Perioperative Care and Operating Room Management 31 (2023) 100315

Contents lists available at ScienceDirect

Perioperative Care and Operating Room Management

journal homepage: www.elsevier.com/locate/pcorm

Tutorial on how to calculating content validity of scales in medical research

Farzan Madadizadeh [a], Sajjad Bahariniya [b,*]

[a] Center for Healthcare Data Modeling, Departments of Biostatistics and Epidemiology, School of Public Health, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

[b] MSc student of Health Services Management, School of Public Health, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

ARTICLE INFO

Keywords:
Content validity
Psychometric properties
Scale

ABSTRACT

Background: Due to the high importance of CV assessment and the lack of user-friendly guide, this study was conducted with the aim of providing a tutorial on content validity.

Methods: This is a tutorial study and the stated information has been collected from reliable sources such as books and similar articles.

Results: The standardization of a tool in research is a necessary condition for the accuracy of research results. One of the methods of standardizing a questionnaire is to check the content validity (CV). CV means how much a scale measures exactly what it was designed for. CV is checked both qualitatively and quantitatively. In the quantitative method, the content validity ratio (CVR) and the content validity index (CVI) are measured. CVR and CVI measure the necessity and relevance of a scale's items, respectively.

Conclusion: Researchers in the field of questionnaire studies need to be familiar with how to calculate content validity in order to improve the quality of research in the field of medical sciences.

1. Introduction

Today, medical questionnaires and scales have an undeniable role in measuring various variables and advancing research goals. Tools, if they belong to another language, need translation, cultural adaptation, and psychometric properties. One of the psychometric features is narrative. Validity means that the content of the questionnaire is in line with the purpose of its design. There are different types of validity in the psychometric field of tools such as face validity, content validity, structural validity, criterion validity and so on.

Examining the content of the questionnaire by experts is one of the best ways to gather evidence in support of the validity of a measured tool. In the validity of the content, the content of the test is examined to ensure that the content of the test represents the structure it claims to measure. Content validity indicates the adequacy of sampling the content of the measuring instrument.[1] CV is the answer to the question of whether the content of this measuring device represents the content or set of features being measured. It is essentially about judgment so that the person or persons judge the identity of the items.[2,3]

Previous study have shown there were two scientifically approaches for examining content validity included qualitative and quantitative method.[2,4] Regarding to Qualitative content assessment, it refers to the necessity, simplicity and relevance of the questionnaire items that are checked by receiving the opinion of experts.[5,6] In terms of Quantitative assessment of content validity, it will be done after receiving qualitative feedback from the experts and correcting the items content. To examine content validity quantitatively, there are two indicators called content validity ratio (CVR) and content validity index (CVI). The criteria evaluated by these two indicators are fundamentally different. In the content validity ratio index, the necessity of an item is examined from the expert's opinion, while in the content validity index, the relevance of the items to the research goal from the experts' point of view is considered. In the following sections of this tutorial article, detailed descriptions of each method are provided.[7]

2. Determining content validity

Basically, this method was introduced in 1983 by Waltz and Basel and is now used in many studies. The relevance index is considered quantitatively as the main criterion for measuring content validity. Rong Tosantham (1990) believes that two points should be considered in examining the validity of content; ensure that the most relevant and accurate content is selected and that the tool items are best designed to measure content. The first case can be calculated by the content validity

* Corresponding author.
E-mail address: sajjadbahari98@gmail.com (S. Bahariniya).

https://doi.org/10.1016/j.pcorm.2023.100315
Received 20 August 2022; Received in revised form 24 February 2023; Accepted 3 April 2023
Available online 6 April 2023

Table 1
Lawshe table to determine the minimum numerical value of CVR.[16]

Number of experts	Acceptable CVR Value
5	0.99
6	0.99
7	0.99
8	0.75
9	0.78
10	0.62
11	0.59
12	0.56
13	0.54
14	0.51
15	0.49
20	0.42
25	0.37
30	0.33
35	0.31
40	0.29

ratio and the second case by the content validity index.[8]

2.1. Content validity ratio (CVR)

This index was first developed by Lawshe (1975). Calculating the content validity ratio index makes the researcher ensure the selection of the most important and best content statistically. In this approach, a group of experts is asked to determine whether an item in a set of other items is necessary and important to operate a theoretical structure.[1,9] The panel of experts is asked to review each item based on the following 3-part spectrum:

1. Necessary 2. Useful, But Not Necessary 3. Not Necessary

The answers for each item are then calculated based on the following formula (Eq. (1)):

$$\mathbf{CVR} = \left(\frac{\mathrm{Ne} - \frac{\mathrm{N}}{2}}{\frac{\mathrm{N}}{2}} \right) \tag{1}$$

In this formula: **Ne = number of specialists who have selected the required option.**

N = Total number of evaluators

Lawshe has prepared a table to determine the numerical value of the content validity ratio (CVR), which is called the "minimum CVR accepted value". The result after the calculation is compared with the criteria in the table according to the number of specialists, which if the number is larger than the table, indicates that the relevant item has an acceptable level of statistical significance ($P < 5\%$) Is essential and important in this tool.[10,11] Baghestani et al. (2019) recalculated the critical values of the Lawshe table using the Bayesian method, taking into account the initial default probabilities.[11] To the best of our knowledge, a one tailed test hypothesis with considering a significant level 5% and power of 80% were used for calculation of critical value. Although the original methods used for calculating CVR critical were not reported in Lawshe's article on content validity, and as both Lawshe and Schipper have since passed away, it is now not possible to gain clarification (Wilson et al., 2012).[12] Unfortunately, the efforts that have been made in recent years to recalculate and solve the problems of these critical values have all been unsuccessful.[13–15]

Although Lawshe's proposed method is somewhat criticized, for example, by increasing the number of experts from 8 to 9, the value of CVR has increased and in other cases decreased. But so far no more accurate method has been proposed to calculate the content validity ratio.[17] The general rule is that if the numerical value of CVR based on the number of experts was higher than the acceptable values mentioned in the Table 1, the content validity of tool will be accepted. As a rule of thumb, if less than half of the experts do not agree about necessity of an item, that item will have no necessary.

Table 2
CVR calculation checklist and assessment questions.

items	Assessment questions by raters		
	Necessary	**Useful, But Not Necessary**	**Not Necessary**
q1			
q2			
q3			
q4			
q5			

Table 3
CVI calculation checklist.

Items	Assessment questions by raters			
	Not relevant	**Somewhat relevant**	**Relevant**	**Completely relevant**
q1				
q2				
q3				
q4				
q5				

Example 1. Suppose that a questionnaire has 5 items. The CVR can calculate as follows:

If this checklist is provided to 10 specialists and experts, the minimum acceptable level for CVR according to Lawshe table is **0.62**. For example, if in q1, 9 out of 10 experts choose the option "necessary"; will have:

$$\mathbf{CVR} = \left(\frac{\mathrm{Ne} - \frac{\mathrm{N}}{2}}{\frac{\mathrm{N}}{2}} \right) \mathbf{CVR} = \left(\frac{9 - \frac{10}{2}}{\frac{10}{2}} \right) = 0.80$$

Therefore, the first question has an acceptable CVR greater than 0.62.

2.2. Content validity index (CVI)

Content Validity Index (CVI), as an important procedure in scale development, is the degree to which the items of a tool are related to the purpose of the tool, which calculated based on two different levels such as at the item level (I-CVI) and at the overall scale level (S-CVI). The number of experts needed to judge the content validity of the tool is at least five people and the maximum is usually 10 people.[1,18,19]

To calculate the content validity index (CVI), we need to use the following formula, which proposed by Waltz and Basel (1983):

$$\mathbf{CVI} = \left(\frac{\text{Number of raters giving a rating of } '3' \text{ or } '4'}{\text{total number of raters}} \right) \tag{2}$$

According to the above formula, usually 5 criteria of ambiguity, simplicity, clarity, specificity and relevance are calculated based on a four-part Likert scale for each of the five dimensions. In the following years, toolmakers initially reduced these 5 criteria to three criteria of clarity, simplicity, and relevance. Content validity index (CVI) score above 0.79 is considered appropriate, between 0.79 and 0.70 is questionable and needs to be corrected and revised, less than 0.70 is unacceptable and should be removed.[20]

Lane, Paulit, and other instrumentation experts (1970) believe that the issue of "relevance" is more important in calculating the Content Validity Index (CVI). Also suggest as follows:(Tables 2, 3)

(1) **Not related**
(2) **is somewhat related**

Table 4
The minimum numerical value of CVI.

Number of panel people	Minimum CVI Value	Number of people need to rate 3 or 4
2	0.8	2
3	1	3
4	1	4
5	1	4
6	0.83	5
7	0.83	5
8	0.83	6
9	0.78	8
10	0.79	8

(3) **is related**
(4) **It is completely relevant**

Example 2. Suppose Questionnaire X has 5 questions. The CVI is calculated as follows:

If this checklist is provided to 10 specialists and experts, the minimum acceptable level for CVI according to Lawshe table is **0.79 (**see Table 4). For example, if in Q1, 9 out of 10 experts choose the option "is related " and "It is completely relevant"; will have:

$$\mathbf{CVI} = \left(\frac{\text{Number of raters giving a rating of } '3' \text{ or } '4'}{\text{total number of raters}}\right) = \frac{9}{10} = 0.90$$

Therefore, the first question has an acceptable CVI greater than 0.79.

The method of calculating the two indicators of CVR and VCVI is summarized in Fig. 1.

2.3. Multidimensional content validity

In multi-dimensional scales, which generally have more questions and more dimensions than normal scales, the calculation of CVI and CVR is similar. In calculating CVI multiple scales are formed of two types I-CVI and S-CVI which are explained in the following section. Only in the case of reliability, Cronbach's Alpha can be calculated separately for each dimension, as well as the overall Cronbach's Alpha for the entire scale.

2.3.1. I-CVI

Item content validity (ICVI) is the content validity that is calculated for each item of the questionnaire, so for each question in the questionnaire, an ICVI will be calculated and then according to the rule of thumb if the CVI value for each item is greater than 0.78 that item remains in the set of questions otherwise it will be the candidate to elimination.[21]

2.3.2. S-CVI

Paulit, Beck and Owen (2007) in a very interesting and comprehensive article suggest new aspects of calculating the content validity index in the continuation of their previous study. In his article entitled "Is the content validity index an acceptable indicator for calculating content validity?" Evaluations and Recommendations» Criticize the previous methods of calculating the content validity index and suggest a new method. He states that the evaluation of the content validity of the whole tool (S-CVI) is one of the basic steps in increasing and improving the validity of the tool structure, and also content validation is one of the important issues for an accurate measurement and determination. Content validity is considered. It was mentioned earlier that there are two different methods for calculating the validity of total tool content (S-CVI), which unfortunately researchers do not mention in their research.[22,23]

S-CVI Calculation

(1) **Universal agreement method (UA)**

In this method, we divide all the items that all retailers have considered completely relevant or partially relevant by the total number of questions. In other words, those questions whose CVI value is one are counted and divided by the total number of questions.[24,25]

$$\text{SCVI} - \text{UA} = \frac{\text{sum of UA scores}}{\text{number of item}} \quad (3)$$

(1) **Averaging method (AVE)**

Another approach to calculating S-CVI is to calculate the I-CVI for each tool item and then consider the average of the total I-CVI across the items. Usually, the acceptance criterion of S-CVI among instrument makers is 0.9 as an excellent criterion and the numerical value of 0.80 as the lower acceptance limit of the content of the instrument.

$$\text{SCVI} - \text{AVE} = \frac{\text{sum of I} - \text{CVI scores}}{\text{number of item}} \text{or} \frac{\text{sum of proportion relevance rating}}{\text{sum of expert}} \quad (4)$$

Due to conservativeness of averaging method (SCVI- AVE) in calculation of S-CVI values, it is more reliable than the SCVI-UA method.

Example. suppose that we have 10 items questionnaire and for

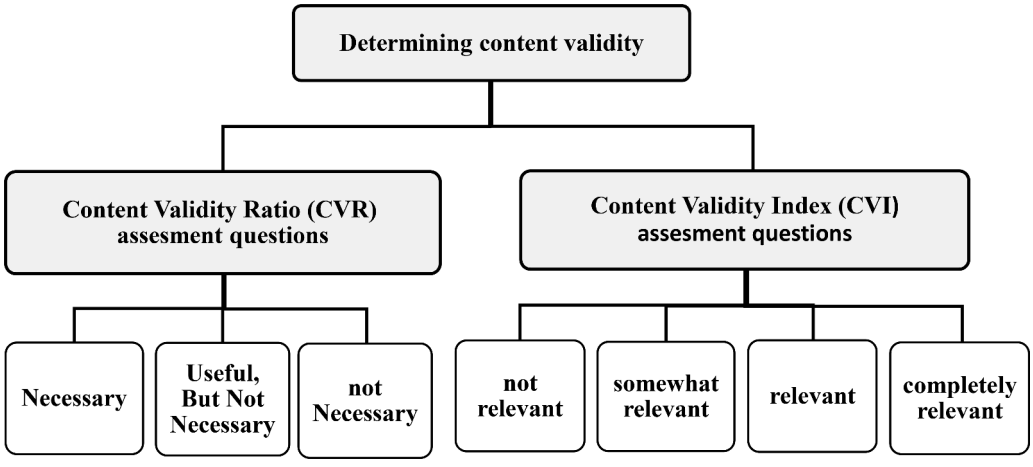

Fig. 1. Assessment questions for determining content validity.

Table 5
Results of ranking of 10 items by 3 rater.

Item	Rater 1	Rater 2	Rater 3	Item CVI	S-CVI (UA)
1	3	3	3	1	#CVI=1/TOTAL ITEMS=7/10=0.7
2	3	3	3	1	
3	3	3	3	1	
4	3	3	3	1	
5	3	3	3	1	
6	3	3	3	1	
7	3	3	3	1	
8	2	3	3	0.67	
9	2	3	3	0.67	
10	3	2	3	0.67	
S-CVI= Mean – I-CVI				**0.90**	

content validity we give them to 3 rates. Results of rating to each item are presented in Table 5. According to both methods of S-CVI calculation we calculated them in following table:

3. Conclusion

Content validity means how relevant the content of a tool to a pre-designed purpose. In short, we have two types of content validity. The first is the content validity ratio, which examines the necessity of each question in the questionnaire. The second is the content validity index, which deals with the relevance of each question in the questionnaire. Researchers in the field of questionnaire studies need to be familiar with how to calculate content validity in order to improve the quality of research in the field of medical sciences.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: none

References

1 Almanasreh E, Moles R, Chen TF. Evaluation of methods used for estimating content validity. *Res Soc Adm Pharm*. 2019;15(2), 214-21.
2 Taherdoost H. Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. How to test the validation of a questionnaire/survey in a research (August 10, 2016). 2016.
3 DeVon HA, Block ME, Moyle-Wright P, et al. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh*. 2007;39(2), 155-64.
4 Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity—Establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—Eliciting concepts for a new PRO instrument. *Value Health*. 2011;14(8), 967-77.
5 Eriksson P, Kovalainen A. *Qualitative Methods in Business research: A practical Guide to Social Research*. Sage; 2015.
6 Tracy SJ. *Qualitative Research Methods: Collecting Evidence, Crafting Analysis, Communicating Impact*. John Wiley & Sons; 2019.
7 Fernández-Gómez E, Martín-Salvador A, Luque-Vara T, et al. Content validation through expert judgement of an instrument on the nutritional knowledge, beliefs, and habits of pregnant women. *Nutrients*. 2020;12(4):1136.
8 Terwee CB, Prinsen CA, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*. 2018;27(5), 1159-70.
9 Almanasreh E, Moles RJ, Chen TF. *A practical approach to the assessment and quantification of content validity*. *Contemporary Research Methods in Pharmacy and Health Services*. Elsevier; 2022, 583-99.
10 Zamanzadeh V, Ghahramanian A, Rassouli M, et al. Design and implementation content validity study: development of an instrument for measuring patient-centered communication. *J Caring Sci*. 2015;4(2):165.
11 Baghestani AR, Ahmadi F, Tanha A, Meshkat M. Bayesian critical values for Lawshe's content validity ratio. *Meas. Eval. Couns. Dev*. 2019;52(1):69–73.
12 Ayre C, Scally AJ. Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation. *Meas. Eval. Couns. Dev*. 2014;47(1):79–86.
13 Wilson FR, Pan W, Schumsky DA. Recalculation of the critical values for Lawshe's content validity ratio. *Meas. Eval. Couns. Dev*. 2012;45(3):197–210.
14 Olsen S. *Instruments For Clinical Health-Care Research*. Jones & Bartlett Learning; 2004.
15 Underwood S.M., editor Instruments for Clinical Health-Care Research. Oncol Nurs Forum; 2005: Oncology Nursing Society.
16 Lawshe CH. A quantitative approach to content validity. *Pers Psychol*. 1975;28(4), 563-75.
17 Koller I, Levenson MR, Glück J. What do you think you are measuring? A mixed-methods procedure for assessing the content validity of test items and theory-based scaling. *Front Psychol*. 2017;126.
18 Zhang L, Basham JD, Carter Jr RA. Measuring personalized learning through the Lens of UDL: development and content validation of a student self-report instrument. *Stud Educ Eval*. 2022;72, 101121.
19 Bull C, Crilly J, Latimer S, Gillespie BM. Establishing the content validity of a new emergency department patient-reported experience measure (ED PREM): a Delphi study. *BMC Emerg Med*. 2022;22(1):1–10.
20 Dolatian M, Ebadi A, Hasanpoor-Azghady SB, NayebiNia AS. Designing and validation of a reproductive health need assessment tool for women experienced domestic violence. *Reprod Health*. 2022;19(1):1–11.
21 Claeys L. *Teacher Motivation to Teach and to Remain Teaching Culturally and Linguistically Diverse Students*. The University of Texas at San Antonio; 2011.
22 Yusoff MSB. ABC of content validation and content validity index calculation. *Resource*. 2019;11(2):49–54.
23 Ab Hamid MR, Mohd Yusof NDB, Buhari SS, Abd Malek K, Md Noor H. Development and validation of educational video content, endorsing dietary adjustments among patients diagnosed with hypertension. *Int J Health Promot Educ*. 2021:1–12.
24 Rodrigues IB, Adachi JD, Beattie KA, MacDermid JC. Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. *BMC Musculoskelet Disord*. 2017;18(1):1–9.
25 Ruhl C, Scheich B, Onokpise B, Bingham D. Content validity testing of the maternal fetal triage index. *J Obstet Gynecol Neonatal Nurs*. 2015;44(6), 701-9.