

Reliability and validity: Importance in Medical Research

Ishtiaq Ahmed¹, Sundas Ishtiaq²

Abstract

Reliability and validity are among the most important and fundamental domains in the assessment of any measuring methodology for data-collection in a good research. Validity is about what an instrument measures and how well it does so, whereas reliability concerns the truthfulness in the data obtained and the degree to which any measuring tool controls random error. The current narrative review was planned to discuss the importance of reliability and validity of data-collection or measurement techniques used in research. It describes and explores comprehensively the reliability and validity of research instruments and also discusses different forms of reliability and validity with concise examples. An attempt has been taken to give a brief literature review regarding the significance of reliability and validity in medical sciences.

Keywords: Validity, Reliability, Medical research, Methodology, Assessment, Research tools.

DOI: <https://doi.org/10.47391/JPMA.06-861>

Introduction

The validity factor in research is about the assessment of how accurate is the measure of data or results, while the reliability factor is about the consistency of a measure of results or data.¹ Reliability and validity are the fundamental concepts in research, used to assess the quality of research which together specify how well a methodology, data-collection technique or data analysis is planned to measure study variables or parameters.² The significance of assessing the research tool's accuracy and consistency i.e., validity and reliability respectively, has been reported in literature, but usually this aspect is ignored or not measured properly in medical sciences by researchers, especially those in developing countries. This limitation has been linked mainly due to the dearth of knowledge among researchers about how to measure the validity and reliability in research.³

The validity refers to how well the data represents the true findings among the participants of a study and among the similar individuals not participating in the study.⁴ This applies to all types of clinical research, like prevalence, diagnosis, interventions or disease-association studies. The

¹Al-Nafees Medical College, Isra University, Islamabad, Pakistan; ²Fauji Foundation Hospital, Foundation University Medical College, Islamabad, Pakistan.

Correspondence: Ishtiaq Ahmed. e-mail: surgish2000@yahoo.com

validity of a research is more important and harder to evaluate than reliability. To gain valuable results, the data-collection methods should be valid, research should be measuring what it plans to measure to ensure that the conclusions drawn from results are also valid.^{5,6}

A meticulous assessment of validity and reliability comprises an assessment of the data-collection methodology which should be precise to present a standard and acceptable research. This provides a good relation to interpret data, especially from the psychometric tools, like questionnaire, symptom scales and observer ratings, which are frequently used in research regarding clinical practice, medical education and administration.^{7,8} The data-collection errors in terms of validity and reliability not only jeopardise the ability to get important results, but can also damage significance of scores in preparing a good research.³

In modern research, validity and reliability are crucial concepts used in enhancing the precision and accuracy of the evaluation and assessment of a research project.^{5,9} Without this, it is very hard to describe the effects of measurement errors over theoretical interactions that are being measured. The reliability and validity of data can be enhanced by using various types of methodology in the data-collection process to obtain the correct information. Researchers usually fail to give an account of the reliability of their tools and are also unable to give the details of inextricable link between validity of scale and effective research.^{10,11} Construction or development of variables and the use of instruments or tests to measure or quantify these variables are crucial steps in research. Using an appropriate or better methodology with high validity and reliability in data-collection enhances the scientific quality of the research. Therefore, the objective of establishing research validity and reliability is basically to ensure that the data is comprehensive, replicable and the results produced are accurate.

The current review was planned to describe and explore comprehensively the reliability and validity of research instruments, and to discuss different forms of reliability and validity with concise examples.

Materials and Methods

The current literature review comprised online search on

PubMed, Medline, Medscape, Research Gate, Excerpta Medica Database (EMBASE), Health Management Information Consortium (HMIC), Cumulative Index to Nursing and Allied Health Literature (CINAHL+), Google search and PakMediNet databases which carry archives of majority of biomedical journals from across the world. Scientific publications published from January 1, 2005, to April 30, 2020, were searched using key words 'reliability', 'validity', 'assessment', 'tools', 'methodology', and 'medical research'. All relevant scientific papers, written in English, published within the stipulated timeframe were included. News articles, non-scientific commentary and reports were excluded.

Results

What is reliability?: Reliability is defined as the consistency of a method in measuring something. The measurement is considered reliable if the same result can be attained consistently by applying the same methodology under similar conditions. For example, in the measurement of water temperature at different times but under the same conditions, the thermometer always displays the same temperature, it proves that the results are reliable. In contrast, if the use of a symptoms questionnaire to diagnose a disease by different clinicians fails to give the same diagnosis, it shows that the questionnaire has a low reliability in terms of measuring that disease.^{4,12,13}

What is validity? Validity is defined as to how accurately a methodology measures a variable that it intends to measure. High reliability is one of the indicators showing that a result is valid and any unreliable methodology of data-collection is probably not valid. For example, if the measurement of temperature with a thermometer each time shows different readings, even under carefully controlled similar conditions, probably the thermometer is malfunctioning and therefore its temperature reading is not valid. Similarly, if the results of a symptoms questionnaire shows a reliable diagnosis when data is collected on different occasions by different clinicians, it indicates that the questionnaire has high validity as a measurement of the diagnosis in that medical condition.^{13,14}

Understanding reliability vs validity:^{4,5,15} Validity and reliability are different terms in meaning, but are closely related to each other. The reliability is the extent to which the results can be reproduced when the research is repeated under the same conditions, whereas validity is the extent to which the results really measure what they are supposed to measure.

In research, it is difficult to assess validity compared to reliability, but validity assessment is more important. It is

very important that methodology and data-collection methods should be valid to achieve useful results and the research must measure what it claims to measure. This also guarantees that the discussion of results and conclusions drawn are also valid.

However, to ensure validity, the element of reliability is not enough on its own, because even a reliable test may not precisely reflect the real situation. A good example in this regard is that a thermometer used to measure temperature gives reliable results, but if it is not calibrated properly it may give false readings. Therefore, in this case the temperature measurement is not valid.

How are reliability and validity assessed? Reliability is assessed by testing the consistency of results across time, between different observers and across the parts of the test itself. It can be done by comparing different versions of the same measurement. The validity of a research is measured by examining how well the results of a study compare with or correspond to the existing recognised theories and other measures of the same idea. Validity is usually difficult to evaluate, but can be assessed by comparing the results to other relevant data or theory. Different methods of estimation of validity and reliability are usually divided into different types.⁵

Types of reliability: There are different types of reliability which can be assessed through various statistical methods.^{12,13}

- a. **Test-retest reliability:** It is consistency of a measure across time, meaning that when you repeat the test, do you get the same results?
An example is a questionnaire designed to measure personality traits completed by a group of people. When the same questionnaire is repeated after days, weeks or months, and give the same responses, it indicates a high test-retest reliability.
- b. **Inter-rater reliability:** It is the consistency of a measure across observers or raters, meaning that when different observers conduct the same assessment, they get the same results.
An example is checklist-based assessment in which different examiners submit substantially different assessments for the same student assignment or project. It clearly shows that the checklist used for the assessment has low inter-rater reliability.

- c. **Internal consistency reliability:** It is the consistency of the measurement itself, meaning that it measures if you get similar results from different parts of a test which is designed to measure the same thing. An example in this regard is that you have designed a questionnaire to measure the self-esteem of a person. If the results are randomly split into two halves, it shows a strong correlation between these two sets of results, while varied results in the two halves indicate a low internal consistency in the results.

Though important, reliability contributes to the validity of measurement tool, like questionnaire etc, it is not an adequate condition for the validity of a tool. The lack of reliability may be due to discrepancy among measuring tools or observers like instability of the variable being measured or a questionnaire^{4,14} which invariably affects the validity of such a tool or questionnaire.^{16,17}

Types of validity: The validity of data can be based on estimation of the three main types of evidences which can be assessed through statistical analysis or expert judgement.^{4,13,14}

- a. **Construct validity:** It is the adherence of an assessment to the existing knowledge and theory of the concept which is being measured. An example is that of a questionnaire on self-esteem which could be assessed by gauging traits, like optimism and social skills, which are known to be associated with self-esteem concept. The strong correlation among the scores for self-esteem and associated traits is highly indicative of high construct validity.
- b. **Content validity:** It is the extent to which an assessment covers all aspects of the idea being measured. For example, a test which assesses the level of French language in students contains writing, reading and speaking, but contains no listening component. The listening is an essential component of language ability, as agreed by experts. As such, this test lacks content validity for measuring the overall level of ability in French language.
- c. **Criterion validity:** It is the extent to which the data measured corresponds to the other valid measures of the same domain or concept.

For example, analysis of the opinion of voters in a locality about their favourite and winning candidate. Subsequently, if the results of this survey accurately predict the outcome of election, it will show high criterion validity of the survey.

- d. **Predictive validity:** It refers to the degree to which operationalisation can predict or correlate with other measures of the same construct that are measured at some time in the future. For example, predictive validity is slightly different in job applicants for selection at the time of interview and then after the individuals work in the job for a year or so. Their test scores are correlated with their first year job performance to see the validity of cognitive tests administered for selection.

To assess cause-and-effect relationship of validity it is important to consider internal and external validity. Internal validity refers to how accurately the results obtained in reality measure what they were considered to measure, whereas external validity shows how precisely the results obtained represent the reference sample or population from which the study sample was drawn.¹⁷

How to ensure validity and reliability in research? The validity and reliability of results depends on conceiving a strong research design, choosing appropriate methodology, proper sample selection and by conducting the research meticulously, consistently and carefully.

Ensuring validity: Validity of research should be assessed at the earliest possible stage, like while deciding how to collect data. If a researcher is using rating scale or scores to measure the variations in parameters like physiological properties or psychological traits, then it is essential that the results should reflect the real disparities or variations as accurately as possible.^{14,18}

To ensure this it is better to select a high-quality, targeted technique to measure data. The technique of data collection should be based on existing or latest knowledge and thoroughly researched. For example, while collecting data on personality trait, it is more appropriate to use a standardised questionnaire which is considered more valid and reliable. If researchers want to develop their own questionnaire, then it should be based on established findings or theory from previous studies, and the questioner should be precisely and carefully worded. Secondly, the sampling methods should be appropriate in sample selection. Thirdly, to achieve generalised and valid results, the subjects or parameters to be assessed should

be clearly defined e.g. specific age group, geographical location or profession. Finally, the size should be adequate which should be representative of the population, procedure or parameter.^{4,18,19}

Ensuring reliability: The reliability of a research should be taken into account throughout the data-collection process. While using a data-collection technique or tool, it is very important that data should be stable, precise and reproducible. One should be sure that methodology should be planned carefully and applied consistently, especially when multiple researchers are involved in data-collection.^{11,19}

It is very important to standardise research or data-collection conditions. While collecting data the conditions should be kept as consistent as possible to minimise the external factors that may influence the process and may lead to result variation. For example, in experimental setup, all participants should be given the same information and assessed under similar conditions.²⁰

Discussion

In research critique, an essential component is to assess how amicably the validity and reliability issues have been addressed and how it influences the decision about whether or not to implement the research findings into medical practice. A good quality research will provide evidence of how all these issues have been addressed in a study. This will help in assessing the research reliability and validity and also help in deciding whether or not to apply these results in clinical practice.^{8,20}

The reliability of research denotes to a measurement which provides reliable, consistent, precise data which is trustworthy and can be repeated, especially in quantitative research.^{21,22} It also designates the extent to which it ensures that reliable data measurement can be applied across time and various parameters without bias, meaning error-free. In qualitative research, it designates as to when a researcher's approach is reliable across different projects and different researchers.^{8,10}

The research study validity refers to how well the study participants' results represent the true findings among the same group of individuals not participating in the study. This concept applies to all kinds of clinical studies, including those related to diagnosis, interventions, association and prevalence of a disease.

The research study validity comprises two domains i.e., internal validity and external validity.¹¹ The internal validity is defined as "the extent to which the observed results represent the truth in the population we are studying and thus are not due to methodological errors". The different

factors can threaten the internal validity of a study, like measurement errors and sample selection, and the researchers should consider all these to avoid errors.^{3,23} After establishing the internal validity of a study, the researcher can proceed to establish the external validity by assessing whether the study results can be applied to similar subjects in a different setting or not. The external validity of a study shows up to what extent the study results are generalisable to other patients in daily practice, especially over the population for which the sample is thought to represent.

Lack of internal validity of a study suggests that the results obtained differ from the reality, so one cannot draw any conclusions from these results. Therefore, if a trial results are not valid internally, the external validity of that study is irrelevant. The lack of external validity denotes that the trial results cannot be applied to patients who differ from the study population. Therefore, they could lead to low treatment adoption which is tested in the trial by other clinicians.^{19,23}

Adequate quality control, carefully ensuring the study planning and implementation of strategies like sample size, data-collection, data analysis and proper recruitment of subjects are essential to enhance the internal validity of a research. The external validity of a research can be increased by using broad-based inclusion criteria in a study population which resembles more closely to real-life patients and in clinical trials, by selecting an intervention which is more feasible to apply.^{23,24} Measurement of errors done simultaneously through multiple sources instead of one source at a time is more reliable in sensory testing.¹⁸ Another study has stressed the important role of construct validity by using scale scoring in driving proper conclusions.²⁵ Similarly, Burns et al.²⁶ and Campos et al.²⁷ have also stressed the importance of validity and reliability in research.

Aoki et al.¹⁰ also emphasised in their systemic review the importance of developing or identifying a patient-reported outcome measure with content validity as a future research agenda by using preliminary evidence for reliability and validity of the Shoulder Pain and Disability Index. Studies have stressed considerable improvement in validity and reliability in designing a valid assessment tool for transitional incidents identification in medical records of primary and secondary care.^{28,29} Similarly, many other studies have also supported the importance of establishing the validity and reliability in research to produce a good quality research.^{6,24,30-33}

Evaluation of the quality of measure: Proper or accurate assessment of research reliability and validity are key

indicators of the quality of a measure or assessment. Therefore, it is crucial to mention that how accurately the reliability and validity issue has been addressed and how it will influence the decision about whether or not to implement the study findings into clinical practice. In quantitative research, the rigour is determined by assessing the reliability and validity of the instrument or tools used in data-collection or measurement. The research will be considered fruitful beyond any doubt if error margins are low and research results are of high standard.^{4,16,20}

The trustworthiness (validity and reliability) of data is the foundation of a research to draw a good conclusion.^{16,34} Kimberlin et al.⁵ stressed that assurance of the quality and integrity of measuring instrument is a prerequisite in reliability and validity of research. A study tried to create an evidence-based assessment tool to measure reliability and validity.³⁵ Another study concluded that patient-centred assessment methodology is a reliable and valid tool in the assessment of disease complexity during the early stage of secondary care hospital admission.³⁶

A good standard research should provide evidence that all these issues have been properly addressed in a study. This will help to assess the research reliability and validity and also help decide whether or not one should apply the research findings in area of clinical practice.^{6,8}

Where to mention validity and reliability in a paper or research thesis: It is appropriate that the validity and reliability issues be discussed in various sections of a research project, dissertation or thesis to make the work more trustworthy and credible. The emphasis should be given while planning the research and interpreting its results. In the literature review section, information about what other researchers have done to plan and improve the methods which are more reliable and valid should be mentioned for guidance. In methodology, the information regarding different measures adopted in sample selection, sample size calculation, sample preparation, external conditions and measuring techniques should be mentioned to ensure the reliability and validity of a study. If reliability and validity is calculated, its values should be added along with the main results. In the discussion section, comments are essential regarding how the results are reliable and valid and if they are consistent and reflect the true values or not? If not, reasons should be given.¹

Conclusion

It is imperative that reliability and validity factors are kept in mind to ensure accurate results, which enhances the probability of arriving at the right conclusion in research. Reliability is essential for the validity of research, but is not adequate on its own.

Disclaimer: None.

Conflict of Interest: None.

Source of Funding: None.

References

- Middleton F. Reliability vs validity: what's the difference? [Online] 2019 [Cited 2020 May 28]. Available from: URL: <https://www.scribbr.com/methodology/reliability-vs-validity/>
- Taherdoost H. Validity and Reliability of the Research Instrument; How to Test the Validation of a Questionnaire/Survey in a Research. SSRN Electronic Journal 2016; 5: 28-36
- Noble H, Smith J. Issue of validity and reliability in quantitative research. Evid Based Nurs 2015; 18: 34-5.
- Haradhan M. Two Criteria for Good Measurements in Research: Validity and Reliability. Annals of Spiru Haret University Economic Series 2018; 17: 58-82
- Kimberlin CL, Winterstein AG. Validity and Reliability of Measurement Instruments Used in Research. Am J Health Syst Pharma 2008; 65: 2276-84.
- West J. Assessing research quality. Research Connections. [Online] 2020 [Cited 2020 May 28]. Available from: URL: <https://www.researchconnections.org/content/childcare/understand/research-quality.html>
- Shekharan U, Bougie R. Research Methods for Business: A Skill Building Approach. 5th ed. New Delhi: John Wiley; 2010, Pp 123-9.
- Heale R, Twycross A. Validity and reliability in quantitative studies. Evid Based Nurs 2015; 18: 66-7.
- Tavakol M, Dennick R. Making Sense of Cronbach's Alpha. Int J Med Educ 2011; 2: 53-5.
- Aoki K, Hall T, Takasaki H. Reporting on the level of validity and reliability of questionnaires measuring Katakori severity: A systematic review. SAGE Open Med 2019; 7: 1-13
- Akturk Z. Reliability and validity in medical research. Dicle Med J 2012; 39: 316-9.
- Reliability (Statistics): Wikipedia. [Online] 2020 [Cited 2020 June 1]. Available from: URL: [https://en.wikipedia.org/wiki/Reliability_\(statistics\)](https://en.wikipedia.org/wiki/Reliability_(statistics))
- Bannigan K, Watson J R. Reliability and validity in a nutshell. J Clin Nurs 2009; 18: 3237-43.
- Validity (Statistics): Wikipedia. [Online] 2020 [Cited 2020 May 25]. Available from: URL: [https://en.wikipedia.org/wiki/Validity_\(statistics\)](https://en.wikipedia.org/wiki/Validity_(statistics)).
- Twycross A, Shields L. Validity and Reliability-What's it All About? Part 2: Reliability in Quantitative Studies. Paediatric Nursing 2004; 16: 36.
- Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. Philadelphia, USA: Lippincott William and Wilkins; 2008: Pp. 128-47.
- Bolarinwa OA. Principles and methods of validity and reliability testing of questionnaires used in social and health science researches. Niger Postgrad Med J 2015; 22: 195-201.
- Moana-Filho E J, Alonso AA, Kapos FP, Leon-Salazar V, Gurand SH, Hodges JS, et al. Multifactorial Assessment of Measurement Errors Affecting Intraoral Quantitative Sensory Testing Reliability. Scand J Pain 2017; 16: 93-8.
- Haradhan M. Two Criteria for Good Measurements in Research: Validity and Reliability. Paper No. 83458, UTC. Munich Personal RePEc Archive. [Online] 2017 [Cited 2020 Jan 03]. Available from: URL: <https://mpra.ub.uni-muenchen.de/83458/> MPRA
- Roberts P, Priest H. M. Reliability and validity in research. Nurs Stand 2006 20; 44: 41-5.
- Blumberg B, Cooper D R, Schindler P S. Business Research Methods. Berkshire: Education. 2005:p25-80.

22. Chakrabartty SN. Best Split-Half and Maximum Reliability. *IOSR J of Res & Method in Edu* 2013; 3: 1-8.
 23. Patino CM, Ferreira JC. Internal and external validity: can you apply research study results to your patients? *J Bras Pneumol* 2018; 44: 183.
 24. Price PC, Jhangiani R, Chant A. Reliability and validity of measurement. In: *Research Methods in Psychology*. 2nd ed. Victoria, B.C.: BCcampus; 2015: Pp 85-92.
 25. Flake J K, Pek J, Hehman E. Construct Validation in Social and Personality Research: Current Practice and Recommendations. *Soc Psychol Personal Sci*. 2017; 8: 1-9.
 26. Burns GN, Morris MB, Periard DA, LaHuis D, Flannery NM, Carretta TR, et al. Criterion-Related Validity of a Big Five General Factor of Personality from the TIPI to the IPIP. *Int J Sel Assess*. 2017; 25: 213–22.
 27. Campos CMC, da Silva Oliveira D, Feitoza AHP, Cattuzzo MT. Reliability and Content Validity of the Organized Physical Activity Questionnaire for Adolescents. *Educ Res* 2017; 8: 21-6.
 28. van Melle MA, Zwart DLM, Poldervaart JM. Validity and reliability of a medical record review method identifying transitional patient safety incidents in merged primary and secondary care patients' records. *BMJ Open* 2018; 8: e018576.
 29. Mohammadi M, Larijani B, Tabatabaei SM, Nedjat S, Yunesian M, Nayeri FS. A study of the validity and reliability of the questionnaire entitled "physicians' approach to and disclosure of medical errors and the related ethical issues". *J Med Ethics Hist* 2019; 12: 2.
 30. Hall BJ, Puffer E, Murray LK, Ismael A, Bass JK, Sim A, et al. The importance of establishing reliability and validity of assessment instruments for mental health problems: An example from Somali children and adolescents living in three refugee camps in Ethiopia. *Psychol Inj Law* 2014; 7: 153–64.
 31. Connell J, Carlton J, Grundy A. The importance of content and face validity in instrument development: lessons learnt from service users when developing the Recovering Quality of Life measure (ReQoL). *Qual Life Res* 2018; 27: 1893–902.
 32. Lachin JM. The role of measurement reliability in clinical trials. *Clin Trials* 2004; 1: 553-66.
 33. Stephanie K, Pihet K. Sandrine. Reliability, validity and relevance of needs assessment instruments for informal dementia caregivers: a psychometric systematic review. *JBI Evid Synth* 2020; 18: 704-42.
 34. Thatcher R. Validity and Reliability of Quantitative Electroencephalography. *J Neurotherapy* 2010; 14: 122-52.
 35. Haynes M C, Ryan N, Saleh M, Winkel A F, Ades V. Contraceptive Knowledge Assessment: Validity and Reliability of a Novel Contraceptive Research Tool. *Contraception* 2017; 95: 190–7.
 36. Yoshida S, Matsushima M, Wakabayashi H, Mutai R, Murayama S, Hayashi T, et al. Validity and Reliability of the Patient Centered Assessment Method for Patient Complexity and Relationship with Hospital Length of Stay: a Prospective Cohort Study. *BMJ Open* 2017; 7: e016175.
-