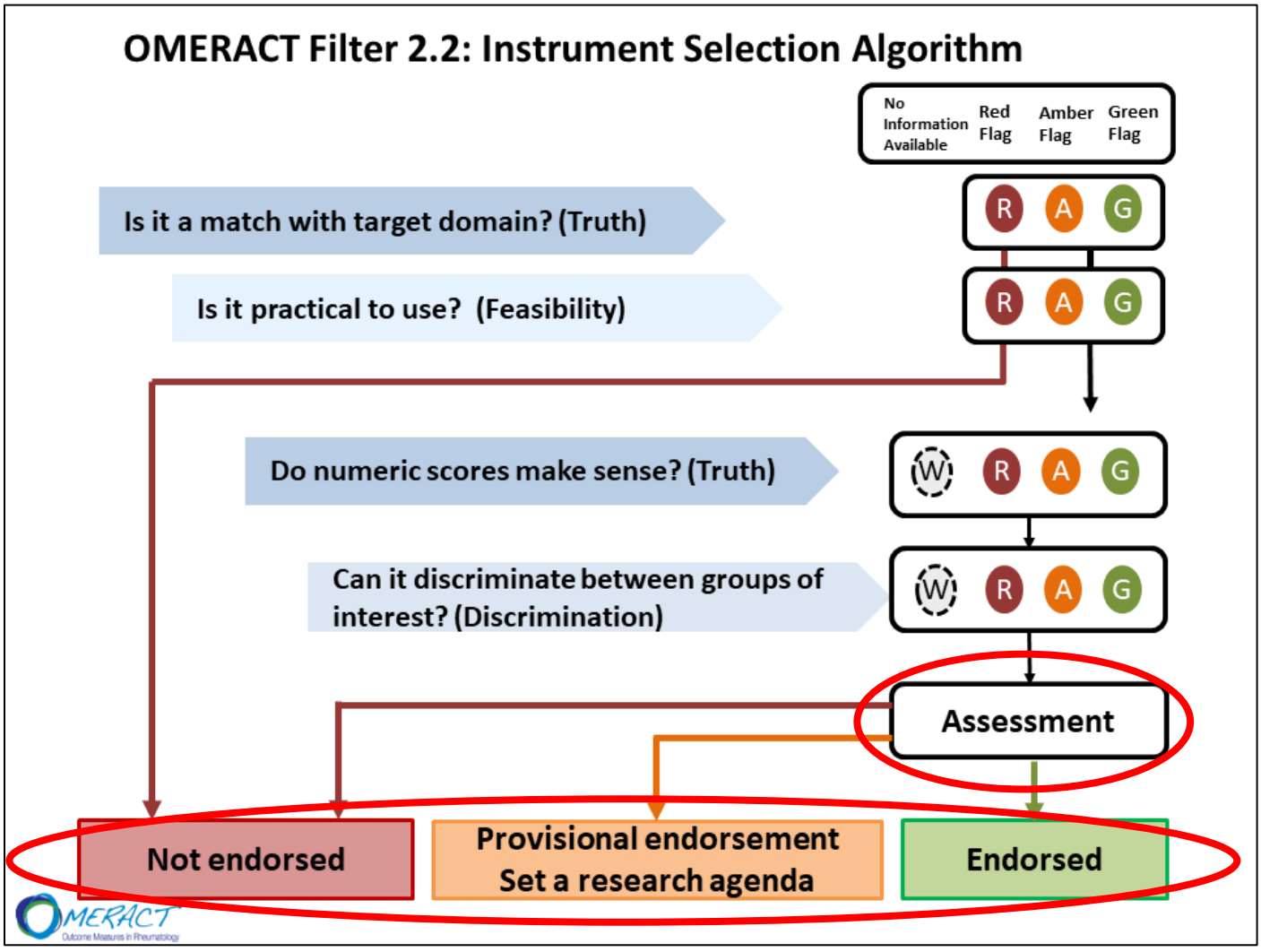


C. OMERACT Master checklist for instrument selection: Step 11 & 12

OMERACT Master Checklist for Instrument Selection ***Name of Instrument*:**		
Step #	**OMERACT Instrument Selection Process Checklist Item**	Mark when complete
Assembly of working group and protocol development		
1	Assemble working group	o
2	Decide on methods protocol for Core Outcome Instrument Set selection	o
3	**Deliverable**: Submit protocol using Instrument Selection Workbook to Technical Advisory Group [TAG]	o
4	Review and approval of final protocol by TAG	o
Review of evidence of instrument performance for existing or new instrument		
Part A: Domain match and Feasibility assessment		
5	Obtain Working Group and others assessment of match with the target domain	o
6	Obtain Working Group and others assessment of feasibility	o
7	Is the instrument a match with the domain AND feasible? Yes ____ → if yes, continue with Part B of checklist below No ____ → If no, set instrument aside (find new one or develop new one)	o
Part B: Review of evidence of performance of an instrument across key measurement properties		
8	Conduct literature search; create PRISMA diagram; place articles of measurement properties in Summary of Measurement Properties (SOMP) Table	o
9	Conduct COSMIN-OMERACT Good Methods check, add findings into the SOMP Table	o
10	Conduct data extraction, create summary reporting tables, fill in SOMP Table with assessment of adequacy of results	o
11	Conduct synthesis across evidence available for each measurement property	o
12	Decide if any gaps exist in evidence of measurement properties If gaps found, draft protocol for new study to fill gaps If no gaps, finish the SOMP Table with proposed level of endorsement	o
Initial submission to TAG: literature review findings & protocol for gaps		
13	**Deliverable**: Submit the Instrument Selection Workbook to TAG	o
14	Receive final response from TAG	o
15	If studies are needed to fill gaps, conduct new measurement property studies, submit to TAG for Good Methods check, add to body of evidence (SOMP) and go back to Step 12 If no studies are needed, put X here: _____and move to Step 16	o
Final submission to TAG for approval		
16	Obtain agreement on final report	o
17	Set timeline for next review of instrument	o
Ratification of level of endorsement by OMERACT Community and communication of results		
18	Ratification of level of endorsement by OMERACT Community	o
19	Implement communication and dissemination plan	o

D. OMERACT Filter 2.2. Instrument Selection Algorithm (OFISA)

The evidence from all the studies is synthesized for each measurement property. Then use the OMERACT Algorithm to determine proposed level of endorsement.



E. Where does synthesis fit on the Summary of Measurement Properties (SOMP) table?

The 'Synthesis rating' row is completed for each measurement property. Then the OMERACT Algorithm is used to determine the overall synthesis statement in the 'OMERACT Endorsement' row.

Instrument: ABC Domain: Physical function					Date completed: 2021-02-11			
Population: rheumatoid arthritis	Intervention(s): drug		Control: placebo/drug		Type of studies: clinical trials			
Author/year	Truth Domain match	Feasibility	Truth		Discrimination			
			Construct validity	Inter-method reliability	Test retest reliability	Long'l construct validity	Clinical trial discrimination	Thresholds of meaning
Working Group Appraisal (n=20 including 7 PRPs)	+	+						
Tugwell 2005			+/–			+		
Shea 2004						+		+
Smith 1999								
Beaton 2015							+	
De Wit 2018							+	
Wells 2004			+					
March 2008							+	+/–
D'Agostino 2011						+/–		+
Bingham 2018			+		+/–			
Singh 2010			+					
Strand 2015			+/–					
Simon 2011						+		+/–
New data from Conaghan 2021					+			
Total available studies for each property			5	N/A	3	5	3	4
Total studies available for synthesis			5	N/A	2	4	3	4
Synthesis Rating	**GREEN From Working group**	**GREEN From Working group**	**GREEN**	**N/A**	**AMBER**	**GREEN**	**GREEN**	**AMBER**
OMERACT Endorsement	**Based on the OMERACT algorithm this instrument is: Provisionally endorsed *More research needed on test-retest reliability and thresholds of meaning.***							

F. Excerpt from OMERACT Handbook, Chapter 5, Instrument Selection (pg. 41-42; 44-45)

https://omeracthandbook.org/

11. Conduct synthesis across evidence available for each measurement property

All studies avoiding risk of bias in their design have now had their findings extracted and compared to the adequacy standards. The Working Group must now consider the synthesis of their information. OMERACT is using the best evidence synthesis approach blending Quality, Quantity, Consistency of findings, and Adequate (or better) Performance. This decision is guided by the work of others in best evidence synthesis groups (NQF 2013; Schellingerhout et al., 2012; Schmitt et al., 2015; Slavin, 1995). Best evidence synthesis looks for consistent evidence of good performance across multiple good quality studies of that property.

Quality has been determined at the level of quality appraisal as only those publications free of fatal flaws (GREEN, AMBER) are included in the synthesis. **Quantity, Consistency**, and **Adequacy** are now considered to complete the synthesis at this stage. For example, multiple high-quality studies could consistently show poor longitudinal construct validity of an instrument suggesting strong confidence against that measurement property for that instrument.

The literature gathered for each measurement property will be assigned a rating of GREEN (good evidence supporting this property, passes this element of the Filter), AMBER (some caution, or perhaps only one study on that property, but good enough to move forward) or RED (stop, evidence against this property or only poor-quality evidence) score. If there is no adequate quality evidence available on that property, it can be assigned a WHITE rating and await the creation of that evidence and future update of the rating.

Working Groups must gather all the evidence that they believe should be included in a synthesis for each of the six measurement properties required in the OMERACT Filter (construct validity, inter-method reliability, test-retest reliability, longitudinal construct validity, clinical trial discrimination, thresholds of meaning). Inter-method reliability (i.e., inter-rater, inter-machine) is new to Filter 2.2 to accommodate the lessons learned when integrating outcomes like imaging outcomes into OFISA. For these types of outcomes, the inter-rater reliability is a critical feature as there can be a lot of discordance between raters. In other situations, like a patient-reported outcome (PRO), sources of variability may not have been found and in that case the column will be marked NA (not applicable) and the related cell in the profile will be GREY and marked NA. This is not a weakness in the tool, just a measurement property that was not needed as a piece of evidence for that instrument.

The algorithm described in Figure 5.8 should be used as a guide for assigning the measurement property syntheses. A Green rating is assigned when there is consistent (at least two studies) evidence from studies with good enough quality supporting the instrument's performance in this measurement property. Note that the consistency of the evidence needs to be assessed across all the studies; it is not enough to find two studies with adequate evidence and decide not to continue reviewing the evidence - the entire body of evidence needs to be considered. A Red rating is assigned if there is an indication that this instrument is not performing well in this population and setting by demonstrating either inadequate findings in studies or if there are only studies deemed to not have good enough methods to provide credible evidence. White is assigned if there is no evidence available. Amber is assigned for all other situations.

Criteria for final rating				**Final rating for this measurement property**
Quality Of studies on measurement properties	*Quantity of good quality studies*	*Consistency across studies*	*Performance in this property*	

Good methods used	+	At least 2 pieces of evidence	+	Consistent findings	+	Adequate or better performance	→	GREEN
Good methods	+	At least 2	+	Consistent or Questionable	+	Inadequate performance	→	RED
Good methods	+	1 study only	…	NA	+	Inadequate performance		
Studies with fatal flaws	…	Not considered	…	Not considered	…	Not considered		
No evidence	…	0	…	NA	…	NA	→	WHITE
All other situations (Final rating not RED or GREEN or WHITE)							→	AMBER

Figure 5.8. Guide for synthesis ratings for each measurement property considering quality, quantity, consistency of findings across studies and adequacy of the performance on that measurement property.

The synthesis rating for each of the measurement properties are recorded on the Summary of Measurement Properties table in the "Synthesis Rating" row.

11	Conduct synthesis across evidence available for each measurement property	○

12.3. If no gaps exist, or if gaps cannot be filled, fill in SOMP Table with proposed level of endorsement of instrument

Working Groups now have a body of evidence that they feel is as complete as possible. Each measurement property has undergone the synthesis step described above and is represented by a GREEN, AMBER, or RED rating. Synthesis of this profile is then the final step in this process.

The algorithm described in Table 4 is used to determine the proposed level of endorsement: Endorsed, Provisionally endorsed, Not endorsed.

A GREEN in the synthesis rating row for every measurement property means a full endorsement of the instrument as having passed the OMERACT Filter 2.2.

A mixture of AMBER and GREEN ratings means provisionally passing the OMERACT Filter 2.2. When the recommendation is going to be AMBER (provisional), a statement of the work that needs to be done to bring it up to a full endorsement must also accompany it. AMBER is provisional not permanent. Working groups should commit to finding the remaining evidence and recognize that the completion of the evidence table could lead to a full endorsement OR to a decision that the instrument is not good.

Any WHITE ratings (a gap in the literature) or RED ratings (poor performance) found in the synthesis ratings across measurement properties means an instrument is lacking the supporting evidence and it would not be recommended for endorsement (do not endorse).

Table 4. OMERACT Algorithm to determine proposed level of endorsement	
Full endorsement	All SOMP columns have a synthesis rating of GREEN. The instrument fulfils the requirements of OMERACT Filter 2.2 for inclusion in a core set.
Provisional Endorsement	There is a mixture of GREEN and AMBER synthesis ratings across the measurement properties. The instrument is endorsed for provisional inclusion in a core set until additional information is obtained. The working group sets a research agenda and continues to work on this instrument to see if it can become a fully endorsed instrument.
Not endorsed	Any of the columns have either RED or WHITE synthesis ratings. No available evidence, large gaps in evidence or flawed instrument performance suggest that this instrument does not yet have the evidence to support its use in a core set at this time.

12	Decide if any gaps exist in evidence of measurement properties If gaps found, draft protocol for new study to fill gaps If no gaps, finish the SOMP table with proposed overall rating of instrument	○

G. Excerpt from Instrument selection workbook (pg. 34)

https://omeracthandbook.org/workbooks

11. Conduct synthesis across evidence available for each measurement property; fill in rating (Green/Amber/Red/White) on SOMP Table

*At this point, your SOMP table shows the results of the Good Methods check using the Red/Amber/Green colour in cells and the adequacy/performance of the results using the symbols "+", "+/-" and "-". Now, the Working Group needs to synthesize the evidence available for each measurement property (i.e. synthesize the evidence down each column). Fill in the row titled "**Synthesis Rating**" with the Working Group's assessment of the evidence for each measurement property.*

***Green** indicates synthesis of at least 2 studies with good methods showing positive support ("+") for the measurement property*

***Amber** indicates synthesis of only 1 study showing either positive or ambivalent support; or 2 or more studies showing ambivalent support or an inconclusive result*

***Red** indicates synthesis of studies with evidence that the instrument did not reach performance standards*

White indicates no studies assessing this measurement property; i.e. a gap in the evidence

12.3 Complete SOMP table with proposed level of endorsement

To the same SOMP table used to track "Good Methods Check" results, the overall adequacy of the results for each study, and the synthesis for each measurement property, the working group now adds the final synthesis statement in the 'OMERACT Endorsement' row.

Below is an example completed SOMP table. Delete this example and replace with your completed SOMP table.