کلیدی الفاظ: ارزیابی و توثیق ابزار، محققین، روشنایی، ثبات، ارزیابی، ابزار، اثرات، تحقیق، آزمون، سوالنامه.
INTRODUCTION

Nursing studies have significant role in enhancing nursing professions and improving services quality to patients (1and2). Just as researches in other fields, these studies have regular steps in which the process of measurement is one of the most important components. Applying measurement subtle principles is the essential part of studies and it gains more importance when the aim is to apply study results as a basis for further actions (3and4). In addition to physiological variables, measurement in nursing researches deals with properties such as quality of life, patients’ compliance with the medicinal and therapeutic regime and patients’ satisfaction which are abstract concepts and are called conceptual factors. In these cases, measurement includes activation of these factors in the form of defined variables and the preparation and application of tools or tests in order to measure these variables (5). An important and essential point in this process is emphasis and focus on reducing errors in measuring process so as to trust on the results(6), since the application of its results involves humans which has a delicate situation with its own special ethical and legal issues. In order to reduce errors in measurement, choosing appropriate research tool is of utmost importance (1, 6 and 7).

Selecting the correct and authentic tools result in valid and accurate measurement of intended variables. On the other hand, inappropriate tools cause non-related data collection and in turn, scientific interpretation of research findings (1). In the meantime, the validity and reliability of measurement are key indicators of measurement tools quality. The reliability of a tool refers to its stability during multiple measurements. It also indicates random error rate of measurement methods (2). According to Namali, Yaghmaie (1385) states that reliability indicates the effectiveness of tools and if the tool is not reliable, it causes error in results. Researchers need indices which are reliable and the resultant measures have less error level. Reliable scales enhance the power of study in discovering the differences and relationships in society. Thus it is important to test the reliability of index before studying (6).

Reliability or validity of a tool indicates to what extent the tool assesses the intended concept or factor. According to the principles of America’s psychological association, validity is indicator of appropriateness, significance and usefulness of inferences from a tool scores. The validity of tools is measured in terms of three validity types in theoretical foundation of research: content validity, predictive validity, and factor validity (9, 10, 11, and 12). Each of these three types of validity also has secondary validity, Too much validity is so confusing, since they are related and not independent. There are different views about the type and number of the validity in the questionaire. For example, based on Norick’s sayings, Yaghmaie writes that we must at least make use of content validity or together with factor or predictive validity in designing a research tool(1). Validity is not similar to reliability to be all or nothing; rather it has range and degree (10, 11, and 13). Several studies are reported and printed daily across the world in which researchers do not pay much attention to the accurate measurement of validity and reliability of questionnaires and have unscientific and incomplete reports about the indicators of questionnaires and therefore cause the readers to doubt the findings of study (1), have no trust on the study results and refrain from putting them into practice. This is a major barrier to the implementation of evidence-based practice and nursing research, since determining the accuracy of the collected data is an essential component of evidence-based practice (14). Thus, despite the progress of nursing researchers’ knowledge regarding the importance of validity and reliability and tools authentication in terms of these two indicators, it appears that psychometric tools used in nursing research within the country is not done appropriately. Yaghmaie study results (1385), which were conducted with the aim of psychometric criticism of nursing study tools, show that researchers do not make use of scientific and accurate principles for the validity and reliability of tools(1). Moreover, Kakhki et al. study (1386) indicates that researchers have not paid attention to scientific principles in their researches in order to determine accuracy, sensitivity and machine errors (6). Thus, the present study has the aim of critical investigation of the validity and reliability of used tools in quantitative studies published in Iran’s scientific and research journals and it seeks to answer this question: “how is the quality of validity and reliability assessment methods of used tools in quantitative studies published in Iran’s nursing scientific and research journals?” We can refer to Yaghmaie (1385) and Kakhki (1386) as similar studies done in this field in Iran. However, in Yaghmaie’s study, he evaluated 12 articles of a nursing foreign magazine published during a year.

In Kakhki study, he evaluated the physiological tools of master’s theses in a college which is not thorough and complete. In these studies, methods of assessment are irregular and traditional without adopting a specific procedure while an appropriate assessment that can lead to an accurate judgment is of particular importance (15). In the present study, we make use of Morse critical evaluation which is a regular method and has particular steps for evaluation. These steps include: 10 clear definition of the purpose of investigation 2. Literature search and text organization 3. Identification of important analytical questions 4. Mixing and reporting of the results of critical review (16).

METHODS

The present study is a critical review of contexts that make use of Morse concepts of critical evaluation method (Morse 2000). In the first step, this evaluation, which includes clear and explicit definition of research goal, the overall goal of study is to criticize the used tools’ validity and reliability evaluation method in quantitative studies published in Iran’s nursing scientific magazine. It was designed in two specific, smaller aim forms including the validity evaluation method of used data in quantitative studies published in Iran’s nursing scientific magazine and the reliability evaluation method of used data in qualitative studies published in Iran’s nursing scientific magazine. Following the above-mentioned
issues, two research questions are raised: 1. How is the quality of validity evaluation method of used data in quantitative studies published in Iran’s nursing scientific magazine? 2. How is the quality of the reliability evaluation method of used data in quantitative studies published in Iran’s nursing scientific magazine? In the second step, which deals with literature search and text organization, the researcher surveyed all the published essays in Iran’s 5 nursing scientific magazine within the year 1391- that received the highest rank among Iran’s nursing scientific magazines in 1391 ranking by the secretariat of medical science publications committee. These articles were: Urmia’s Nursing and Midwifery College periodical, Journal of critical care nursing, Iran’s nursing journal, Nursing and Midwifery College journal in Tehran’s Medical Science University and Isfahan’s nursing and midwifery research journal. The researcher omitted essays with quantitative methodology and quantitative essays in which no tools and/ or just physiological ones were used and on the whole, 197 articles were chosen for investigation. In the third step, which discusses identification and design of important analytical questions, the researcher made a list of related questions regarding the validity and reliability of tools after studying texts and articles related to tools’ psychometric and by opinion-seeking from researchers specialized in assessment. In order to prepare this list, nursing research methodology textbooks such as Burns and Grove, Pollt and social science research methodology books were studied. Then, the list was given to 10 professionals in this field including research methodology and tool-making professionals and nursing and midwifery university teachers who have taught this course for many years. Afterwards, by taking into account the opinions of teachers, the final check list was prepared and was rendered to the teachers for content validity verification.

All teachers were asked to present their ideas regarding the mentioned questionnaire questions- in a table provided for checking content validity relative coefficient; the options were non-related, needs serious review, related but needs consideration and completely related. Therefore, questionnaire’s content validity relative coefficient was considered CVI=0.86 after concluding teachers’ opinions. Moreover, a second test was applied for test reliability verification so that 10 articles were evaluated in two time span by the prepared checklist and the correlation coefficient of these two articles’ marks were studied, which was p < 0.001 with r=0.91. Moreover, Cronbach’s alpha index was used for estimating checklist’s internal consistency. Considering that alpha 0.87 was p<0.001, the results show that checklist and all its items have significant and desired consistency. The fourth step, which involves mixing and reporting the critical review report results, we investigate the chosen articles from 5 mentioned magazines by using this checklist, after preparing valid and reliable checklist for studying validity and reliability assessment method of used tools in studies. Then, all the information was entered into spss software — 16th edition and was analyzed by applying analytical and descriptive statistical tests. The level of significance was 0.05 in all statistical tests.

### RESULTS

Of 197 articles from 5 intended magazines, 63 articles was in Urmia’s Nursing and Midwifery College periodical , 36 ones are related to journal of critical care nursing, 27 articles was in Iran’s nursing journal, 29 ones from Nursing and Midwifery College journal in Tehran’s Medical Science University and 29 articles were of Isfahan’s nursing and midwifery research journal. Other information regarding these articles are provided in table 1.

The tools used in 197 articles were 280 including 245 (%87.5) questionnaires and 35 (%12.5) checklists. Other information regarding these tools is provided in table 2.

Moreover, from 180 foreign tools used in study, in 38 cased (%21.1), normalization test in Iran was mentioned and there was no discussion regarding CVI= 0.86 after concluding teachers’ opinions. Moreover, a second test was applied for test reliability verification so that 10 articles were evaluated in two time span by the prepared checklist and the correlation coefficient of these two articles’ marks were studied, which was p < 0.001 with r=0.91. Moreover, Cronbach’s alpha index was used for estimating checklist’s internal consistency. Considering that alpha 0.87 was p<0.001, the results show that checklist and all its items have significant and desired consistency. The fourth step, which involves mixing and reporting the critical review report results, we investigate the chosen articles from 5 mentioned magazines by using this checklist, after preparing valid and reliable checklist for studying validity and reliability assessment method of used tools in studies. Then, all the information was entered into spss software — 16th edition and was analyzed by applying analytical and descriptive statistical tests. The level of significance was 0.05 in all statistical tests.

### Table 1. The status of articles from author’s scientific level, study’s field and kind point of view

| Author’s scientific level (doctorate/ M.S/ B.S) | 5(%1.8)/97(%49.2)/95(%48.2)/113(%57.4)/42(%21.3)/42(%21.3) |
| Type of study (experimental/ semi-experimental/ analytical descriptive) | №4(%27.4)/45(%22.8)/9(%4.6)/19(%9.6)/70(%35.5) |
| Field of study (clinical/ educational/ professional/ management/ health improvement) | 54(%27.4)/45(%22.8)/9(%4.6)/19(%9.6)/70(%35.5) |

| Tool source (researcher-made/ existing tools/ adjusted tools) | 5(%1.8)/175(%62.5)/100(%35.7)/50(%17.9)/230(%82.1) |
| Mentioning tools dimension in the article (Yes/ No) | 64(%22.9)/216(%77.1) |
| Mentioning tools scoring method in article (Yes/ No) | 68(%24.3)/212(%57.7) |
| Mentioning the effectiveness of tools in article (Yes/ No) | 68(%24.3)/212(%57.7) |
The most applicable method for validity verification of tools was content validity method 148(52.85) among which only in 7(4.7) cases content validity index was mentioned. Furthermore, the most applicable method for reliability verification of tools was internal consistency assessment method using Cronbach's alpha test 132(47.11); among which only in 2(1.5) cases level of significance were mentioned.

**DISCUSSION**

The present study results regarding the first specific goal of the present study, which was criticizing the validity and reliability assessment methods of used tools in quantitative studies published in Iran's nursing scientific and research journals, shows that there is no hint to the validity and reliability of tool's original version in the major investigated foreign tools (60%) while the existence of validity and reliability in the original version is an obvious matter (17). Moreover, according to the results of this study, tools content validity authentication method were not mentioned in 52.9% cases and none of validity verification methods were used while according to Burns and Grove, validity must be tested in all studies because the validity differs from one case and situation to the other; in fact, this test relates to the validation of using an assessment tool for a specific group or goal and not for the validation of the tool itself (10). A tool may be valid in a specific situation but not in the other one. According to the results of this study, there was not discussion regarding the validity authentication method in 29% of these articles, from 100 researcher-made tools in these articles and also tools source construction in 9% cases were not specified. Polight mentions that those who intend to develop new tools must start from factor conceptualization so that measurement encompasses the whole domain sufficiently (11). Nursing has extended to other fields so their conceptualization probably results from a quantitative validation of using an assessment tool for a specific group or goal. Although Yaghmaie's study (1385) is secondary type of studies, Mohamadbeigi (17) (1393) performed studies abroad in this field such as Yaghmaie's study results or from surveying the texts. The results of this study correspond to Yaghmaie's study results (1385). Yaghmaie, who discusses and criticize articles published in the journal of advanced nursing in the year 2001, states that in the surveyed articles content validity lacks applying scientific and accurate principles in most cases and just 2 articles makes use of content validity index from 12 ones. Both articles make use of different method for content validity and the number of professionals for content validity designation was also different (1). In the present study, content validity designation was not performed at all and in only 7 tool cases (4.7) content validity index was mentioned- in the tools had content validity designation. The advantage of the present study to Yaghmaie research is that more articles (19 against 19) were studied from several Iran's nursing magazines (5 against 1). Regard the second specific goal, which was criticizing the validity and reliability assessment methods of used tools in quantitative studies published in Iran's nursing scientific and research journals, results indicate undesirable quality of reliability evaluation of tools so that in 89(31.8) cases no method of reliability evaluation of tools was used, while the most applicable method for reliability verification of tools was internal consistency assessment method using Cronbach’s alpha test and only in 2 cases (1.5) alpha’s level of significance was specified. Whereas Burns and Grove considered tool reliability essential before studying and stated that reliability estimation is performed regarding the tested statistical case and thus high reliability in a statistical case does not mean that it is the same in another statistical case (10). Therefore, we must have scale reliability test in each statistical analysis and must report the reliability level. Although Yaghmaie's study (1385) was also indicative of insufficiency in reliability assessment method of essays tools- so that from 12 case study only 5 cases make use of internal consistency for reliability measurement(1)- the quality of reliability assessment method in much lower in the present study since in Yaghmaie's study, more than 50% of essays (7 from 12 cases) makes use of more than one method for reliability authentication while in the present study only in 23 cases from 290 tool cases) made use of two reliability assessment method. The results of the present study also correspond to Barvishpoor Kakhki et al, (1396) (6). His study results is also indicative of researchers’ lack of attention to accuracy and authenticity of tools used in the study. The difference of the present study with the mentioned researchers’ ones is that in Barvishpoor Kakhki et al study, physiological tools were used and investigated while in the focus of the present study is on paper tools and is test type. It is worth mentioning that the researcher found nothing related to the present study while reviewing the extensive texts in valid scientific sites, except Yaghmari and Barvishpoor Kakhki researches. Thus the performed studies in this field are limited to these two cases. Other related studies were simply reviews which deal with including Yaghmari (2) (1382), Mohamadbeigi (17) (1393). Performed studies abroad in this field such as Kimbrell (2009) (5) are secondary type of studies.

Results of the present study indicate insufficiency in reliability and validity assessment method of used tools in Iran’s valid magazines. Since the implementation of these study results is based on high validity of studies and since...
Validity and reliability of research tools

Validity and reliability authentication of used tools in study has a key role in gaining results and accurate findings, then the results of this research can be an incentive for nursing researchers and those involved in searching and education so as to take an important step toward solving this problem and by taking into account the tools psychometrics, being equipped with it and put it into practice, we paved the way for greater evidence-based care.

Since the present study had no other alternatives than making use of restated reports in this field in the studied articles in order to survey the quality of evaluation method in used tools in studies, then the trust on these reports was one of the limitations of the present study.

ACKNOWLEDGEMENT

We appreciate their supports and help especially their financial support. Moreover, we would like to express our gratitude to Mrs. Simin Sharafi (M.S. of nursing) for her earnest help.

Research committee approval and financial support:
This article refers to approved research project in vice presidency research and Research Ethics Committee of Mashhad Medical Science University (research code: 921593).

Conflict of interest: The authors declare no conflict of interest.

REFERENCES

17. Mohammadbeigi A, Mohamadsalehi N, Aligol M. Validity and reliability of the instruments and types of measurement in health applied researches. Journal of Rafsanjan University of Medical Sciences 2015; 13(10): 1153-70. [In Persian].