Psychometric properties of clinical competency assessment instruments for psychiatric nurses: A systematic review of literature

Effat Sheikhbahaeddinzadeh[1], Tahereh Ashktorab[1], Abbas Ebadi[2,3]

Abstract:

BACKGROUND: Facilitating the healing process of patients with psychiatric disorders depends on high-quality mental health care and expert psychiatric nurses. A valid tool is required to objectively evaluate the quality of performance and competency of psychiatric nurses. This systematic review aimed to investigate and critique the psychometric properties of some psychiatric nursing clinical competency assessment instruments with Consensus-based standards for the selection of health status measurement instruments (COSMIN) checklist and according to Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA).

MATERIALS AND METHODS: To retrieve published articles using Persian and English keywords "psychiatric nursing," "competence," "competency," "tool," "checklist," "scale," "questionnaire," "psychiatric mental health nurse," were searched in databases without time limitation. Then, psychometric properties of selected instruments were evaluated using the COSMIN checklist and reported according to the PRISMA statement.

RESULTS: Most of these tools did not entirely and desirably report psychometric properties. It is not designed as a special tool for postgraduate psychiatric nursing.

CONCLUSION: The methodology of existing instruments does not meet the COSMIN checklist criteria; therefore it needs to develop. To assess the competency of psychiatric nursing postgraduate, a tool tailored to the cultural and social context and with acceptable psychometric properties is necessary.

Keywords:

Checklist, competency, instrument, psychiatric nursing, psychometrics, systematic review

[1]Department of Nursing Management, Faculty of Nursing and Midwifery, Tehran Medical Science Islamic Azad University, Tehran, Iran, [2]Behavioral Sciences Research Center, Life Style Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran, [3]Department of Management, Nursing Faculty, Baqiyatallah University of Medical Sciences, Tehran, Iran

Address for correspondence:
Dr. Tahereh Ashktorab, Department of Nursing Management, Faculty of Nursing and Midwifery, Tehran Medical Science Islamic Azad University, Tehran, Iran.
E-mail: taherehashk@gmail.com

Introduction

"Competency" is the nurse's ability to integrate knowledge, skills, judgment, and professional characteristics to act safely and ethically in accordance with the nursing scope,[1] and in specific clinical situations.[2] The clinical competency of psychiatric nurses is to integrate carefully and continuously therapeutic communication with technical skills, knowledge, clinical reasoning, emotions, and values in the context of psychiatric nursing interventions.[3] Therefore,

psychiatric nurses are expected to improve the ability of criticism, apply evidence-based interventions and develop diagnostic reasoning skills to support and confirm clinical decisions.[4,5]

Growing changes in the last few decades have affected the health condition,[6-9] has increased the expectation of nurses to carry out professional responsibilities, support and provide continuous, high-quality, and safe care.[7,10] Awareness of clinical practice leads to receiving feedback and increase awareness of strengths and weaknesses of performance in clinical settings. In

this way, more efforts will be made to acquire the necessary skills and competencies before entering clinical environments.[11,12] However, nursing educators and managers face challenges to assess clinical competency. Traditional clinical evaluation relies on direct observation. It may lead to subjective evaluation, bias, and limited reliability. The evaluation may be risked due to the lack of standardization and investigation of a limited number of clinical problems.[13] Furthermore, traditional clinical examination tools often do not have psychometric properties.[14,15] As a result, the validity and reliability of traditional clinical studies are questionable.

Some researchers have examined nurses` competency for caring of people with psychiatric disorders, including Stockman *et al*., Robbins and Hoke.[4,13] However, with the best knowledge of researchers, limited studies have examined the competency of postgraduate psychiatric nurses to care for people with psychiatric disorders. Nowadays, the selection of appropriate and relevant measurement tools is emphasized, which is as important as the research and the science production.[16,17] Therefore, before using a tool, it is necessary to evaluate its psychometric properties based on appropriate criteria.[18] Studies have shown that some designed tools are not accepted by users and cannot be used in many situations or there is no agreement about their terms, definitions, and psychometric properties.[16] Hence, the question arises whether the psychometric properties of existing tools on examining postgraduate psychiatric nurse clinical competency are approved based on the Consensus-Based Standards for the Selection of Health Status Measurement Instruments "(COSMIN)" comprehensive checklist. A valid and reliable instrument helps teachers and managers to assess the postgraduate psychiatric nurses' competency objectively. COSMIN checklist contains the list of required design standards and preferred statistical methods is a useful tool for evaluating the quality of measurement tools.[19] The present systematic review has been conducted with the aim of introducing psychometric properties of psychiatric nurses' clinical competency assessment tools based on the COSMIN comprehensive checklist.

Materials and Methods

The current study is a systematic review that reviews articles related to psychiatric nurses" clinical competency assessment tools and evaluates them based on the COSMIN checklist. This study was conducted using two instructions:

1. The use of Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement which has been developed by a group of epidemiologists and editors of medical journals to improve the quality of research papers.[20-22] These guidelines, in addition to helping researchers to write manuscripts, can also help reviewers to evaluate articles[20,23]

2. COSMIN: This checklist was designed by Mokkink *et al*. using the Delphi method to assess the methodological quality of studies on measurement properties.[24,25]

This systematic review was performed in 2021, by searching Persian and English keywords "Psychiatric nurse," "Psychiatric mental health nurse," "Competence," "Competency," "Tool," "Checklist," "Scale," "Questionnaire," "Instruments," based on MeSH in PubMed, PsycINFO, Cochrane library, and Google scholar databases without time limitation.

The search strategy was: ((("psychiatric nursing"[Title/Abstract]) OR ("psychiatric mental health nurse"[Title/Abstract])) AND ("clinical competency"[Title/Abstract]) OR ("clinical competence"[Title/Abstract]) AND (instrument [Title/Abstract]) OR (tool[Title/Abstract])) OR (scale[Title/Abstract]) OR (questionnaire[Title/Abstract])).

Methodology and writing of the review study were conducted using PRISMA statement.[20] Published articles in English and Persian were retrieved.

The first and second authors of the study independently reviewed the abstracts and titles of the studies based on inclusion criteria. Then, the full text of the selected articles was reviewed by two researchers. The difference of opinions was discussed until the consensus. If necessary, the third author was asked.

After completing the search, duplicate articles were deleted using Mendeley software. In the sampling phase, the research team selected all studies that reported or applied psychiatric nursing clinical competency tools.

Inclusion criteria

Articles which met the following criteria were reviewed: (1) Peer-reviewed articles, (2) Articles published in English and Persian journals, (3) Access to the full text of the article, (4) Having validity or reliability.

Exclusion criteria

(1) Articles without psychometric properties, (2) Articles in other languages, (3) Duplicate articles, (4) Conference papers, book chapters, dissertations, and thesis.

The paper selection process was carried out in several steps [Figure 1]. There were 465 articles in the initial search. After deleting duplicate articles, the number of articles decreased to 449 articles. In studies reviewing, only articles related to the design, and psychometric evaluation of tools, or the use of the psychiatric
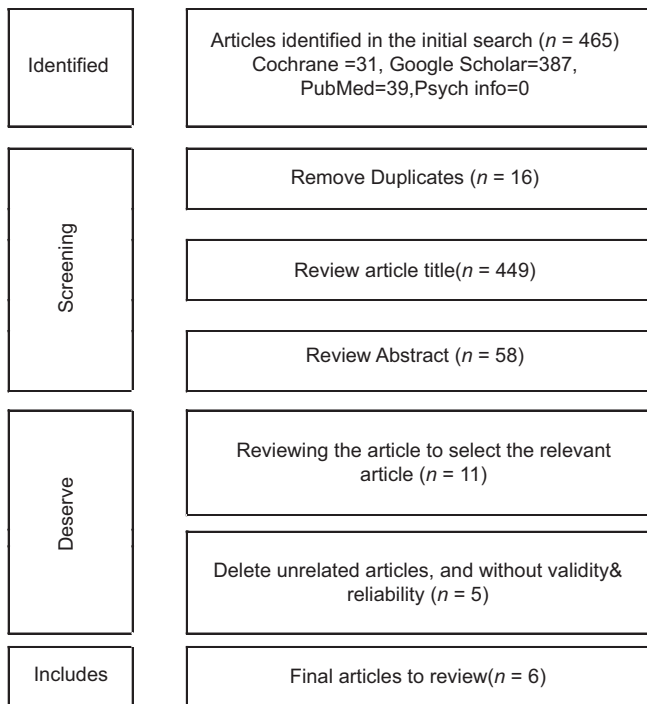
Figure 1: Process of Checking, Screening and Searching

nurse clinical competency assessment tools were preserved (*n* = 11). Finally, after reviewing the full text of articles and deleting articles without validity and reliability, 6 articles were reviewed.

At first, the full text of 6 articles was reviewed. Data were extracted according to the following topics: The purpose of the study, the target population, the main findings, and the instrument characteristics [Table 1].[13,26-30]

Data quality assessment

The methodological quality of the studied tools was evaluated based on the version 2020 COSMIN checklist [Table 2][31] and their psychometric properties were evaluated. COSMIN checklist consists of 12 boxes in 4 domains as follows: "Validity," "Reliability," "Responsiveness," "Generalization and interpretability." The validity includes three domains of content validity (contains in face validity), construct validity (hypothesis test, structure validity, and cross-cultural validity (in translated tools)), and criterion validity. Reliability domain is: three characteristics of internal consistency, reliability (i.e., test-retest, inter

Table 1: Summary of the final articles review

Authors	Year	The purpose of study	Target group	Main findings and instrument characteristics
Bondy *et al.*	1997	The development and testing of the PsychNPAI of the	Baccalaureate students exposed to the psychiatric population	80 competencies were categorized into six subscales Basic knowledge/critical thinking, nursing process, nursing interventions, communication skills, professional socialization behaviors, self-evaluation[26]
Mohtashami *et al.*	2014	Developing and psychometric properties check list of clinical competency in mental health nursing students	Undergraduate students having the psychiatric nursing internship	The 73-item check list with CVR=0.83, CVI=0.98. and Cronbach's alpha for internal consistency=0.98, inter-observer correlation=0.70 includes general and specific competencies[27]
Moskoei *et al.*	2017	Assessing the competence of mental health nurses	Nurses working in psychiatric wards of hospitals	The 45-item questionnaire included 2 factors emotional/moral and special care competencies. CVR=0.88, CVI=0.97, Cronbach's alpha=0.98, ICC=0.98[28]
Chen *et al.*	2018	Developing and psychometric testing scale for assessing the competence of psychiatric nursing case management	Nurses working in psychiatric wards (Diploma, Bachelor, Master)	The 18-item scale in which essential activities of case care were classified into 6 groups: Processes and Services of case care, use of resources and management, social and economic psychological support, rehabilitation activity, effectiveness of assessment and ethics and law. EFA defined 2 factors. Item total correlation ranged from. 78 to. 90. The inter item correlations within the scale and between the two dimensions ranged from 0.60 to 0.88, and 0.55 to 0.79, respectively. The test-retest reliability for 30 participants, with 2-week interval was between 0.90 and 0.92[29]
Feng *et al.*	2018	To explore the competencies needed to provide nursing care for people with psychiatric disabilities in a hospital environment	Nurses working in psychiatric wards of hospitals	The 17-item questionnaire included 5 factors, which accounted for 68.60% of the total variance: Sense of responsibility, vocational identification, agreeableness, cooperation capacity, and carefulness; (α=0.85, 0.85, 0.74, 0.80, 0.77, respectively). The competencies were on attitudes, values, and traits[30]
Stockman *et al.*	2019	To explore the psychometric properties of a mental health OSCE for nursing students	Undergraduate mental health nursing students	Preliminary validity and reliability measures such as content and criterion validity and reliability were performed among evaluators. Content validity was done by 3 specialists. I-CVI=1, S-CVI=1.13 Stations were established with the subject of mental state, therapeutic communication and nurse relationship[13]

OSCE=Objective structured clinical evaluation, ICC=Intra class correlation, CVR=Content validity ratio, CVI=Content validity index, EFA=Exploratory factor analysis

Table 2: Summary of psychometric properties of measurement tools for psychiatric nursing clinical competency

Tool	Years	Content validity	Structure validity	Criterion validity	Test hypothesis	Cross cultural	Inner consistency	Reliability	Measurement error	Responsibility
PsychNPAI	1997	+ (the lack of κ*)	NR	NR	NR	NR	+	?	NR	NR
Clinical competency in mental health nursing students	2014	+ (the lack of κ*)	NR	NR	NR	NR	+	+	NR	NR
CCMHN	2017	+ (the lack of κ*)	+	NR	NR	NR	+	+	NR	NR
Case management competence scale	2018	? (the lack of CVR, κ*)	± EFA with PCA	NR	NR	NR	+	+	NR	NR
MHOSCE	2018	?	NE	?	NR	NR	NE	+	NR	NR
Questionnaire to assess nursing competencies for the care of people with psychiatric disabilities in a hospital environment	2019	?	+	NR	NR	NR	+	NE	NR	NR

NE=Not evaluated, based on Terwee *et al.* rankings, (2007), κ*=Modified Kappa, +=Positive, ?=Indeterminate rating, ±=Conflicting data, NR=Not reported, MHOSCE=Mental health objective structured clinical examination, CCMHN=Clinical competent mental health nurse, CVR=Content validity ratia

and intra rater agreement), and measurement error.[31-34] Other psychometric properties include reproducibility, floor and ceiling effect.[35]

In the field of responsiveness, only one responsiveness feature is measured in both internal and external categories.[16,19] Responsiveness is the sensitivity and ability to respond to changes.[16,35]

The last one is devoted to generalization and interpretability, which is the rate of the qualitative significance of the minimum significant changes in the instrument score.[16,35] Interpretability is not a measured property, but it is an important characteristic.[31]

The methodological quality of each study should be evaluated separately by ranking all COSMIN checklist standards. Therefore, the COSMIN checklist should be used as a modular tool. This means that when evaluating the quality of the tool described in an article, there may be no need to complete the entire checklist.[31] However based on COSMIN, the measurements the article are evaluated, and are determined which them need to be completed.[24]

In this study, the ranking was done based on Terwee *et al.* method,[36] Each item is graded according to "Poor," "Fair," "Good", and "Excellent" grading system. Cordier *et al.* scoring method was used to better understand the psychometric properties of the instruments [Table 3].[37] The results were presented as percentage (weak = 0.25%, fair = 25.1%–50%, good = 50.1%–75%, excellent = 75.1%–100%). Cases which did not meet all of the COSMIN criteria, the following formula was used to calculate the total score for each psychometric property to more accurately obtain the quality of the psychometric properties.

The total score for the psychometric property:[37]

$$\frac{\text{Total score obtained} - \text{Minimal possible score)}}{\text{Maximum score possible} - \text{Minimum score possible)}} \times 100$$

Results

All the articles were mixed method. Half of the articles (50%, *n* = 3) were between 2018 and 2019, 33.3% 2014–2015 and one article in 1997. Half of the study tools (*n* = 3, 50%) are for assessing the clinical competency of nurses working in the psychiatric ward, and 50% assess the clinical competency of undergraduate nursing students. Two instruments (33.3%) were self-evaluated[29,30] and the rest were performed by the observer. Quantitative face validity, and modified kappa statistic declining raters agreement by the chance,

Table 3: Psychometric quality ranking criteria adapted from Terwee *et al.* (2007)[36,37]

Psychometric property	Score	Quality criteria
Content validity	+	A clear description is provided of the measurement aim, the target population, the concepts that are being measured, and the item selection AND target population and (investigators OR experts) were involved in item selection
	?	A clear description of above-mentioned aspects is lacking OR only target population involved OR doubtful design or method
	–	No target population involvement
	±	Conflicting results
	NR	No information found on target population involvement
	NE	Not evaluated
Structural validity	+	Factors should explain at least 50% of the variance
	?	Explained variance not mentioned
	–	Factors explain <50% of the variance
	±	Conflicting results
	NR	No information found on structural validity
	NE	Not evaluated
Hypothesis testing	+	Specific hypotheses were formulated AND at least 75% of the results are in accordance with these hypotheses
	?	Doubtful design or method (e.g., no hypotheses)
	–	<75% of hypotheses were confirmed, despite adequate design and methods
	±	Conflicting results between studies within the same manual
	NR	No information found on hypotheses testing
	NE	Not evaluated
Criterion validity	+	Convincing arguments that gold standard is "gold" AND correlation with gold standard 0.70
	?	No convincing arguments that gold standard is "gold" OR doubtful design or method
	–	Correlation with gold standard <0.70 despite adequate design and method
	±	Conflicting results
	NR	No information found on criterion validity
	NE	Not evaluated
Internal consistency	+	Factor analyses performed on adequate sample size (7* items and _100) AND Cronbach's alpha(s) calculated per dimension AND Cronbach's alpha(s) between 0.70 and 0.95
	?	No factor analysis OR doubtful design or method
	–	Cronbach's alpha(s) <0.70 or >0.95, despite adequate design and method
	±	Conflicting results
	NR	No information found on internal consistency
	NE	Not evaluated
Reliability	+	ICC or weighted κ 0.70
	?	Doubtful design or method (e.g., time interval not mentioned)
	–	ICC or weighted κ <0.70, despite adequate design and method
	±	Conflicting results
	NR	No information found on reliability
	NE	Not evaluated
Measurement error	+	MIC < SDC OR MIC outside the LOA OR convincing arguments that agreement is acceptable
	?	Doubtful design or method OR (MIC not defined AND no convincing arguments that agreement is acceptable)
	–	MIC_SDC OR MIC equals or inside LOA, despite adequate design and method;
	±	Conflicting results
	NR	No information found on measurement error
	NE	Not evaluated

Scores=+=Positive rating, ?=Indeterminate rating, –=Negative rating, ±=Conflicting data, NR=Not reported, NE=Not evaluated (for study of poor methodological quality according to COSMIN rating, data are excluded from further evaluation). Doubtful design or method is assigned when a clear description of the design or methods of the study is lacking, sample size smaller than 50 subjects (should be at least 50 in every subgroup analysis), or any important methodological weakness in the design or execution of the study. Hypothesis testing: All correlations should be statistically significant (if not, these hypotheses are not confirmed) and these correlations should be at least moderate ($r > 0.5$). Measurement error: MIC (Minimal important change), SDC=Smallest detectable change, LOA=Limits of agreement, ICC=Intra class correlation, SD=Standard deviation

were not conducted in any article studied. Similiarity, measurement error, sensitivity, and responsiveness were not investigated. Except for Chen *et al.* study, other articles did not mention the time interval between the two evaluations in the reliability.[29] Only Stockman *et al.*[13] presented criterion validity, but the questionnaire used (Pierr *et al.* questionnaire 2004) as criterion had no validity and reliability.

Table 3 shows the ranking summary of psychometric properties of studies based on the qualitative criteria of the COSMIN checklist and the scoring method proposed by Terwee *et al*. table.[37]

Table 4 is the psychometric properties of the instruments studied based on the ranking criteria of Terwee *et al*.[36]

Discussion

In this study, the psychometric properties of psychiatric nurse clinical competency assessment tools were investigated. The results showed that several tools have been designed to measure the competence of nurses working in psychiatric ward or the competence of undergraduate mental health nursing or nursing students. However, none of them check the postgraduate psychiatric nursing clinical competency. As knowledge about clinical competency in postgraduate education is limited,[13] no study was found on the design and psychometrics of MSc psychiatric nursing clinical competency assessment tool. Merely there were 6 tools in the field of psychiatric nursing cares. To design or application a tool, paying attention to its psychometric quality is very important.[38] All tool developers must report the validity and reliability of their instruments to make it possible to use, as well as to develop that tool.[39] The four main psychometric properties of tools are validity, reliability, and ability to respond to changes and interpretability.[40] However, according to Tables 3 and 4, it seems that the existing instruments based on COSMIN checklist criteria have not fully investigated or reported psychometric properties. Quantitative face validity, modified kappa statistics, and measurement errors were not calculated in none of the studies reviewed. Furthermore, there were no reports about missing value and management of it, responsibility and interpretability. Except for Stockman *et al*. (2019) article, the criterion validity has not been performed in other articles. In this article, the questionnaire used was not approved as a "golden standard" because its validity and reliability were not measured.

Content validity is considered the most important property of measurement because it should be clear that the items of the instrument are relevant, comprehensive, and understandable, taking into account the structure and population studied.[41] Face validity (as part of content validity) is determined as the amount of understanding that the construct evaluated provides from the content of the questionnaire.[25] In content validity section, the target population in three tools[27-29] has no clear description. For example, in the "Checklist of Clinical Competency in Mental Health Nursing Students," the undergraduate mental health nursing student was defined as the target population. Such educational course is not held in

Table 4: Methodological quality of psychiatric nursing clinical competency tool according to scoring Terwee *et al.* (2007)									
Author(s)	Year	Internal consistency	Reliability	Measurement error	Content validity	Structural validity	Hypotheses testing	Cross cultural validity	Criterion validity
Bondy *et al.*	1997	Fair (45.45)	?	NR	Good (63.15)	NR	NR	NR	NR
Mohtashami *et al.*	2014	Good (54.54)	Good (60.00)	NR	Fair (47.36)	NR	NR	NR	NR
Moskoei *et al.*	2017	Excellent (81.8)	Good (73.68)	NR	Good (57/14)	Excellent (88.88)	NR	NR	NR
Chen *et al.*	2018	Good (72.72)	Excellent (84.21)	NR	Fair (47.36)	±	NR	NR	NR
Feng *et al.*	2018	Excellent (90.00)	NE	NR	?	Excellent (77.77)	NR	NR	NR
Stockman *et al.*	2019	NE	Good (73.68)	NR	?	NE	NR	NR	?

The studies quality which evaluated each instrument' psychometric properties were evaluated according to the COSMIN rating. Four-point scale was used (poor=1, fair=2, good=3, excellent=4) and the outcome was presented as percentage of rating (0%-25.0%=poor, 25.1%-50%=fair, 50.1%-75%=good, 75.1%-100%=excellent). NR=Not reported, NE=Not evaluated, ?=Indeterminate rating

Iran, and the target population of researchers was the undergraduate nursing students after passing mental health internship. The modified kappa statistic is a new approach to eliminate evaluators' random agreement on how items are related.[41-44] However, the articles in the study had no modified kappa statistics at all.

The instruments are categorized into formative and reflective types. The items and variables of reflective instruments are interdependent and are usually self-evaluated, i.e., the patient or person studied response. Considering that in reflective instruments correlations between items and variables are investigated, it is important to calculate Cronbach's alpha in these instruments.[34] In this study, two tools of "Nursing Competencies for the Care of People with Psychiatric Disabilities".and "A Nurse Case Management Competence Scale" are self-evaluated which were performed by nurses of psychiatric settings, and internal stability was reported >0.7.

Formative tools are more objective and dependent on the observer. In these instruments, performing factor analysis does not make sense.[34] Construct validity consists of "structure validity," "cross-cultural validity" and "hypothesis test." In the construct validity section, description of these components is important. Cross-cultural is related to the translation of tools. If article is not translated it is not necessary.[31] The two "CCMHN" and "Checklist of Clinical Competency in Mental Health Nursing Students" tools are suitable for the cultural context of Iran. Although, researchers did not find anything regarding cultural competency, which is an essential part of providing effective and cultural services to each client individually. Since the meaning of a concept varies from one culture to another, tool-developing experts believe that the tool should be directly due to the target population and suitable for the culture and lifestyle of the country consuming. This should be considered at the time of the development of the items.[19] So, the Iranian postgraduate psychiatric nurse competency should be examined by the tools designed for the same culture and context.

Because the instruments under study are as checklist or observational and formative, the best method of assessing the construct validity of these instruments is "known group validity."[44,45] For example, evaluation of nurses' clinical competency in general wards compared to psychiatric wards. In the "PsychNPAI" tool, Bondy *et al*.[26] has been compared competence between the two groups of freshman and senior nursing students, and in "Case Management Competence Scale" has been done between the two groups taught and without training about case management. "Objective Structured Mental Health Clinical Examination (MHOSCE")" tool and "Checklist of Clinical Competency in Mental Health Nursing Students" have no construct validity, but three tools[28,29,30] have factor analysis. Chen *et al*.[29] used principal component analysis to extract factors instead of the maximum likelihood method, though. The results of this study showed that most studies have not evaluated the basic characteristics of construct validity based on COSMIN taxonomy.

Reliability (internal consistency, stability, and measurement error) is also an important part of the COSMIN checklist. Since the tools of this study are observational and formative, the inter-rater and intra-rater reliability is very important.[33] The reliability between the observers refers to the level of agreement between different observers in scoring or judging a situation or behavior, by a same tool and within the same predetermined criteria.[16] The observational instruments "MHOSCE," "CCMHN," "Checklist of Clinical Competency in Mental Health Nursing Students" and "PsychNPAI" reported correlation coefficients above 0.7; although, none mentioned the time interval between the two observations. In the "PsychNPAI" tool, it is not known which type of correlations (inter and intra rater) have been investigated. The reliability of self-evaluate tools "A Nurse Case Management Competence Scale" and "Nursing Competencies for the Care of People with Psychiatric Disabilities in a Hospital Environment" were supported with retest coefficient (≥ 0.90).[29,30] Except for brief point in Bondy *et al.* (1997)[26] study, in any of the articles under study has not been reported measurement error, which is a systematic and random error of patient score, and is not attributed to actual changes in the measured construct.[41,43]

Responsiveness is tool ability for detecting and measuring construct changes over time.[31] The effect of ceiling and floor is when more than 15% of respondents have the highest or lowest achievable score, respectively. The presence of ceiling and floor effects indicates that the items related to the minimum and maximum severity of the phenomenon are probably not included in the questionnaire. It indicates insufficient content validity.[42] Although these criteria for COSMIN checklist are the main foundations for investigating the application of the instrument, it has not been reported in any of the studied instruments. As a result, it is suggested to develop these tools.

Furthermore, in reviewing other characteristics of each tool, it can be noted separately including: two tools "CCMHN" and "Checklist of Clinical Competency in Mental Health Nursing Students" are observational, and have a high number of evaluators for content validity which are the strengths of them. As these tools assess the expected competence for nurses working in

the psychiatric ward and the undergraduate nursing students, therefore they are not suitable for assessing postgraduate psychiatric nurses' clinical competency. In these tools, expected abilities which are stated in the postgraduate psychiatric nursing curriculum and approved by the nursing intervention classification system such as reminiscence, humor therapy, reality orientation therapy, play therapy, family therapy, group therapy, psychiatric nursing counseling, etc., have not been investigated.

"Nursing Competencies for the Care of People with Psychiatric Disabilities tool" pays attention to the attitude, value, and characteristics of psychiatric nurses, which has been less noticed in most studies.[44]

Since objective structured clinical evaluation (OSCE) is a simple method of competency measurement, it is widely used in medicine[45] in fact due to its validity and reliability of OSCE, it is a clinical evaluation objectively. However, in MHOSCE, validity was done with the minimum number of samples which was three specialists. However, with increasing the number of evaluators, the probability of random agreement decreases.[46] Furthermore, internal consistency was not evaluated. Studies have suggested gender as one of the factors affecting the student's caring behaviors and clinical competency,[47] while in Stockman *et al*. (2019) study sampling was done only on female students.

One of the advantages of "Psychiatric Nurses of a Nurse Case Management Competence" tool is to assess the competence of case management and emphasis on obtaining this qualification for nurses working in psychiatric ward. But it lacks quantitative content and face validity, measurement error, missing value, management of missing value, and the effect of ceiling and floor.

On the other hand, some existing tools should be tested with larger sample numbers, including the study of Bondy *et al*. (1997), Stockman *et al*. (2019), Mohtashami *et al*. (2014), and Moskoei *et al*. (2017). It should be noticed that however there is no "undergraduate of psychiatric nursing" course in Iran, the expected qualification of an undergraduate nurse for mental health care is similar to the qualifications required for the undergraduate mental health nurse in other countries. The learning process in the 3 or 4-year "undergraduate Psychiatric nursing" course in the representing countries is through a curriculum focused on clinical learning of prevention, promotion, and rehabilitation in mental health. However, these qualifications won't meet postgraduate psychiatric nurses' qualifications. To sum up, the existing instruments do not meet all the COSMIN criteria, so they need to develop. One of the limitations of this study is the lack of access to EMBASE and ISI sites.

Conclusion

The results showed that the psychometric properties of existing studies are not according to the COSMIN criteria, which indicates the need to develop tools. It is recommended that COSMIN checklist be used when designing and developing the tool. Due to the extensive changes of recent decades, increasing health problems, and increasing the need for specialized mental health care, an educational system appropriate to the cultural and social context is necessary. To assess postgraduate psychiatric nurse competency is required to evaluate based on psychometric properties and adapted to the COSMIN checklist. Systematic review can help health managers and decision-makers to choose the most appropriate tool for evaluating psychiatric nurse competency.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References

1. McKnight SE. Mental health learning needs assessment: Competency-based instrument for best practice. Issues Ment Health Nurs 2013;34:459-71.
2. Govaerts MJ, Schuwirth LW, Pin A, Clement ME, van der Vleuten CP. Objective assessment is needed to ensure competence. Br J Midwifery 2001;9:156-60.
3. Sheikhbahaeddinzadeh E, Ashktorab T. Clinical competence of psychiatric nurse: A concept analysis in Rodgers 'evolutionary method. Iran J Psychiatr Nurs 2020;8:86-103.
4. Robbins L, Hoke M. Using objective structured clinical examinations to meet clinical competence evaluation challenges with distance education students. Perspect Psychiatr Care 2008;44:81-8.
5. Naegle MA, Krainovich-Miller B. Shaping the advanced practice psychiatric-mental health nursing role: A futuristic model. Issues Ment Health Nurs 2001;22:461-82
6. Adams S. Psychiatric mental health nursing: "A seat at the table." J Am Psychiatr Nurs Assoc 2015;21:34-7.
7. Fukada M. Nursing competency: Definition, structure and development. Yonago Acta Med 2018;61:1-7.
8. Alaee Karahroudy F, Sheikhbahaeddinzadeh E. A comparison of the psychiatric nursing masters curriculum in Iran and Canada. Strides Dev Med Educ 2020;17:1-6.
9. Vahedanian-Azimi A, Rahimi A. Concept of spirituality: A conventional content analysis. Qual Res J 2012;2:11-9.
10. Adib Hajbaghery M, Eshraghi Arani N. Assessing nurses' clinical competence from Their own viewpoint and the viewpoint of head nurses: A descriptive study. Iran J Nurs 2018;31:52-64.
11. Negarandeh R, Pedram Razi S, KM. Effect of clinically competent nurses services on safety and patients' satisfaction in an emergency department. HAYAT 2013;19:53-64.
12. Ebadi A, Tabanejad Z, Pazokian M. Clinical competence among MSc students of critical care nursing. Iran J Med Educ 2015;14:1036-46.
13. Stockmann C, Diaz DA, Murphy D, Dever K, Marchini M,

Huffman-frazee J, *et al*. Development of a mental health Objective Structured Clinical Examination (OSCE): A pilot study. Simulation & Gaming ,2019;50:448-60. https://doi.org/10.1177/1046878119866878.

14. Mitchell ML, Henderson A, Groves M, Dalton M, Nulty D. The Objective Structured Clinical Examination (OSCE): Optimizing its value in the undergraduate nursing curriculum. Nurse Educ Today 2009;29:398-404.

15. Watson R, Stimpson A, Topping A, Porock D. Clinical competence assessment in nursing: A systematic review of the literature. J Adv Nurs 2002;39:421-31.

16. Shirinabadi Farahani A, Rassouli M, Yaghmaei F, Alavi Majd H. Index for selecting an appropriate instrument to conduct research in health sciences: Introducing the COSMIN checklist. JHPM 2015;4:1-13.

17. Andresen E. Criteria for assessing the tools of disability outcomes research. Arch Phys Med Rehabil 2000;81 Suppl 2:S15-20.

18. Bombardier C, Tugwell P. Methodological considerations in functional assessment. J Rheumatol Suppl 1987;14 Suppl 15:6-10.

19. Yazdani N, Sharif F, Elahi N, Ebadi A, Hosseini S. Psychometric properties of quality of life assessment tools in morbid obesity: A review of literature. Evid Based Care J 2017;7:7-21.

20. Asar S, Jalalpour S, Ayoubi F, Rahmani M, Rezaeian M. PRISMA; Preferred reporting items for systematic reviews and meta-analyses. J Rafsanjan Univ Med Sci 2016;15:63-80.

21. Rezaeian M. How to report studies relevant to diagnostic accuracy. J Rafsanjan Univ Med Sci 2013;12:679-80.

22. Rezaeian M. A review on the review studies within the health domain. Res J Heal Syst 2014;9:1149-57.

23. Lang T. Just who are we and what are we doing, anyway? Needed research in medical writing. AMWA J 2009;24:106-12.

24. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, *et al*. Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. BMC Med Res Methodol 2006;6:2.

25. Mokkinka LB, Terweea CB, Patrickb DL, Alonsoc J, Stratforde PW, Knola, DL *et al*. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol [Internet]. 2010;63(7):737–45.

26. Bondy KN, Jenkins K, Seymour L, Lancaster R, Ishee J. The development and testing of a competency-focused psychiatric nursing clinical evaluation instrument. Arch Psychiatr Nurs 1997;11:66-73.

27. Mohtashami J, Salsali M, Pazargadi M, Manoochehri H, Alavi Majd H. Developing and psychometric properties check list of clinical competency in mental health nursing students. IJPN 2014;2:1-8.

28. Moskoei S, Mohtashami J, Ghalenoeei M, Nasiri M, Tafreshi MZ. Development and psychometric properties rating scale of "clinical competency evaluation in mental health nurses": Exploratory factor analysis. Electron Physician 2017;9:4155-61.

29. Chen SC, Lee SK, Rong JR, Wu CC, Liu WI. The development and psychometric testing on psychiatric nurses of a nurse case management competence scale in Taiwan. J Nurs Res 2018;26:72-9.

30. Feng D, Li H, Meng L, Zhong G. Development of a questionnaire to assess nursing competencies for the care of people with psychiatric disabilities in a hospital environment. Psychiatr Q 2018;89:699-706.

31. Mokkink LB, Boers M, van der Vleuten CP, Bouter LM, Alonso J, Patrick DL, *et al*. COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: A Delphi study. BMC Med Res Methodol 2020;20:293;

32. Mokkinka LB, Terweea CB, Patrickb DL, Alonsoc J, Stratforde PW, Knola DL, *et al*. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol 2010;63:737-45.

33. Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, de Vet HC, *et al*. COSMIN Study Design checklist for Patient-reported outcome measurement instruments. Amsterdam, The Netherlands. 2019. J Nurs Meas 2021;29:1-22.

34. Mokkink LB, Terwee CB, Knol DL, Stratford PW, Alonso J, Patrick DL, *et al*. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. BMC Med Res Methodol 2010;18(10):28; 1-8. [doi: 10.1186/1471-2288-10-22].

35. Mokkink LB, Prinsen CA, Patrick DL, Alonso J, Bouter LM, de Vet HC, Terwee CB. COSMIN Study Design Checklist for Patient-Reported Outcome Measurement Instruments. Amsterdam, The Netherlands; 2019:1-36

36. Terwee CB, Bot SD, de Boer MR, van der Windt DA, Knol DL, Dekker J, *et al*. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol 2007;60:34-42.

37. Cordier R, Chen YW, Speyer R, Totino R, Doma K, Leicht A, *et al*. Child-report measures of occupational performance: A systematic review. PLoS One 2016;11:e0147751.

38. Kardong-Edgren S, Adamson K, Fitzgerald C. A review of currently published evaluation instruments for human patient simulation. Clinical Simulation in Nursing. 2010;6 (1):e25-35

39. Colton D, Covert R. Designing and Constructing Instruments for Social Research and Evaluation. 1st ed. San Francisco: John Wiley & Sons JB; 2007.

40. Prinsen CA, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HC, *et al*. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1147-57.

41. Polit DF, Yang F. Measurement and the Measurement of Change: A Primer for the Health Professions. Philadelphia: Wolters Kluwer; 2016. p. 147.

42. Kirwan B. A Guide to Practical Human Reliability Assessment. London: CRC Press. 2017

43. Mokkink LB, de Vet HC, Prinsen CA, Patrick DL, Alonso J, Bouter LM, *et al*. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Qual Life Res 2018;27:1171-9.

44. Taghizadeh Z, Ebadi A, Montazeri A, Shahvari Z, Tavousi M, Bagherzadeh R. Psychometric properties of health related measures. Part 1: Translation, development, and content and face validity. J Iran Inst Health Sci Res 2017;16:343-57.

45. Wallace J, Rao R, Haslam R. Simulated patients and objective structured clinical examinations: Review of their use in medical education., Adv Psychiatr Treat [Internet]. 2012;8:342–8.

46. Terwee CB, Prinsen CA, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, *et al*. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. Qual Life Res 2018;27:1159-70

47. Liu NY, Hsu WY, Hung CA, Wu PL, Pai HC. The effect of gender role orientation on student nurses' caring behaviour and critical thinking. Int J Nurs Stud 2019;89:18-23.