JCHR

Journal of Community Health Research 2021; 10(2): 100-102.

A Brief Review of the Types of Validity and Reliability of scales in Medical Research

Sajjad Bahariniya [1], Mohammad Ezatiasar [2], Farzan Madadizadeh [3*]

1. Masters student of Health Services Management, School of Public Health, Shahid Sadoughi University of Medical Sciences, Yazd, Iran
2. School of Public Health, Semnan University of Medical Sciences, Semnan, Iran
3. Center For Healthcare Data Modeling, Departments of Biostatistics and Epidemiology, School Of Public Health, Shahid Sadoughi University Of Medical Sciences, Yazd, Iran

ARTICLE INFO

Letter to the Editor

Received: 9 March 2021
Accepted: 25 May 2021

OPEN ACCESS

Corresponding Author:
Farzan Madadizadeh
f.madadizadeh@ssu.ac.ir

How to cite this paper:
Bahariniya S, Ezatiasar M, Madadizadeh F. A Brief Review of the Types of Validity and Reliability in Medical Research. J Community Health Research. 2021; 10(2): 100-102.

Dear Editor in Chief

Reliability and validity are concepts used to evaluate the quality of psychometric properties of a scale in medical research. Reliability is related to reproducibility, and validity is related to the accuracy of a scale. It is important to check reliability and validity when using a scale or questionnaire. The present study provides a brief overview of the types of validity and reliability in medical research.

- **Reliability**

Reliability deals with whether the measuring instrument in different repetitions with the same conditions and the same subject group offers the same results or undergoes a drastic change (1). In other words, reliability refers to the stability of the results in the repetitive measurement of a scale.

The different types of reliability measurement methods are as follows (2)

1- Test-retest reliability

In this method, the subjects are measured twice at intervals of one or two weeks with the same instrument. Finally, the correlation (usually intraclass correlation coefficient (ICC) through two way mixed-effect model with absolute agreement type) is calculated between the scores recorded in the first and second measurements.

2- Interrater reliability

This method examines whether different people will achieve the same results if they use the desired tool or not. ICC and Kappa statistics used to measure interrater reliability for ordinal/interval and nominal responses, respectively (2).

3- Internal consistency

In this method, the internal consistency of the questions is examined. In other words, the question is answered whether the results of different parts of the tool are in line with the general results of the tool or not.

Kuder-Richardson Formula 20 (KR20), ' 'Cronbach's alpha (α), and person separation reliability (R) are three criteria for calculating internal consistency. The internal consistency of a tool is strongly influenced by random responses so that by increasing random responses in the data, the degree of internal consistency decreases (3).

Copyright: ©2021 The Author(s); Published by Shahid Sadoughi University of Medical Sciences. This is an open-access article distributed under the terms of the Creative Commons Attribution License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[Downloaded from jhr.ssu.ac.ir on 2026-11-19]

[DOI: 10.18502/jchr.v10i2.6582]

- **Validity**

Validity is a term that refers to the goal that the test is designed to achieve. It is adapted from the word "correct and permissible," and its meaning is "correctness." In other words, validity means how accurately a scale measures what is designed to be measured. If the research is highly valid, it means that it produces results that are consistent with reality.

The different types of validity measurement methods are as follows:

1- Face validity

It deals with the appearance of the tool, and the question is whether the scale you are considering is valid or not. Usually, word shifting can improve face validity to some extent. In fact, face validity examines the level of difficulty, the degree of appropriateness, and ambiguity of questions in scale or questionnaires. Face validity can be measured in two ways, qualitatively and quantitatively. In the qualitative method, several subjects or specialists (usually 5 to 10 people) are interviewed about the level of difficulty, appropriateness, and ambiguity in the questions. In a quantitative method, respondents are asked to rate each question in the questionnaire from "**Not at all important**," "**Slightly Important**," **Important,**" "**Fairly Important**," and "**Very Important**." Finally, the average rating given to each question is calculated and multiplied by the percentage of people who considered that question important and very important, and finally, a score is calculated for each question. Questions with a score of less than 1.5 will be removed from the set of questions (4).

2- Content validity

In this method, the content of the scale is examined. Content validity can be evaluated qualitatively and quantitatively. In the qualitative method, experts are asked to express their views on the observance of language grammar and how to score each question's options. If corrections are needed, corrections should be made according to their opinion (5). In the quantitative method, the content validity ratio (CVR) is calculated by examining the necessity of having a question, and the content validity index (CVI) index is calculated to check the relevance of the question to the purpose of the research (6).

3- Construct validity

Construct validity refers to the degree of efficiency of a scale and seeks to answer whether the results presented by the tool in question are consistent with theoretical evidence. This type of validity can be checked by factor analysis (exploratory and confirmatory factor analysis) (7).

4- Criterion validity

This type of validity compares the scale results with the results of other existing scales with the same meaning. The two most common forms of validity are predictive validity and concurrent validity (other types are convergent and discriminant). In predictive validity, based on the current scale, the characteristics of individuals in the future are predicted, and in simultaneous validity, the current instrument simultaneously examines the status of the participants with an existing instrument and compares their results. Pearson correlation coefficient is a measure to check this type of validity (8).

In summary, the validity and reliability of the tool are major concerns for health researchers. Tools that do not have the required validity and reliability do not provide reliable results. Researchers need to be familiar with the concepts of validity and reliability that are briefly presented in this article to be used in research desirably and thus pave the way to increase health research quality.

Author's contribution

S.B. and F.M. conceived of the presented idea. S.B. wrote the manuscript with support from F.M., S.B., and F.M. read the manuscript and verified it.

Conflict of interest

The author had no conflict of interest.

[Downloaded from jhr.ssu.ac.ir on 2026-11-19]

[DOI: 10.18502/jchr.v10i2.6582]

References

1. Carson SH, Peterson JB, Higgins DM. Reliability, validity, and factor structure of the creative achievement questionnaire. Creativity research journal. 2005;17(1):37-50.
2. Golafshani N. Understanding reliability and validity in qualitative research. The qualitative report. 2003;8(4):597-607.
3. Anselmi P, Colledani D, Robusto E. A comparison of classical and modern measures of internal consistency. Frontiers in psychology. 2019;10:2714.
4. Holden RR. Face validity. The corsini encyclopedia of psychology. 2010:1-2.
5. Beckstead JW. Content validity is naught. International journal of nursing studies. 2009;46(9):1274-83.
6. Lawshe CH. A quantitative approach to content validity. Personnel psychology. 1975;28(4):563-75.
7. Brown JD. What is construct validity. Available at jalt. org/test/bro_8. htm (accessed 10 June 2013). 2000.
8. Golafshani N. Understanding reliability and validity in qualitative research. The qualitative report. 2003;8(4):597-607.

[Downloaded from jhr.ssu.ac.ir on 2026-11-19]

[DOI: 10.18502/jchr.v10i2.6582]