JAPAN JOURNAL OF
NURSING SCIENCE

Japan Journal of Nursing Science (2016) **13**, 46–54 doi:10.1111/jjns.12085

ORIGINAL ARTICLE

Development of a comprehensive clinical performance assessment system for nursing students: A programmatic approach

Masoomeh IMANIPOUR[1] and Mohammad JALILI[2]

[1]*Nursing and Midwifery Care Research Center (NMCRC), Critical Care Department, School of Nursing and Midwifery and* [2]*Emergency Medicine Department, Center for Educational Research in Medical Sciences (CERMS); Critical Care Department, Tehran University of Medical Sciences, Tehran, Iran*

Abstract

Aim: Evaluation of achievement of learning objectives needs an accurate assessment program. Hence, nursing educators should move away from the use of individual assessment methods to apply a programmatic approach. The aim of this study was to develop a comprehensive assessment system for nursing students in their critical care rotation based on a programmatic approach.

Methods: The population of this study was nursing students in their critical care course. The learning objectives of the course were determined using an expert panel and classified into three categories. Suitable assessment methods were identified for each category according to the consensus of experts. Then, the assessment tools were designed and the content validity was established using content validity ratio (CVR) and index (CVI). The reliability was determined by Cronbach's alpha coefficient. The satisfaction of the participants was investigated using a questionnaire.

Results: According to the findings, all items of the assessment system had a high CVR ($P < 0.05$) and CVI ranged 0.93–0.97. The alpha coefficient of the whole system was more than 0.90 and for subsystems ranged 0.72–0.96. The findings showed that 87.5% of the instructors and 89.47% of students believed that the new assessment system had a positive impact on learning. In addition, the majority of them were satisfied with the new assessment system.

Conclusion: A programmatic approach should be used for effective evaluation of clinical performance of nursing students in critical care settings because of high validity and reliability, multidimensionality, positive educational impact, and acceptability.

Key words: clinical performance, educational assessment, nursing student.

INTRODUCTION

Along with the current global reform in health professions' education, student assessment also needs to be changed with more emphasis on learning outcomes and skill training (Wass, Mcgibbon, & Van der Veleuten, 2001).

Correspondence: Masoomeh Imanipour, Critical Care Department, School of Nursing and Midwifery, Tehran University of Medical Sciences, Dr Mirkhani Street, Tohid Square, Tehran 6459, Iran. Email: m_imanipour@tums.ac.ir

Received 22 September 2014; accepted 23 April 2015.

Learning in clinical settings is less structured than preclinical learning, so there are particular challenges to overcome in developing an effective assessment program for clinical settings (Driessen, Tartwijk, Govaerts, Teunissen, & Van Der Vleuten, 2012). Because of the complexity of the educational objectives and the plurality of the tasks and skills that medical and nursing students must learn and perform during their clinical education, institutions have the responsibility to provide a comprehensive and valid assessment system to demonstrate adequate acquisition of those objectives.

Assessing clinical performance of nursing students continues to be challenging. In nursing education, there

are multiple complex clinical skills and tasks that nursing students encounter and have to be competent in them. As a result, it is difficult to assess students' ability to carry out all aspects of such tasks (Calman, Watson, Norman, Refern, & Murrells, 2002; Dolan, 2003).

Clinical assessment of nursing students has traditionally been based on unstructured and empiric observation of students' performance by a clinical preceptor and has been permanently in doubt because it is subjective, thus running the risk of unfairness and observer bias (Calman *et al*., 2002; Dolan, 2003; Karayurt, Mert, & Beser, 2008; Norman, Watson, Murrells, Calman, & Redfern, 2002). In a study in Scotland, all nursing students believed that their clinical assessment was open to bias and that how they were assessed depended on the assessor's personality. Poor agreement between assessors and insufficient attention to psychomotor skills were some complaints of students regarding their assessment process (Calman *et al*., 2002). The findings of some other studies have shown that most nursing students are dissatisfied with clinical performance assessment tools and methods with little confidence in their results (Alavi, Irajpour, & Abedi, 2007; Baraz Pordaniani, Fereidooni Moghadam, & Loorizade, 2008; Davis, Ponnamperuma, & Ker, 2009; Imanipour & Jalili, 2012; Wishnia, Yancy, Silva, & Kern-Manwaring, 2002).

In an attempt to overcome this problem, many institutions and educators have tried to improve their clinical assessment methods. The first endeavors were focused on replacing their assessment methods with more reliable ones such as the Objective Structured Clinical Examination (OSCE), portfolio, logbook, and 360-Degree scale instead of unplanned subjective methods. Although it was a great effort, each of these methods has specific implications with strengths and weaknesses, whereas clinical performance cannot be broken down into simple parts and should be considered as a whole. As a result, a considerable conceptual shift has occurred gradually in assessment approaches (Dijkstra, Van Der Vleuten, & Schuwirth, 2010; Schuwirth & Van Der Vleuten, 2011; Van Der Vleuten & Schuwirth, 2005). Viewing assessment as a method for improvement and learning instead of viewing it as only a tool for accountability is the most important change that occurred in thinking about assessment (Norcini *et al*., 2011; Schuwirth & Van Der Vleuten, 2011; Van der Vleuten *et al*., 2012). In traditional assessment programs, clinical performance is usually evaluated using only a single assessment method. However, it is difficult to make a valid and reliable judgment about students' performance without having a comprehensive assessment system that is aligned with the education program. Thus, nursing educators need to move away from using individual assessment methods for each competence. The programmatic approach has been proposed as a result of this need and has originated from a holistic perspective on assessment (Dijkstra *et al*., 2010). In this approach, traditional methods of assessment are not just replaced by the modern ones. The central distinction is that in programmatic assessment a whole picture of students' performance can be obtained by a systematic approach that contains a careful selection of assessment methods, combination of different sources of data, design of organization systems, and formulation of the rules (Schuwirth & Van Der Vleuten, 2011).

Although there are various methods that have been used to measure students' clinical performance (Norcini & Burch, 2007), nursing educators need a comprehensive and effective assessment system of clinical performance in nursing. With regard to this need and current dissatisfaction with existing assessment methods in nursing education, and according to the complexity of nursing skills in critical care settings, a programmatic approach that uses multiple informational resources and various methods intertwined with the educational curriculum is required to make a comprehensive decision about students' clinical performance. Therefore, the aim of this study was to develop a comprehensive clinical assessment system for nursing students in intensive care settings with a programmatic approach and measure its validity, reliability, and educational impact as well as the stakeholders' satisfaction with the process and results.

METHODS

Study design

This study was conducted to design a comprehensive assessment system to evaluate clinical performance of nursing students in their intensive care rotation. The methodology was approved by the research committee of the educational development center and institutional review board of the research deputy of the university that the authors belong to. In addition, the ethics review board approved it for ethical considerations. All senior nursing students participating in the intensive care internship course in spring semester 2013 comprised the study population. The intensive care course in the 4 year degree program for undergraduate nursing education is provided in the last semester. This training course is done in a clinical setting consisting of the coronary care unit (CCU), intensive care unit (ICU), and intensive care

unit for open heart surgery (ICUOH). Students learn the theoretical concepts and basic relevant knowledge before entrance to the clinical setting. They are supposed to train in basic and advanced care of critical patients.

Steps of the study

The project was conducted in multiple steps. The first step was to determine the educational objectives based on the nursing curriculum, besides the opinion of instructors and nurses with clinical experience in critical care. After determining all educational objectives, they were divided into three categories of cognitive knowledge, clinical skills, and professional behavior (Table 1).

In the second step, a list of appropriate assessment methods for each category was provided according to the current medical education published work such as oral and written examinations, Mini-Clinical Evaluation Exercise, 360 degree, portfolio, clinical work sampling (CWS), global rating, OSCE, direct observation procedural skill (DOPS), and log book (Amin & Eng, 2006; Amin, Seng, & Eng, 2006; Dent & Harden, 2005; Epstein, 2007; Nitko, 2001; Norcini & Burch, 2007; Shahid, 2011; Swandwich, 2010) to decide which of them is most suitable for assessment of the educational objectives.

Step three was selection of the most appropriate assessment method for each category based on the learning context by an expert focus group. The participants in the expert group were five nurses of the critical care department and five medical educationists specializing in student assessment. At the start of the expert meeting, a brief explanation about implications, strengths, and weaknesses of each assessment method mentioned in above was presented by one of the investigators. Then, the experts were asked to weight different assessment methods with attention to their efficiency and applicability in the learning context. In round one, the first three priorities of assessment methods in each category were determined. In the second round, the results of the weighting were discussed by members of the expert group, and the most appropriate assessment method was selected for each educational objective category based on their consensus. Those included oral examination and CWS for cognitive knowledge and professional behavior, respectively. To evaluate clinical skills, two assessment methods were selected as the most appropriate methods including global rating and DOPS.

Upon agreement, the new assessment system was shaped which was a combination of oral examination, CWS, global rating, and DOPS, and which should be applied throughout the training course. For the next

Table 1 Educational objectives of intensive care clinical course for nursing students

Objectives	
Cognitive	– Sign and symptoms and nursing care of common diseases in intensive care units – Indications and complications of drugs and nursing care in drug prescription – Reading the cardiac rhythm and recognizing dysrhythmias – Reading results of laboratory test and reports of diagnostic/therapeutic procedures and recognizing abnormal results – Principles of oxygen therapy, airway management, and cardiopulmonary resuscitation – Medical devices in intensive care units (mechanical ventilators, defibrillators, pacemakers, monitors and electrocardiography machine)
Psychomotor	– Taking history and assessing the patients – Patient and/or their relatives' education – Working with medical devices – Special procedures, for example, endotracheal suctioning, central venous pressure measurement, internal feeding, taking arterial blood gas – All nursing care for critically ill patients
Professionalism	– Be honest, patient, trustful, confident, and responsible – Advocate for patients and/or their relatives and support them – Have commitment to duties – Work with others and have good collaboration – Adhere to intensive care setting rules and academic regulations – Respect patients, families, medical teams, and other staff – Communicate with others in a good and suitable manner

step, the assessment tools were prepared. To do this, the draft of the tool for each chosen assessment method was developed through reviewing the relevant medical education published work and similar tools (Amin *et al*., 2006; IBTICM; Norcini & Burch, 2007; Radiation Oncology Residents; RCOphth Workplace Based Assessment; Turnbull, Macfadyen, Barneveld, & Norman, 2000; Waltz, Strickland, & Lenz, 2005; Wass, Mcgibbon, & Van der Veleuten, 2001; Wishnia *et al*., 2002). The tools were designed in a 5 point Likert form for coordination, with a complete explanation at the beginning of each. The first draft of the tools was revised several times by educational assessment experts and then finalized.

The two subsequent steps were determining the validity and reliability of the new assessment system. After determining content validity of the designed tools, the new assessment system was implemented. High-quality implementation needed some preparation. Therefore, a number of meetings were provided for students and their instructors in the training course who were the study population in order to familiarize them with the new assessment system. In these orientation meetings, all components of the new assessment system, its importance and aims, how it would be applied, and the policy and procedures were explained in detail. Furthermore, a sample of assessment tools, and examinee's and examiner's guide were given to both the students and their instructors. Head nurses of clinical settings, faculties, clinical supervisor, vice chancellor of education, and dean of the nursing school attended some of these meetings for more coordination and to acquire their technical and administrative support. After preparation and orientation that took place over 3 weeks, the new assessment system was applied for all enrolled students in the intensive care course for one complete semester. These students participated in the clinical rotation and training was done under the supervision of the instructors who assessed the students' clinical performance during the course using the new assessment system. Concurrently with implementation of the new assessment system, the opinions of the study population were surveyed. As the last step and after collecting all data, the reliability was calculated.

Validity

In order to determine validity of the designed assessment system, content validity ratio (CVR) and content validity index (CVI) were used. Some educational content experts of critical care nursing were asked to evaluate all items of the tools for necessity (CVR) and for being relevant, clear, and simple (CVI). The CVR had three categories (necessary, beneficial but not necessary, not necessary), and each part of the CVI was divided into four categories (i.e. completely relevant, need major revision, need minor revision, not relevant).

Reliability

The reliability of the whole assessment system and its components was determined after implementing the new assessment system in one semester for all the students. This was accomplished using the internal consistency of items in the designed tools by Cronbach's alpha coefficient.

Educational impact and satisfaction of participants

The instructors and students were surveyed using a researcher-made questionnaire and a 10 point scale to elicit their opinions about the new assessment system. The questionnaire consisted of questions about the objectivity, feasibility, learning effect, strengths, and weaknesses of the new assessment system. Satisfaction of the participants was assessed in three domains (0–3, unsatisfied; 4–6, moderate satisfaction; 7–10, satisfied).

RESULTS

Study group

The participants were all bachelor nursing students in their fourth academic year who had registered in the critical care internship course consisting of CCU, ICU, and ICUOH. The sample of the study was equal to the study population (20 females and 18 males). All instructors who were responsible to train this group of students in the clinical settings also participated in the study. All instructors (four women and four men) had a master's degree in nursing with 7–20 years of clinical and educational experience in the critical care field.

Assessment system components

To evaluate the clinical performance of the students, a comprehensive assessment system including multiple methods was designed according to the consensus of experts in the expert group meeting. The selected assessment methods for each category of educational objectives are shown in Table 2.

The method for evaluating cognitive knowledge of students was oral examination. The examination was structured in a form with seven subtitles (Table 2) to ensure uniformity and full coverage of cognitive

Table 2 Components of assessment system

Categories	Assessment method	Items of each assessment tool
Oral examination	Cognitive knowledge	– Drugs information – Diseases – Lab. Tests interpretation – O_2 therapy and airway management – Cardiac monitoring and electrocardiography – Cardiopulmonary resuscitation algorithm – Equipment application
Clinical skills	GRF	– History taking – Patient assessment – Documentation/reporting – Patient education – Vital sign measurement – Drug prescription – i.v. infusion therapy – Intake/output measurement – Care before or after diagnostic/therapeutic procedures – Laboratory sampling
	DOPS	– Preparation of patient, equipment, and environment – Communication with patient and explanation of the procedure – Clean or aseptic technique – Correct technique – Correct sequence – Post-procedure management
CWS	Professionalism	– Interpersonal communication – Respectful behavior – Honesty – Patient advocacy – Trustworthiness – Self-confidence – Responsibility – Team work and collaboration – Adherence to academic and setting rules

CWS, Clinical Work Sampling; DOPS, Direct Observation of Procedural Skills; GRF, Global Rating Form.

objectives. In each clinical setting (CCU, ICU, ICUOH), the most appropriate and specific questions under each heading were asked and scored in a range of excellent (5), good (4), moderate (3), poor (2), and bad (1). The oral examination was performed once for each student at the end of the course and lasted approximately 15 min.

To evaluate clinical skills, two assessment methods were suggested by experts including DOPS and global rating. The procedures of the DOPS examination in each clinical setting were determined by consensus of clinical nurse specialists and nursing faculties. These consisted of some generic procedures like dressing, blood sampling, and urinary catheterization, and two specific-setting procedures for each setting (Table 2). Each student had to take two DOPS examinations in each setting with at least one specific procedure. The DOPS examinations were performed during the course, whenever the students themselves announced that they were ready, and were performed according to the existing guidelines (Amin *et al*. 2006; Dent & Harden, 2005; Swandwich, 2010). Each DOPS examination took place for approximately 15 min.

Other clinical skills were assessed according to the student's performance throughout the course and direct/indirect observation of the instructor using a Global Rating Form. The scoring ranged from very good (5) to very poor (1). These clinical skills are shown in Table 2.

The selected method to assess professional behavior was students' CWS. A special form was developed to

Table 3 Validity value of each assessment tool

Assessment category (assessment tool)	Content validity	
	CVR	CVI
Cognitive knowledge (oral examination)	0.87	0.97
Clinical skills (GRF)	0.95	0.98
Procedural skills (DOPS)	1	0.94
Professionalism (CWS)	0.91	0.93

CVI, Content Validity Index; CVR, Content Validity Ratio; CWS, Clinical Work Sampling; DOPS, Direct Observation of Procedural Skills; GRF, Global Rating Form.

Table 4 Reliability of whole assessment system and each tool

Assessment Tools	Cronbach's alpha coefficient		
	CCU	ICU	ICUOH
Oral examination	0.79	0.72	0.72
GRF	0.90	0.91	0.92
DOPS	0.80	0.87	0.92
CWS	0.79	0.90	0.96
Whole assessment system	0.90	0.95	0.92

CCU, coronary care unit; CWS, Clinical Work Sampling; DOPS, Direct Observation of Procedural Skills; GRF, Global Rating Form; ICU, intensive care unit; ICUOH, intensive care unit for open heart surgery.

document the instructor's observation of real behavior of students through the course in a range of excellent (5), good (4), acceptable (3), below expectation (2), and bad (1) (Table 2).

Validity and reliability

The results of the validity assessment are shown in Table 3. The reliability calculated by Cronbach's alpha coefficient are shown for the whole assessment system and each tool in Table 4.

Educational impact and satisfaction of participants

In the final survey, 87.5% of instructors and 89.47% of the students believed that this approach in evaluation had a positive educational impact on learning. They also stated that this system can be effective in improving learning considering the opportunity to provide feedback and hence being informative. According to participants' viewpoints, DOPS and structured oral examination had the highest educational impact. Seventy-five percent of the instructors and 76.3% of the students believed that the assessment system had acceptable objectivity. These values were 62.5% and 75.5% for its feasibility, respectively.

The majority of participants were satisfied with the new assessment system according to a 0–10 point scale. The mean score of satisfaction was 7.66 ± 1.50 and 8.70 ± 1.82 for instructors and students, respectively.

DISCUSSION

In this study, three key points were considered and the results were in accordance with all of them:

Multidimensional assessment system

According to the findings of several studies in assessment of students' clinical performance, using one assessment method is not enough to perform a valid assessment and the judgment will likely be unfair and unreliable. For example, a qualitative study on the perception of nursing students in Iran showed that students believed that the final judgment about their clinical performance should be multidimensional and cover all aspects of clinical performance rather than focusing on the knowledge. They believed that the assessors should use multiple methods and sources of information (Alavi *et al*., 2007).

Van der Vleuten and Schuwirth (2005) believe it is too simplistic to think that nursing educators can provide a comprehensive assessment by only one instrument and that this can result in reductionism. They explained that effective assessment requires a systemic and programmatic approch. A programmatic approch and using several sources of information and various methods help us to perform a high-quality assessment. Of course, programmatic assessment is not a random selection of various methods; rather, it is a proccess that is initiated by determining educational objectives and that fitness to purpose is the cornerstone of the design of a high-quality assessment system (Dijkstra *et al*., 2010).

In this study, the aim was designing a comprehensive assessment system based on a programmatic approch. Therefore, the present authors selected a series of various assessment methods in alignment with the educational objectives which were set a priori. The present authors tried to develop a system with different assessment methods to achieve full coverage of educational objectives and take a full picture of the students' clinical performance.

Oral examination was the chosen method to assess cognitive objectives in this study. The oral examination is a traditional assessment method and has longstanding application in medical education (Amin & Eng, 2006). For example, in assessment of surgery residents,

oral examinations have been used in addition to current clinical examinations (Smith, Dollase, & Boss, 2003).

One of the common assessment methods to evaluate "does" is global rating. Global rating is used usually to determine quality of randomly observed behavior and real performance of students that is rated on a Likert scale (Amin *et al*., 2006). In the present authors' study, global rating was used to assess general clinical skills in the real clinical settings on a 1–5 Likert scale.

For assessment of procedural skills, DOPS was the method suggested by experts. The main goal of DOPS is evaluation of "does" in Miller's pyramid (Amin *et al*., 2006; Swandwich, 2010). While DOPS is commonly used for evaluation of physicians and surgery residents (Beard, Strachan, & Davies, 2005; Newble, Paget, & Mclaren, 1999), its application has grown in other disciplines like nursing (Beard *et al*., 2005; Swandwich, 2010).

Clinical work sampling was used to assess communication skills, teamwork and collaborative skills, and professionalism. CWS is an in-training evaluation method that collects data on observed real behavior (Amin *et al*., 2006). It has been studied in different researches. For example, Turnbull *et al*. (2000) developed a CWS form in a five part Likert scale and asessed 62 third year students during their internal medicine inpatient experince. The collection of ongoing performance data was reasonably feasible, reliable, and valid.

Utility criteria of the assessment system and its component

Van der Vleuten describes utility criteria of usefulness of an assessment method: validity, reliability, its acceptability, educational impact, and cost-effectiveness (Epstein, 2007; Shumway & Harden, 2003). To verify the first utility criteria, validity, the present authors determined CVR and CVI. CVR and CVI are quantitative methods of determining content validity of an assessment scale which measure the neccessity and relevance of each of the items based on experts' agreement. (Polit & Beck, 2006). According to "minimum values of CVR", which is suggested by Lawshe (1975), when a content evaluation panel is composed of 13 members (the number of experts in this study), a minimum CVR of 0.54 is required to accept an item is essential with a *P*-value of less than 0.05 (Lawshe, 1975). As a result, all items of all developed assessment tools in this study were essential at a confidence level of 95%.

After verifying nesseccity of all items of new tools, CVI was measured by calculating the proportion in agreement about relevance (computing the number of items with a score of 3 or 4, divided by the number of experts) (Polit & Beck, 2006). The minimum accepted value of CVI is 0.80 (Davis, 1992) or 0.90 (Waltz *et al*., 2005). According to these authors, the content validity of all tools of the new assessment system was proved. This could be the result of the programmatic approach in designing the new assessment system. The programmatic approach to assessment, by triangulating various methods and different information sources, can result in unbiased judgment with high validity and reliability (Van der Vleuten & Schuwirth, 2005; Wilkinson *et al*., 2011).

There are different methods to determine reliabilty, the second utility critera of an assessment tool, such as internal consistency and Cronbach's alpha coefficient. The findings of this study showed that all assessment tools and the whole system have high reliability ($\alpha > 0.70$ and $\alpha > 0.90$, respectively). Educational specialists emphasize that nursing educators should focus on the comprehensive assessment to increase reliability rather than individual methods (Van der Vleuten & Schuwirth, 2005). Wass *et al*. (2001) showed by combination of different tests, the coverage of all content curriculum and educational objectives will be maximized and the reliability will increase.

Educational impact and satisfaction with the assessment system

Educational impact and acceptability to stakeholders are the other utility criteria of a particular assessment method (Epstein, 2007; Van der Vleuten & Schuwirth, 2005). In the current study, the majority of students and instructors were satisfied with the new assessment system. They believed the assessment system was objective, clear, and feasible with positive impact on learning.

Assessing stakeholders' satisfaction is a technique to study acceptability of a method, instrument, service, or program. Amin (2012) believes that having a programmatic approach in student assessment is more likely to increase different stakeholders' satisfaction. Some other factors which possibly affected the students' and instructors' satisfaction were clear explanation of the assessment system and its components, how it would be performed, what its rules were, the scoring method, and its objectivity and fairness. Baartman, Bastiaens, Kirschner, and Van Der Veleuten (2006) explain clearness, meaningfulness, and fairness in viewpoint of stakeholders as criteria of quality of an assessment system that reasonably can lead to increased satisfaction.

The findings also showed that the majority of students and instructors believed this new assessment system had

a positive educational impact, because it was formative, multidimensional, and gave constructive feedback. The positive educational impact and giving feedback is one of the criteria of a good assessment system (Epstein, 2007; Norcini *et al*., 2011; Shumway & Harden, 2003; Van der Vleuten & Schuwirth, 2005). As explained, a paradigm shift is occurring from assessment of learning to assessment for learning. Thus, to move in the direction of learning and positive educational impact, a good assessment system should be formative and give feedback (Schuwirth & Van Der Vleuten, 2011; Van der Vleuten & Schuwirth, 2005). The endeavors of this new assessment system were guided toward this.

CONCLUSION

This study aimed at developing a comprehensive clinical performance assessment system for nursing students. A programmatic approach was adapted, which not only resulted in full coverage of educational objectives, but also led to high validity, reliability, feasibility, and acceptability, as well as stakeholders' satisfaction.

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the members of Tehran University of Medical Sciences who supported or assisted this study, especially Dr Azim Mirzazadeh, Dr Nahid Dehghan Nayeri, and Mr Hamid Haghani, for all their contributions. In addition, the authors would like to acknowledge the valuable contribution of all students and instructors who participated in this study. This research has been supported and funded by Tehran University of Medical Sciences grant No. 10264.

DISCLOSURE

No conflict of interest has been declared by the authors.

AUTHOR CONTRIBUTIONS

M.I. designed the study, applied the study and data collection, analyzed the data and drafted the manuscript. M. J. designed and supervised the study, drafted the manuscript and gave technical support.

REFERENCES

Alavi, M., Irajpour, A. R. & Abedi, H. A. (2007). Some concepts in the evaluation of clinical education: A qualitative study on the experiences of nursing students and clinical teachers. *Strides in Development of Medical Education*, *4*, 10–18 (in Persian).

Amin, Z. (2012). Purposeful assessment. *Medical Education*, *46*, 3–12.

Amin, Z. & Eng, K. H. (2006). *Basics in Medical Education*. Singapore: World Scientific.

Amin, Z., Seng, C. Y. & Eng, K. (2006). *Practical Guide to Medical Student Assessment*. Singapore: World Scientific.

Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A. & Van Der Veleuten, C. P. M. (2006). The wheel of competency assessment: Presenting quality criteria for competency assessment programs. *Studies In Educational Evaluation*, *32*, 153–170.

Baraz Pordaniani, S., Fereidooni Moghadam, M. & Loorizade, M. (2008). Clinical education status according to the nursing and midwifery students' point of view, Tehran university of medical sciences. *Strides in Development of Medical Education*, *5*, 102–111 (in Persian).

Beard, J., Strachan, A. & Davies, H. (2005). Developing an education and assessment framework for the foundation program. *Medical Education*, *39*, 841–851.

Calman, L., Watson, R., Norman, I., Refern, S. & Murrells, T. (2002). Assessing practice of student nurses: Methods, preparation of assessors and student views. *Journal of Advanced Nursing*, *38*, 516–523.

Davis, L. (1992). Instrument review: Getting the most from your panel of experts. *Applied Nursing Research*, *5*, 194–197.

Davis, M. H., Ponnamperuma, G. G. & Ker, J. S. (2009). Students perceptions of a portfolio assessment process. *Medical Education*, *43*, 89–98.

Dent, J. A. & Harden, R. M. (2005). *A Practical Guide for Medical Teachers*. Churchill: Elsevier.

Dijkstra, J., Van Der Vleuten, C. P. M. & Schuwirth, L. W. T. (2010). A new framework for designing programmes of assessment. *Advance in Health Science Education*, *15*, 393–397.

Dolan, G. (2003). Assessing student nurse clinical competency: Will we ever get it right? *Journal of Clinical Nursing*, *12*, 132–141.

Driessen, E. W., Tartwijk, J. V., Govaerts, M., Teunissen, P. & Van Der Vleuten, C. P. M. (2012). The use of programmatic assesment in the clinical workplace: A maastricht case report. *Medical Teacher*, *34*, 226–231.

Epstein, R. M. (2007). Assessment in medical education. *The New England Journal of Medicine*, *356*, 387–395.

Imanipour, M. & Jalili, M. (2012). Nursing students' clinical evaluation in students and teachers views. *Iranian Journal of Nursing Research*, *7*, 17–26 (in Persian).

Karayurt, O., Mert, H. & Beser, A. (2008). Astudy on development of a scale to assess nursing students' performance in clinical settings. *Journal of Clinical Nursing*, *18*, 1123–1130.

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*, 563–575.

Newble, D., Paget, N. & McLaren, B. (1999). Revalidation in Australia and New Zealand: Approach of the Royal Australasian college of physicians. *British of Medical Journal*, *319*, 1185–1188.

Nitko, A. J. (2001). *Educational assessment of students*. Upper Saddle, NJ: Merrill Prentice Hall.

Norcini, J., Anderson, B., Bollela, V., Burch, V., Costa, M. J., Duvivier, R. *et al*. (2011). Criteria for good assessment: Consensus statement and recommendations from the ottawa 2010 conference. *Medical Teacher*, *33*, 206–214.

Norcini, J. & Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Medical Teacher*, *29*, 855–871.

Norman, I. J., Watson, R., Murrells, T., Calman, L. & Redfern, S. (2002). The validity and reliability of the methods to assess the competence to practise of preregistration nursing and midwifery students. *International Journal of Nursing Studies*, *39*, 133–145. http://dx.doi.org/10.1016/S0020-7489(01)00028-1.

Polit, D. F. & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, *29*, 489–497.

Schuwirth, L. W. & Van Der Vleuten, C. P. M. (2011). Programmatic assessment: From assessment of learning to assessment for learning. *Medical Teacher*, *33*, 478–485.

Shahid, H. (2011). Faculty development: DOPS as workplace-based assessment. *Education in Medicine Journal*, *3*, e32–e43.

Shumway, J. M. & Harden, R. M. (2003). The assessment of learning outcomes for the competent and reflective physician. AMEE Guide No. 25. *Medical Teacher*, *25*, 569–584.

Smith, S. R., Dollase, R. H. & Boss, J. A. (2003). Assessing student's performance in a competency-based curriculum. *Academic Medicine*, *78*, 97–107.

Swandwich, T. (2010). *Understanding medical education: Evidence, theory and practice*. Malaysia: Wiley-Blackwell.

Turnbull, J., Macfadyen, J., Barneveld, C. & Norman, G. (2000). Clinical work sampling: A new approach to the problem of in-training evaluation. *Journal of General Internal Medicine*, *15*, 556–561.

Van Der Vleuten, C. P. M. & Schuwirth, L. W. T. (2005). Assessing professional competence: From methods to programs. *Medical Education*, *39*, 309–317.

Van Der Vleuten, C. P. M., Schuwirth, L. W. T., Driessen, E. W., Dijkstra, J., Tigelaar, D., Baartman, L. K. J. *et al*. (2012). A model for programmatic assessment fit for purpose. *Medical Education*, *34*, 205–214.

Waltz, C. F., Strickland, O. L. & Lenz, E. R. (2005). *Measurement in nursing and health research*. New York: Springer Publishing Co.

Wass, V., McGibbon, D. & Van der Veleuten, C. P. M. (2001). Composite undergraduate clinical examinations: How should the components be combined to maximize reliability? *Medical Education*, *35*, 326–330.

Wilkinson, T. J., Tweed, M. J., Egan, T. G., Ali, A. N., McKenzie, J. M., Moore, M. *et al*. (2011). Joinning the dots: Conditional pass and programmatic assessment enhances recognition of problems with professionalism and factors hampering student progress. *BMC Medical Education*, *11*, 29.

Wishnia, G. S., Yancy, P., Silva, J. & Kern-Manwaring, N. (2002). Evaluation by exception for nursing students. *Journal of Nursing Education*, *41*, 495–497.

IBTICM Direct Observation of Procedural Skills (ICM DOPS). [Cited 18 Mar 2012.] Available from URL: https://www.nwpgmd.nhs.uk/specialties/sites/default/files/ICM%20DOP.pdf

Radiation Oncology Residents Global Rating Forms. [Cited 13 Jul 2013.] Available from URL: https://www.acgme.org/acgmeweb/Portals/0/430_RadOnc_GlobalRev.pdf

RCOphth Workplace Based Assessment Direct Observation of Procedural Skills (DOPS). 2007.02. [Cited 12 Mar 2012.] Available from URL: http://curriculum.rcophth.ac.uk/learning-outcomes/pract_skills/ps11

