The 6th Asian Academic Society International Conference (AASIC)
A Transformative Community:
Asia in Dynamism, Innovation, and Globalization

RELIABILITY OF CLINICAL INDICATORS IN NURSING DIAGNOSIS: ACUTE PAIN

Fatin Hapsah Afifah[1] , Intansari Nurjannah[2], Ery Yanuar Akhmad Budi Sunaryo[3]
[1] Bachelor Student 1. School of Nursing, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia
[2] Mental Health and Community Nursing Department, Faculty of Medicine 2. Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta, Indonesia
[3] Emergency and Critical Care Nursing Department 3. Faculty of Medicine, Public Health and Nursing Universitas Gadjah Mada, Yogyakarta, Indonesia
e-mail: fatin.hapsah.a@mail.ugm.ac.id

Abstract

Background: Determine accurate nursing diagnosis based on patient's data is the one of nurse's responsibility. Patient's response in nursing field called clinical indicator. Reliability of clinical indicator research is important to be conducted to help nurse determine accurate nursing diagnosis.
Purpose: The aim of this study is to investigate inter-rater reliability score of clinical indicators in nursing diagnosis: acute pain.
Methods: Respondents of this study is patients undertaking hemodialysis. Two raters assess 30 respondents with instruments based on clinical indicators of acute pain according NANDA-I taxonomy. The data analyzes use Cohen's Kappa.
Results: Ten items of clinical indicator in nursing diagnosis: acute pain was unreliable or low reliability score (<0,40). Otherwise, the number of items with moderate reliability score (0,41-0,60) and high reliability score (0,61-1,00) was 13 items.
Conclusions: More than 50% items of indicator in nursing diagnosis: acute pain had moderate and high reliability score.
Keywords: Clinical Indicators, Inter-Rater Reliability, Nursing Diagnosis, Pain

INTRODUCTION

Nursing diagnosis is the one part of nursing process (Herdman & Kamitsuru, 2014). Accurate nursing diagnosis is description of conformity between diagnostic statements with patient's data (Lunney, 2008). Human's response becomes fundamentals things for establishment nursing diagnosis. Establishment nursing diagnosis based on clinical evidence and clinical indicator is crucial. This is crucial because when nurse identification clinical indicators in patients will determine nursing diagnosis and influence plan of nursing care (Sousa *et al*, 2012). Nurses have responsibility to establish accurate nursing diagnosis for patients.

Identification of accurate nursing diagnosis is something important to do but the process to identification accurate diagnosis is challenge due to humans is unique and complex (Sousa *et al*, 2012). That's why, advanced development ability of critical thinking on establish nursing process needed by nurses (Sousa *et al*, 2012). So, be required instruments to help nurses for identification accurate clinical indicators. NANDA-I or North American Nursing Diagnoses Associations-International is nursing diagnosis used as clinical judgment related acute and potential response on individu, family, group and community in health conditions or life process (Herdman & Kamitsuru, 2014).

NANDA-I provide some diagnosis that can establish by nurses with consider sign and symptoms on patients to establishment accurate diagnosis (Lunney, 2008). Nurses from many countries in the world have been guided on NANDA-I when establishment diagnosis includes nurses in Indonesia.

Nurses will identification clinical indicators on NANDA-I with situation and characteristic of patient's responses that's will influence the hypothesis for specific diagnosis (Lopes et al., 2012). Limitations knowledge about component of nursing diagnosis (definitions, defining characteristics, and related factors or risk factors) influences risk of patient's safety and accuracy on establishment diagnosis (Herdman & Von Krogh, 2012).

Dr. Marjory Gordon in 1982-1987 state about importance identifications clinical indicators on acute or potential nursing diagnosis as a response from individu. However, there is little research about clinical indicators on patients (Lopes et al., 2012).

PURPOSE

The purpose of this study is to investigate reliability of clinical indicators in nursing diagnosis: acute pain

MATERIALS AND METHODS

Design and setting

This is quantitative research with cross sectional design. This study was conducted in the one of Academic Hospital in Yogyakarta, Indonesia.

Participants

Respondents in this reliability test were 30 patients who underwent routine hemodialysis twice in a week at the Hemodialysis Installation at UGM Hospital. Respondents involved in this reliability test were over 18 years old, underwent hemodialysis process for at least 3 months, and patients with routine hemodialysis therapy twice a week, and were able to communicate well. Data collection was conducted on January 20 to January 26 2017 at UGM Hospital.

Recruitment of respondents was done through the head of the Hemodialysis Installation at UGM Hospital. Prospective respondents willing to become respondents were asked to fill out informed consent. Then, the researcher assistant began to conduct interviews using the Visual Analogue Scale instrument to find out whether the respondent experienced pain or not during the last 6 months. If the respondent experienced pain, then the interview and observation was continued with the clinical indicator of nursing diagnosis (NANDA-I) acute pain, but if the respondent did not experience pain then the respondent was excluded. The assistant then conducted interviews and observations of other respondents until the number of respondents reached 30 people.

Measurement

The preparation phase of the research instrument started from the process of nursing diagnosis back translation (NANDA-I) acute pain. It was done after doing the back translation instrument developed based on clinical indicators of nursing diagnosis (NANDA-I) acute pain. Back translation is one strategy that can be used to maintain the similarity of content on an instrument that experiences cross-culture (Cha, Kim, & Erlen, 2007).

Data collection

Data collection techniques used in this study is visual analogue scale and checklist compiled based on clinical indicators from the diagnosis of NANDA-I acute pain. The content in this research is a form of checklist using closed questions by observation and interview. Visual Analogue Scale are used to find out the pain scores on respondents conducted by interview. Data collection is done through two steps:

The first step is data collection to measure content validity

After the instrument preparation process, the researcher tested the validity of the checklist prepared based on clinical indicators of the diagnosis of NANDA-I acute pain. This study uses content validity test using CVI (content validity index) by testing the instrument by an expert. Content validity is used to find out how far an instrument has a sample of the right item to build something that is measured (Polit, Beck, & Owen, 2007). Azwar (2015) in his book states that to test content validity requires the opinion of experts (expert judgment). The number of experts that can be used to test for good content validity is three or more (Polit, Beck, & Owen, 2007).

The expert judgment in this study consists of 3 people, who are 2 experts holding a master's degree and 1 expert were nurses who has more than one year of clinical experience and has a certificate in the clinic area of dialysis installation. Experts assess the level of relevance, accuracy, clarity, ease of understanding, and similarity by giving values of each item 1-4 with assessment criteria score 1 is not relevant, 2 is slightly relevant and requires very significant changes, 3 are relevant but require little change or modification, 4 is very relevant. The value is then calculated using Microsoft Excel to get the value of Content Validity Index for Items (I-CVI) and Content Validity Index for Scale (S-CVI). After the instrument is declared relevant or valid, the reliability testing process is carried out at UGM Hospital in January 2017.

The second step is data collection to measure reliability

This study conducts reliability testing done by two researchers namely researchers and 1 research assistant on January 20 to January 26 2017 at UGM Hospital. The process of collecting reliability test data begins with the perception equation with research assistants. The researcher carried out the same perception of the contents of the questionnaire. The perception / training equation is carried out to equalize the perception between researchers and research assistants on the course of this research so that it is expected to reduce the bias in data collection. The perception equation is carried out twice with the duration of each meeting approximately 2.5 hours. So, the total activity of the perception equation is 5 hours. This is consistent with the opinion of Johnson et al (2008) which states that a minimum of 5 hours of training sessions will be more effective than less than 5 hours. Researchers carried out similarities in perception with research assistants, namely on January 16 and January 19, 2017. In addition, equality of perception was also carried out by learning together with experts. After that, researchers and research assistants conduct observations together on the same object and also use the same instrument.

Data analysis

Data analysis to measure content validity from clinical indicators of nursing diagnosis (NANDA-I) Acute Pain was done using I-CVI. While reliability measured using inter-rater reliability. Inter-rater reliability is an assessment by observing two different people using the same assessment instrument (Kimberlin & Winterstein, 2008). Inter-rater reliability is used because in this study to find out whether there is a clinical indicator of acute pain in hemodialysis patients is done by observation.

RESULTS
Demographic Data
The characteristic of respondents can be seen in Table 1 below:

Table 1. Characteristics of respondents (n: 30)

Characteristics	Mean±SD	Frequency (f)	Percentage(%)
Age	42,30±10,11		
22-28 years old		3	10%
29-35 years old		4	13%
36-42 years old		7	23%
43-49 years old		10	33%
50-56 years old		4	13%
57-63 years old		2	7%
Gender			

Male		16	53%
Female		14	47%
Length of hemodialysis therapy (in year)	2,73±2,54		

Based on the table above, the average respondent's characteristics were 42.30 years old (SD ± 10.11). The majority of respondents in the age range of 43 to 49 years is 33% (f = 10). Meanwhile, the age of respondents with the least number in the age range of 57 years to 63 years is 7% (f = 2).

There were no significant differences between male and female respondents, namely respondents with male sex 53% (f = 16) and female respondents 47% (f = 14).

Respondents involved in testing the validity and reliability of this study on average underwent hemodialysis for 2.73 years (SD ± 2.54). Respondents who were involved in the validity and reliability testing all underwent routine hemodialysis 2x in one week.

Results Calculation: Validity and Reliability

Test results of the content validity on clinical indicators according nursing diagnosis instruments (NANDA-I): acute pain is as follows:

Table 2. The Results of Content Validity Test

		RELEVANCE	
No	**Item**	**Number of Agreement**	**I-CVI**
1	Item 1	3	1
2	Item 2	3	1
3	Item 3a	3	1
4	Item 3b	3	1
5	Item 3c	3	1
6	Item 4a	3	1
7	Item 4b	3	1
8	Item 4c	3	1
9	Item 5	3	1
10	Item 6	3	1
11	Item 7	3	1
12	Item 8	3	1
13	Item 9a	3	1
14	Item 9b	3	1
15	Item 9c	3	1
16	Item 9d	3	1
17	Item 10	3	1
18	Item 11a	3	1
19	Item 11b	3	1
20	Item 11c	3	1
21	Item 11d	3	1
22	Item 12	3	1
23	Item 13	3	1
24	Item 14	3	1
25	Item 15	3	1
26	Item 16	3	1
	S-CVI		1
	Total agreement		26

Validity test was conducted on December 22, 2016 until January 17, 2017. Content validity test obtained the results of the relevance of I-CVI = 1 and S-CVI = 1 which means the results of the validity is very good.

Reliability test instrument for clinical indicators of nursing diagnosis (NANDA-I) Acute Pain was performed at UGM Hospital on January 20 to January 26 2017 at UGM Hospital. Reliability testing was carried out with 30 respondents. The formula used in the inter-rater reliability test is by using Cohen's kappa.

According to Gisev et al (2013) for the nominal or categorical data type with two examiners then using the Cohen's Kappa formula. Here's the formula for Cohen's Kappa according to Hallgren (2012):

$$K = \frac{P(a) - P(e)}{1 - P(e)}$$

Information:
P (a): shows the proportion of the number of examiners / rater who agreed
P (e): shows the proportion of the expected agreement from the assessment

According to Altman (1991) the kappa coefficient value is categorized into several categories including (Altman, 1991 cit McCray, 2013):

Table 3. Altman's (1991) Benchmark Scale for Kappa (McCray, 2013)

Value	Interpretations
<0.20	Poor
0.21-0.40	Fair
0.41-0.60	Moderate
0.60-0.80	Good
0.81-1.00	Very Good

Table 4. The Results of Kappa Reliability Test

Item	Clinical Indicators	Kappa Score
1	Pupil Dilation	-
2a	**Eyes Lack Luster**	**0,571**
2b	Grimace	0,167
2c	Beaten Look	-
3	**Self Focused**	**0,889**
4a	**Restlessness**	**0,925**
4b	**Moaning**	**1,000**
4c	Crying	-
4d	Vigilance	-
5a	**Change in Blood Pressure**	**0,559**
5b	**Change in Heart Rate**	**0,888**
5c	Change in Oxygen Saturation	-
5d	Change in End-Tidal CO_2	*Not measured*
6	Positioning to ease pain	0,210
7	**Protective Behavior**	**0,473**
8	**Guarding Behavior**	**0,651**
9	**Diaphoresis**	**0,685**

No		Kappa
10	**Distraction Behavior**	**1,000**
11a	Difficulty Time Perception	-
11b	**Difficulty Thought Processes**	**0,870**
11c		**0,474**
12	**Difficulty Interaction with People and Environment** **Self-report of Intensity Using Standardized Pain Scale** *(Visual Analogue Scale)*	**1,000**
13	**Self-report of Pain Characteristics Using Standardized Pain Instruments (PQRST Mnemonic)**	**1,000**
14	**Proxy Report of Pain Behavior/Activity Changes (e.g., Family, Member, Caregiver)**	**0,557**
15	**Appetite Change**	**1,000**
16	Hopelessness	-

Information:

(-): The kappa value doesn't come out

Based on the above table, it can be seen that there are several items that have low kappa values (<0.40), namely items 2b and 6. Items that are kappa values come out such as items 1, 2c, 4c, 4d, 5c, 11a, and 16. In addition, there is one item that is item 5d which is due to the limitations of the tool so that it is not measured. As for items that have moderate kappa values (0.41-0.60), items 2a, 5a, 7, 11c, and 14. Items that have a high kappa value (0.61-1.00) are item 3, 4a, 4b, 5b, 8, 9, 10, 11b, 12, and 13.

DISCUSSION

The reliability of an instrument shows the stability and consistency of the results of measuring instruments (Brockopp& Marie, 2000). Inter-rater reliability is an assessment by observing two different people using the same assessment instrument (Kimberlin&Winterstein, 2008).

Table 1 shows the results for the age range of respondents in this study, namely age 43 years to 49 years. According to data from the Indonesian Nephrology Society or PERNEFRI (2014) the age of patients with the most hemodialysis is 45-64 years (61%).

There were no significant differences in male and female respondents. This is in accordance with data from PERNEFRI (2014) the number of patients undergoing hemodialysis with male and female sex did not have a significant difference, 55.77% for male and 44.23% for female.

The average respondents underwent hemodialysis therapy approximately 2.73 years. Whereas in Claxton's study (2010) in which the average respondent had undergone hemodialysis for 4 years. The results of the content validity test were obtained by I-CVI = 1 and S-CVI = 1 which meant that the validity results were very good. S-CVI value can be accepted or said to be valid if the value of S-CVI is 0.8 or more (Davis, 1992 cit. Polit& Beck, 2006).

The reliability test in this study shows that the kappa value each items is different. Chiang et al (2018) in his research explained that there were several factors that could influence the kappa value in inter-rater reliability including training or rater experience, scoring criteria and research design.

CONCLUSION

In this study, the results of items that were not reliable or less reliable were 10 items. While for clinical indicators that have moderate to high reliable values there are 13 clinical indicators.

ACKNOWLEDGMENTS

Thank you for support from School of Nursing, Faculty of Medicine, Community Health and Nursing Universitas Gadjah Mada and also respondents which make this research possible.

REFERENCES

Brockopp, D.Y., & Marie, T.H. (2000). Dasar-dasar riset keperawatan edisi 2. Jakarta: *EGC*.

Cha, E. S., Kim, K. H., & Erlen, J. A. (2007). Translation of scales in cross-cultural research: issues and techniques. *JAN Research Methodology* , 386–395.

Chiang, H.-Y., Lu, W.-S., Yu, W.-H., Hsueh, I.-P., & Hsieh, C.-L. (2018). Inter-rater and Intra-rater reliability of the Balance Computerized Adaptive Test in Patients with Stroke. *American Congress of Rehabilitation Medicine* .

Claxton, R. N., Blackhall, L., Weisbord, S. D., & Holley, J. L. (2010). Undertreatment of Symptoms in Patients on Maintenance Hemodialysis. *Journal of Pain and Symptton Management* , 211-218.

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches and applications. *Elsevier* , 330–338.

Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial . *NIH* , 23–34

Herdman, T. H., & Kamitsuru, S. (2014). NANDA International Nursing Diagnoses: Definitions & Classifications, 2015-2017. Oxford: *Wiley Blackwell*.

Johnson, R., L, Penny, J., A, & Gordon, B. (2008). Assesing Perfomance. New York: *The Guilford Press*.

Kimberlin, C. L., & Winterstein, A. (2008) Validity and reliability of measurement instrumens used in research. *Am J Health-System Pharmacy*, Vol 65.

Lopes, M. V., Silva, V. M., & Araujo, T. L. (2012). Methods for Establishing the Accuracy of Clinical Indicators in Predicting Nursing Diagnoses. *International Journal of Nursing Knowledge* , Vol 23 No 3.

Lunney, M. (2008). Critical Need to Address Accuracy of Nurse's Diagnoses. *OJIN: The Online Journal of Issues in Nursing* Vol. 13 No. 1

McCray, G. (2013), Assessing inter-rater agreement for nominal judgment variables. Paper Presented at the Language Testing Forum. Nottingham: University of Lancaster. Available from: http://www. norbertschmitt.co.uk/uploads/27_528d02015a6da191320524.pdf

PERNEFRI. (2014). Konsesus Dialisis Edisi VII. Jakarta: P*enerbit Perhimpunan Nefrology Indonesia*. http://www.pernefri-inasn.org/index.html diakses pada tanggal 10 Mei 2016.

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Focus on Research Methods : Is the CVI an Acceptable Indicator of Content Validity? Appraisal and Recommendations. *Research in Nursing and Health* .

Sousa, V. E., Lopes, M. V., Araujo, T. L., Rolim, I. L., Nascimento, R. V., & Oliveira, T. F. (2012). Clinical Indicators of Ineffective Airway Clearance for Patients in the Cardiac Postoperative Period. *International Journal of Nursing Knowledge*, 193-200.