

Appendix 1. COSMIN definitions of domains, measurement properties and aspects of measurement properties[12]

Domain	Term		Definition
	Measurement property	Measurement property aspect	
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g., using different sets of items from the same OMI (internal consistency); over time (test-retest); by different persons on the same occasion (inter-rater); or by the same persons on different occasions (intra-rater)
	Internal consistency		The degree of interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is due to 'true' [†] differences between patients
	Measurement error		The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Validity			The degree to which an OMI measures the construct(s) it purports to measure
	Content validity		The degree to which the content of an OMI is an adequate reflection of the construct to be measured
		Face validity	The degree to which (the items of) an OMI indeed seems to be an adequate reflection of the construct to be measured
	Construct validity		The degree to which the scores of an OMI are consistent with hypotheses (e.g., with regard to internal relationships, relationships to scores of other OMIs, or differences between relevant groups) based on the assumption that the OMI validly measures the construct to be measured
		Structural validity	The degree to which the scores of an OMI are an adequate reflection of the dimensionality of the construct to be measured
		Hypotheses testing	Idem construct validity
		Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted OMI are an adequate reflection of the performance of the items of the original version of the OMI
	Criterion validity		The degree to which the scores of an OMI are an adequate reflection of a gold standard
Responsiveness			The ability of an OMI to detect change over time in the construct to be measured
	Responsiveness		Idem responsiveness
Interpretability*			The degree to which one can assign qualitative meaning (i.e., clinical or commonly understood connotations) to an OMI's quantitative scores or change in scores

COSMIN: COnsensus-based Standards for the selection of health Measurement INstruments

* Not considered a measurement property, but an important characteristic of a measurement instrument

† The word “true” must be seen in the context of the CTT, which states that any observation is composed of two components—a true score and error associated with the observation. “True” is the average score that would be obtained if the scale were given an infinite number of times. It refers only to the consistency of the score and not to its accuracy.[70]