

Assessment criteria for evaluating content validity (Terwee et al., 2017)

Ten criteria for good content validity

Relevance 1. Are the included items relevant for the construct of interest? 2. Are the included items relevant for the target population of interest? 3. Are the included items relevant for the context of use of interest? 4. Are the response options appropriate? 5. Is the recall period appropriate?
Comprehensiveness 6. Are no key concepts missing?
Comprehensibility 7. Are the PROM instructions understood by the population of interest as intended? 8. Are the PROM items and response options understood by the population of interest as intended? 9. Are the PROM items appropriately worded? 10. Do the response options match the question?

Each criterion is rated as sufficient (+), insufficient (-), or indeterminate (?). The general rule to give a sufficient rating per criterion is:

+ : > 85% of the items of the PROM (or subscale) fulfill the criterion.

– : < 85% of the items of the PROM (or subscale) does fulfill the criteria.

? : No(t enough) information available or quality of (part of a) the study inadequate.

Guidance for determining the relevance rating, comprehensiveness rating, and comprehensibility rating per study:

	Relevance	Comprehensiveness	Comprehensibility*
+	At least criteria 1 and 2 are rated + AND at least two of the other three criteria on relevance are rated +. Criteria 1 and 2 (relevance for construct and population) are considered the most important criteria and therefore they need to be rated +. A maximum of 1 criterion rated – is allowed, but reviewers can also rate ± in that case.	Rating of criterion 6	Both criteria 9 and 10 are rated +.
-	At least criteria 1 and 2 are rated – AND at least two of	Rating of criterion 6	Both criteria 9 and 10 are rated –.

	the other three criteria on relevance are rated –.		
?	At least two of the criteria are rated ?.	Rating of criterion 6	At least one of the criteria is rated ?.
±	All other situations	Rating of criterion 6	One criterion is rated + and one is rated –.

Guidance for determining the content validity rating per study:

Rating	Description
+	The relevance rating is +, the comprehensiveness rating is +, and the comprehensibility rating is +.
–	The relevance rating is –, the comprehensiveness rating is –, and the comprehensibility rating is –.
±	At least one of the ratings is + and at least one of the ratings is –.
?	Two or more of the ratings are rated ?.

* Since we were unable to ask the population of interest directly, questions 7 and 8 were excluded from the analysis.

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., Bouter, L. M., de Vet, H. C. W., & Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation*, 27(5), 1159–1170. <https://doi.org/10.1007/s11136-018-1829-0>