Development and Psychometric Testing of the Gerontological Nursing Clinical Reasoning Scale

Pratsani Srikan, Sirisuda Tachavises*

Abstract: The assessing and evaluating of quality in nursing care of older adults is important to promise high standard patients' outcomes. Reliable and appropriate instrument measured clinical reasoning skills in gerontological nursing was not readily available. The purpose of this study was to develop and validate the Gerontological Nursing Clinical Reasoning Scale in Thailand. A cross-sectional study was performed with a psychometric evaluation of this new developed scale. A pool of 18 short clinical vignettes composed of 72 items with a 5-response Likert type scale was initially developed in the format of the Script Concordance Test. The Script theory and the hypothetical-deductive theory were applied. The scale development process had six steps, including content validity and reliability assessment. Then, it was tested in 80 participants who were in equal groups of senior nursing students and registered nurses in Phayao province. The construct validity by known group technique was used. Twelve advanced practice nurses who were specialized in gerontological nursing and considered as a panellist group were involved. The scores among groups were compared by using the one-way analysis of variance with a Scheffe's post hoc test. The 14-vignettes the Gerontological Nursing Clinical Reasoning Scale which comprised 47 items was finalized. The findings showed that the panellists, registered nurses, and nursing students had statistical different mean scores. The panellists had the highest scores, followed by the registered nurses, and nursing students, respectively. The Scale had an acceptable level of construct validity and internal consistency and could distinguish clinical reasoning skills among the three groups of nurses sampled. Nursing educators can introduce this scale to measure clinical reasoning in gerontological nursing courses but further testing with other populations is needed.

Pacific Rim Int J Nurs Res 2019; 23(3) 243-257

Keywords: Clinical reasoning, Gerontology, Gerontology nursing, Nursing students, Script Concordance Test

Received 4 October 2018; Accepted 24 February 2019

Correspondence to: Pratsani Srikan*, *RN, PhD, Nursing instructor, Boromarajonani College of Nursing, Phayao, Thailand.*
E-mail: *pratsanii@gmail.com*
Sirisuda Tachavises, *RN, MNS, Instructor, Boromarajonani College of Nursing, Phayao, Thailand.* ***E-mail:*** *sirisuda.t@bcnpy.ac.th*

Introduction

In most countries, people are living longer on average and often develop multiple illnesses which need advanced care. Nurses, both newly graduated and experienced, have to be sufficiently qualified and skilled in order to meet the needs of the aging population.[1] In Thailand, the setting of this study, gerontological nursing is one of eight subjects in which nursing students are required to gain experience

and then take an examination before obtaining the Thai Nursing License. Being faced with complex, unpredictable and dynamic clinical care situations is a challenge for nurses, especially newly graduated nurses. They are expected to be able to speedily analyze, make decisions in a timely manner, and provide effective and proper care for patients in multifaceted and uncertain circumstances in order to guarantee positive patient outcomes.[2] Clinical reasoning (CR) is a basic cognitive process for nurses to use in patient care and they are expected to be able to integrate a patient's problems with skilled interventions using accurate reasoning.[3] Levett-Jones and associates defined clinical reasoning as a thinking and decision-making process which involves considering a situation, gathering cues and processing the information to identify problems. This leads to the development of plans, applying solutions, evaluating, and reflecting on lessons learned.[4] According to Simmons, clinical reasoning is a complex cognitive process involving both formal and informal rational approaches to collect and scrutinize patient data.[5] This reasoning includes in its process, legal, ethical, and professional components.[4, 6] In this study, the clinical reasoning in gerontological nursing of nursing students was defined as a capacity of senior nursing students to interpret, calculate and analyze information derived from scenarios related to older adults. They then were required to use reasoning to determine appropriate assessment, hypothesis, intervention, and evaluation based on the vignettes provided.

Nursing education programs need to assist students to gain confidence and achieve nursing competency for safe and effective care in varied situations. During their learning journey, students are expected to develop clinical reasoning skills based on a cognitive learning model in the classroom and in the clinical environment by putting their efforts evaluating clinical facts so that a clear perception of a patient's problems emerges.[7] Giving appropriate justifications to provide suitable care for individuals, especially older adults who have multiples diseases, is even more challenging for nursing students in clinical settings. A study by Staydt and Merriman showed that 421 nursing students had a wide variety of clinical placements, yet nearly two-thirds did not have enough confidence in providing nursing interventions to meet patients' safety.[8] During four years in nursing school, students might have variable levels of supervision as well as different clinical settings and inconsistent opportunities to practice.[8] In another study the level of clinical reasoning skills assessed in senior nursing students in Thailand was rated as only "moderate" when they graduated and were waiting to take their nursing license examination.[9] During four years in nursing school, students might have different frequency of opportunities to practice, level of supervision, and clinical settings.[8] This can affect the level of their confidence and clinical reasoning skills.

The process of training nursing students to attain educational outcomes, particularly in clinical reasoning, has been observed and evaluated by various conventional techniques such as direct observation,[10,11] multiple choice questionnaires (MCQs), oral examination,[12] and objective structured clinical examination (OSCE).[13] Direct behavioural observation lacks a clear pattern. The objective examination is also problematic as there is a selection bias.[2] A MCQs is reliable in assessing the technical reasons, but it cannot be used to assess reasoning skills in the clinical situation in which there is not enough information or circumstances that are highly uncertain.[14] Although being an alternative choice, the oral examination also has its limits in setting the standard for assessing and rating, as well as application in large groups. The clinical reasoning skill in the OSCE can similarly be influenced by different format and amount of inter-station.[15]

The Script Concordance Test (SCT) is a tool to assess clinical reasoning. It was developed by a group of medical professors in Canada and the Netherlands around 18 years ago and based on cognitive psychology script theory and hypothetical-deductive theory.[16] The former is used to explain patterns or scripts of human behaviour, while the latter one involves information processing, testing hypotheses, and clinical decision making.[17] Using cognitive psychology script theory, new clinical information of patients emerges and is valued by clinicians in interpretation and making

clinical judgments. The SCT has been used in medical education as an alternative approach when using reasoning in unclear circumstances. A respondent must decide on a possible diagnosis to find alternative approaches, or investigative and treatment options which are selected on a Likert scale.[18,19] This tool is used to evaluate specific aspects of clinical reasoning and the ability to translate the medical information under conditions of uncertainty.[20,21]

Methods

Study design

A cross-sectional study composed of two phases, the development of the Geron-NCRS and psychometric testing.

Samples and Settings

During the scale development phase, five advanced practice nurses (APNs) were interviewed. During the next phase of the psychometric testing, there were 40 RNs from a general hospital located in the northern region in Thailand who had been working in medical or surgical departments and had experienced in nursing for at least two years. Also, 40 senior nursing students from a nursing college in the same region were recruited to the study. These students were required to have finished their clinical placements in community and hospital-based settings. Also, 12 of a total of 25 APNs specializing in gerontological nursing throughout the country agreed to complete the scale.

Ethical considerations

Study approval was received from the Research Ethical Review Committee of the Phayao Hospital (COA No.9), and Boromarajonnani College of Nursing, Phayao (10/2016), Thailand. Data collection took place April-May 2017. All participants signed informed consent forms and had the right not to participate in the research. They were able to withdraw from the research at any time. They could do so without advanced notifications or apprehensions of losing any entitlements. All research data was de-identified and stored in a secure cabinet whilst computer files were protected by a password.

Development of the Geron-NCRS

The development of the Geron-NCRS comprised six steps: 1) Clarifying definition and elaborating of the clinical vignette/scenario, 2) generating an item pool, 3) choosing the format of measurement, 4) content validation, 5) evaluation of vignettes and items, and 6) examinees' testing and scoring assembly. [30]

Step 1: Clarifying definition and elaborating of the clinical vignette/scenario

The scope of scale development was elucidated by a literature review and the Thai Nursing Council licensing test blueprint in gerontological nursing being taken into account. Then, five APNs with expertise in gerontology were interviewed to determine key aspects of gerontological nursing required for undergraduate students. A scale development outline was achieved. The scale was used to measure competency in knowledge, skills, and attitudes to care for older adults who displayed signs of the geriatric syndrome typical of aging. These include insomnia, malnutrition, osteoporosis, hearing and visual impairment, dementia/delirium, depression, chronic illness (such as diabetes, hypertension, cerebrovascular disease, benign prostate hypertrophy, and cancer), end of life care, and poly-pharmacy. Health assessment, communication, health promotion, and rehabilitation skills were also encompassed both in the community and institutionalized situations.

Step 2: Generating item pool

A pool of 18 short clinical vignettes with a total of 72 Likert scale items in Thai to assess clinical competence in gerontological nursing was developed. It contained patient assessment, nursing diagnosis, and nursing intervention domains.

Step 3: Choosing the format of measurement

The Likert type scale had five response options (-2, -1, 0, +1, and +2), ranging from completely contraindicated (-2) to completely indicated or absolutely necessary (+2).

Step 4: Content validation

The research team screened, discussed and evaluated the 72 items within the 18 vignettes. Three vignettes were deleted due to redundancy and being not a match for gerontology, leaving 15 vignettes and

60 items. A panel of three content experts, who were specialists in gerontology, clinical reasoning, and nursing education, were then asked to review and rate the 60-item Geron-NCRS for relevance, comprehension, and clarity. No item needed to be removed, but some minor revisions were required.

Step 5: Evaluation of vignettes and items

Ten nursing students were then asked to complete the scale in order to measure its readability. Test instructions were given to the students for partaking, rating and commenting on the instrument and individual vignette and item. The feedback was verified and taken to scale revision for its clarity and understandability.

Step 6: Examinees' testing and scoring assembly

The known group validity method was adopted. In the beginning, 20 of 25 APNs specialized in gerontology throughout the country were contacted and invited by phone to be in a panellist group. Twelve panellists accepted invitations. According to Gagnon and Charlin, a panel of 10-15 expert members is appropriate to produce credible and reliable scores.[31] The scoring grid was then generated and calculated, with a maximum of 1 for each item. Any item with bi-modal, uniform divergence or discrete outlier response was discarded. Some extreme answers were deleted and some "median" answers were discussed or deleted.[32] To assess the psychometric properties of the Geron-NCRS with a paper-pencil format, a cross-sectional study was implemented by recruiting two more groups of participants. The sample size was calculated with a power of 0.95, an error of 0.05, and a medium effect size of 0.25. It was determined that 40 subjects were needed per group.

The first group was 40 senior nursing students from a nursing college; the second group was made up of 40 RNs with no expertise in gerontology, recruited from a tertiary hospital in the northern region of Thailand. For the RNs and students, convenience sampling on a voluntary basis was applied and invitation letters were sent to participants. All nursing students and RNs had experience in both community and hospital settings. The student group had finished community work two months before recruitment while the RNs worked in a community setting for much longer.

In addition, demographic data record forms for the three groups were established. In the student group, age, gender, average grade point (GPA), confidence while studying a gerontological nursing course, and confidence to pass the gerontological nursing comprehensive examination were asked. For the RNs, age, gender, work department, years of work experience, activeness of providing care of older adults, and preferences and confidence in taking care of older adults were included. For the panellists, age, gender, work setting, and years of experience as an APN in gerontology were also inquired. The content validity of these forms was also achieved by three experts, a nursing faculty member who specialized in clinical reasoning and the script concordance test, an APN in gerontology, and a nursing faculty who specialized in gerontology.

Data Analysis

Statistical analysis was done with the Statistical Package for the Social Sciences (SPSS), Predictive Analytics SoftWare (PASW) Statistics Program version 20. Descriptive and inferential statistics were applied during data analysis. Statistical significance at the 0.05 level was considered acceptable. The demographics of the three groups of participants were calculated and the normality of score distributions was evaluated by a Shappiro-Wilk's test. A t p-value greater than 0.05 along with illustration of a histogram, normal Q-Q plots and box plots indicated that the scores of each group were normally distributed. The homogeneity of group variance was estimated with the Levene's test (p>.05) and showed that this assumption was not violated (p=0.10). To compare differences within and between mean scores of the three groups, one-way analysis of variance (ANOVA) with a Scheffe's post-hoc test were used due to unequal samples.

Reliability and Internal Consistency

The content validity of the Geron-NCRS was at 0.90. The reliability was set using the Cronbach's coefficient at 0.75, indicating its satisfactory reliability coefficient.[33] Pass and fail cut-off scores were also

performed from the panellists' mean score and standard deviations. For the undergraduate students, the cut-off score at 3-4 standard deviations (-4*SD* and-3*SD*) below the panellists' mean scores were suggested and a smaller number of standard deviations were considered in recent graduates.[34, 35] The cut-off score at -4 and -2 standard deviations (-4*SD* and-2*SD*) were used in this study.

Results

Instrument Development

To obtain the scoring grid of the 60-item Geron-NCRS, the scores from 12 APNs were calculated. There were 7 out of 15 vignettes that had each item showed a single modal response. Also, bi-model responses were found in vignette 1 (items 2 and 3) and vignette 15 (items 57, 59 and 60). Consequently, four items (57, 58, 59 and 60) in vignette 15 were deleted because it was left with one item (56). Nine more items (2, 3, 6, 9, 15, 20, 22, 39, and 43) from 8 of the vignettes (numbers 1, 2, 3, 4, 5, 6, 10, and 11) were also deleted[32] because of a single or bi-modal response. The final version of the Geron-NCRS comprised 14 vignettes with 47 items; each vignette has 2-4 nested items shown in **Table 1**. The scores ranged from 4.46 to 47, with higher scores indicating more clinical reasoning skill. It took 45-60 minutes to complete the Scale. The scoring grid of the reference panel for the 47-item Geron-NCRS is shown in **Table 2**. The process of scale development is displayed in **Figure 1**.

Table 1 Example vignettes

You work as community nurse. You visit Mrs. Lin, a 70 year old woman with the history of diabetes and osteoarthritis. Mrs. Lin tells you that she does not want to go out and join the community activity because of frequent toileting and urine leakage. She normally needs to urinate 7-8 times during the day and 4-5 times at night.

Item	Column 1	Column 2	Column 3
1.	If you thought to ask Mrs. Lin about her routine medications.	And then Mrs. Lin says "I haven't had any surgery or experienced any allergy"	The relevance of this assessment becomes: ¨ completely or partially contraindicated ¨ not very useful or possibly harmful ¨ neither more nor less useful ¨ useful ¨ necessary or absolutely necessary
2.	If you thought to ask Mrs. Lin what help she needs the most.	Mrs. Lin says "I has been taking good care of myself and I don't understand why it happens to me"	The relevance of this assessment becomes: ¨ completely or partially contraindicated ¨ not very useful or possibly harmful ¨ neither more nor less useful ¨ useful ¨ necessary or absolutely necessary
3.	If you plan to assess Mrs. Lin's stress level	Mrs. Lin feels that she has become her family burden.	The relevance of this assessment becomes: ¨ rejected ¨ less relevant or possibly less appropriate ¨ the information has no effect on the assumption ¨ needs to be explored in the near future ¨ needs to be explored in the immediate future
4.	If you plan to talk with Mrs. Lin's daughter about bathroom modifications	Mrs. Lin says "I like gardening and it my stress killers"	The relevance of this intervention becomes: ¨ completely or partially contraindicated ¨ not very useful or possibly harmful ¨ neither more nor less useful ¨ useful ¨ necessary or absolutely necessary

Table 2 Scoring grid from the panellists

Vignettes	Items	Responses				
		−2	−1	0	+1	+2
1.	Chronic kidney disease and Gout					
	1. AVF and arm exercise	0.00	0.22	0.11	1.00	1.00
	2. Physical injury risks	0.00	0.40	0.60	0.40	1.00
	3. Follow up and medication adherence	0.00	0.00	0.13	0.38	1.00
2.	Chronic pulmonary disease and Vision					
	4. Cataract and treatment	0.50	1.00	0.30	0.50	0.80
	5. Taking medication assessment	0.00	0.33	0.67	0.00	1.00
	6. Pulmonary rehabilitation	0.00	0.20	0.40	0.80	1.00
3.	Diabetes mellitus					
	7. Foot examination	0.00	0.00	0.83	0.17	1.00
	8. Depression screening	0.00	0.00	0.60	0.80	1.00
	9. Assessment of insulin injection practice	0.50	0.50	0.50	0.50	1.00
4.	Cerebrovascular disease and Pressure Sores					
	10. Pressure sore assessment	0.14	0.43	0.14	0.00	1.00
	11. Discharge planning needs	0.00	0.00	0.33	1.00	0.70
	12. Caregiver support and stress reduction	0.00	0.00	0.43	0.29	1.00
5.	Hypertension and Osteoporosis					
	13. Perception the risks of hip fracture	0.00	1.00	0.40	0.20	0.80
	14. Skin assessment	0.20	1.00	0.60	0.00	0.60
	15. Gathering and giving information to families	0.00	0.00	0.22	0.11	1.00
6.	Liver cancer and End of life care					
	16.Giving information to families	0.00	0.11	0.11	0.11	1.00
	17. Pain assessment	0.00	0.67	0.33	0.00	1.00
	18. Preparing families for loss	0.11	0.00	0.11	0.11	1.00
	19. Non pharmacological relaxation techniques	0.14	0.14	0.14	0.29	1.00
7.	Breast cancer and Depression					
	20. Caregiver burnout assessment	0.00	0.00	0.11	0.22	1.00
	21. Chemotherapy side effects	0.00	0.00	0.14	1.00	0.60
	22. Chemotherapy administration	0.17	0.50	0.33	0.00	1.00
	23. Integrating spiritual care	0.00	0.00	0.11	0.22	1.00

Table 2 Scoring grid from the panellists (Cont.)

Vignettes	Items	Responses				
		-2	-1	0	+1	+2
8.	Dementia and Cerebrovascular disease					
	24. Assessment of activity of daily living	0.00	0.13	0.13	0.25	1.00
	25. Assessment of Caregiver's knowledge	0.22	0.00	0.11	0.00	1.00
	26. Assessment of patients' knowledge	0.00	0.00	0.57	0.14	1.00
	27. Aggressive behaviour approach	0.00	0.40	0.80	0.20	1.00
9.	Diabetes mellitus, Hypertension, Poly-pharmacy and Cerebrovascular disease					
	28. Smoking history- taking skill	0.17	0.33	0.33	0.17	1.00
	29. Medicines and side effects	0.00	0.29	0.43	0.00	1.00
	30. Giving information of symptom progression	0.00	0.00	0.11	0.11	1.00
	31. Family education	0.00	0.50	1.00	0.00	0.50
10.	Chronic kidney disease and BPH					
	32. Knowledge the causes of disease	0.00	0.29	0.29	0.14	1.00
	33. BPH and urinary incontinence	0.20	0.20	1.00	0.60	0.40
	34. Skin care and urinary incontinence	0.00	0.33	0.50	0.17	1.00
11.	Diabetics and Hypertension and					
	35. History taking and examination	0.00	0.17	1.00	0.33	0.50
	36. Alcohol consumption assessment	0.00	0.43	0.29	0.00	1.00
	37. Fall risks	0.25	1.00	0.25	0.75	0.75
	38. Medication non-adherance	0.00	0.20	0.80	0.40	1.00
12.	Cerebrovascular disease and Constipation					
	39. Wound care	0.17	0.50	0.00	0.33	1.00
	40. Constipation assessment	0.20	1.00	0.20	0.20	0.80
	41. Physical rehabilitation	0.00	0.40	1.00	0.20	0.80
	42. Constipation and diet	0.22	0.11	0.00	0.00	1.00
13.	Colon cancer, Surgery and Delirium					
	43. Delirium and medication	0.00	0.00	0.83	0.17	1.00
	44. Delirium and restraints	0.75	.50	0.75	0.00	1.00
	45. A family approach to delirium	0.00	0.00	1.00	0.60	0.80
14.	Hypertension, Arthritis and Urinary incontinence					
	46. Medication and urinary incontinence	0.00	0.40	0.80	0.20	1.00
	47. Home modification and urinary incontinence	0.20	1.00	1.00	0.00	0.20

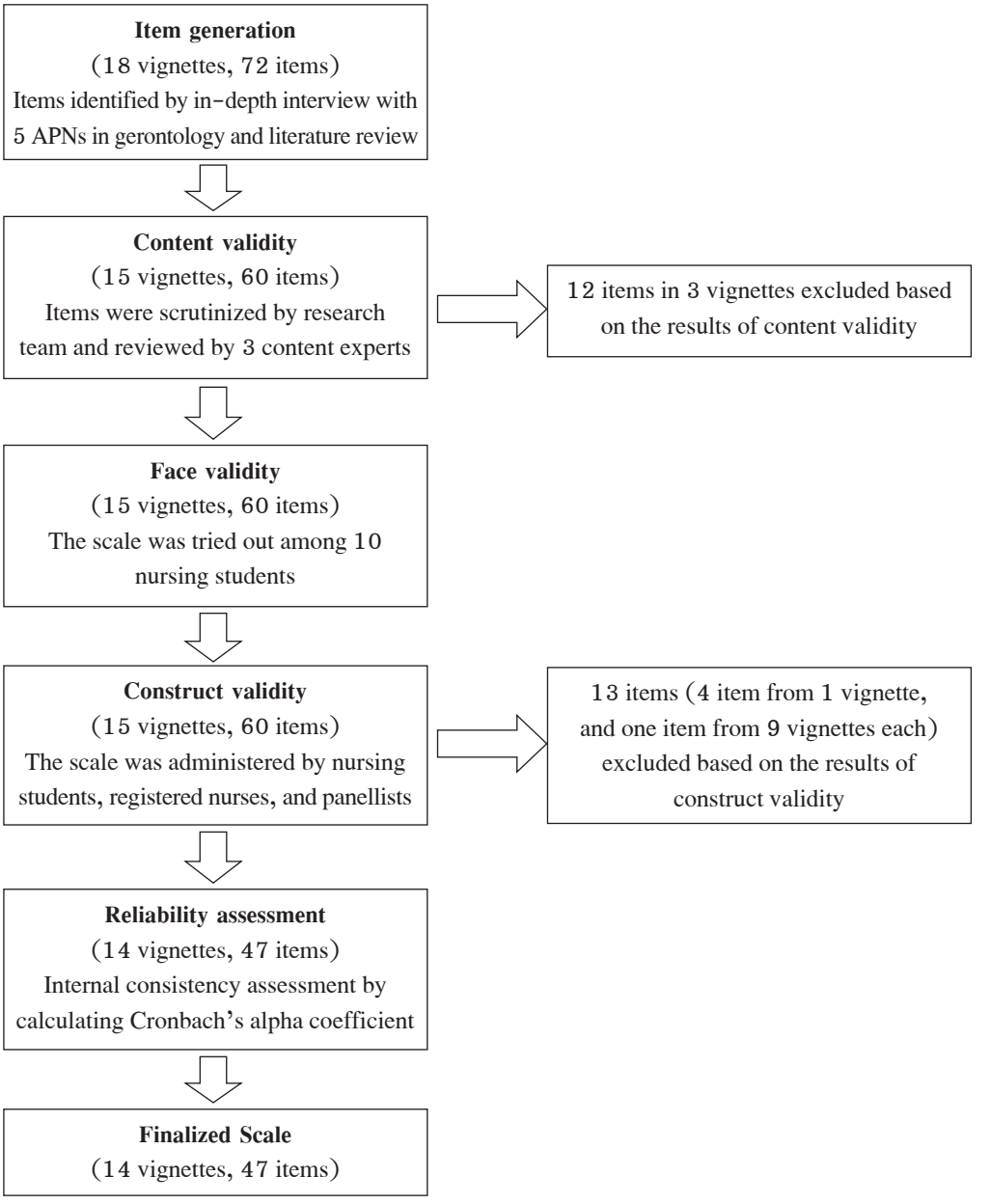


Figure 1 Diagram of scale development

Psychometric Testing

Demographic data of participants

The majority of the students were female (n=36) and their ages ranged from 21-22 years. Sixty (n=24) and 25% (n=10) of them had grade point average (GPA) at 3.01-3.50 level and 2.51-3.00 level, respectively. Seventy-seven percent (n=31) of the student participants felt confident while studying the gerontological nursing courses. Sixty-five percent (n=26) felt moderately confident that they would pass

the gerontological nursing comprehensive examination before they graduated, and 75% (n=30) passed that examination at the first attempt.

For the RN group, 95% were female (n=38) and 5% male (n=2). Ages ranged from 25-53 years (mean=36.95, *SD*=6.59). Half of the group was working in a medical department; the rest in a surgical department, and the nurses' work experience ranged from 3-31 years (mean=16.24, *SD*=8.20). They were actively caring for older adults in the current workplace. The preferences in providing care for older adults among RNs were neutral (50%), some preference (47.5%) and high preference (2.5%), while the confidence in providing care for seniors were some confidence (52.5%), average confidence (37.5%) and some unconfidence (10%).

For the panellist group, 92% were females (n=11) and 8% male (n=1), ages ranged 32-45 years (mean= 42.5, *SD*=4.58). Nine out of twelve were working in tertiary hospitals, in the areas of medical, surgical, orthopedic, ophthalmology and otorhinolaryngology departments. The rest were working in community hospitals. The APNs' experience in years ranged from 4-8 years (mean= 5, *SD*=3).

Psychometric results

To test if the tool could differentiate between experts, RNs and nursing students in their clinical reasoning skill in gerontology, the scores of each group were compared. The scores measured by the finalized Geron-NCRS revealed statistically significant variations between groups as calculated by one way ANOVA, ($F(2, 89)=20.09$, $p<0.001$), indicating that not all groups had the same level of clinical reasoning skill. In a Scheffe's post hoc procedure to determine the pair, the group means differed significantly. The findings are revealed in **Tables 3 and 4** and the clinical reasoning score of panellist group (mean= 35.05, *SD*=3.64) showed greater scores than students (mean= 25.01, *SD*=4.25) and RNs (mean= 27.40, *SD*=5.57) ($p<0.001$).

Table 3 Mean and standard deviation among groups of clinical reasoning scores

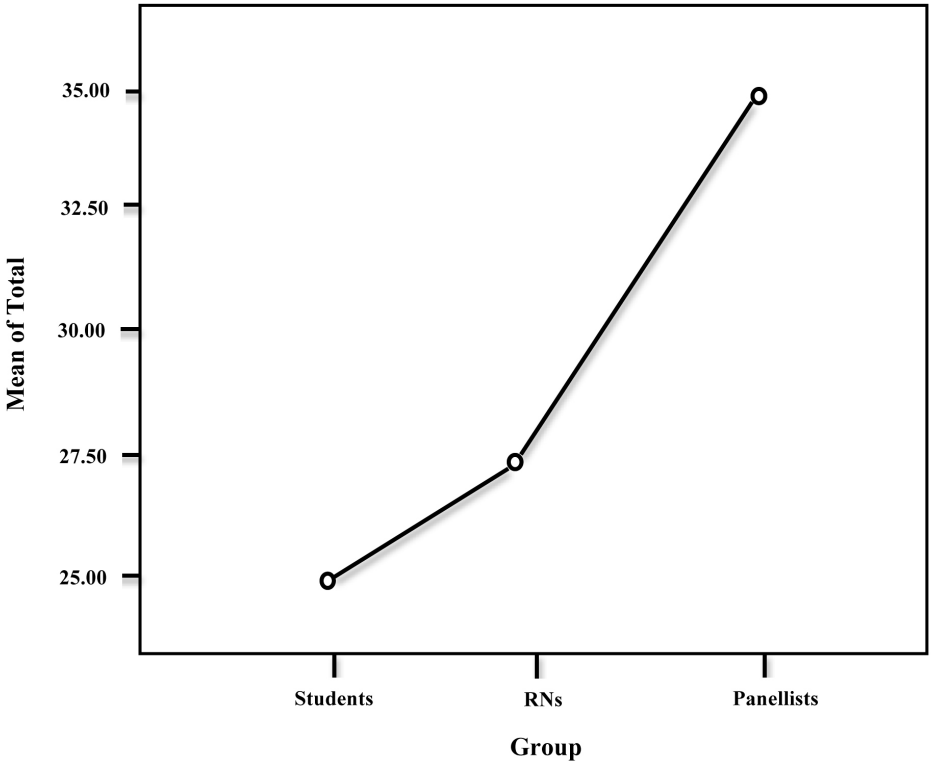
Group	n	Mean	SD
Students	40	25.01	4.25
Nurses	40	27.40	5.57
Panellists	12	35.05	3.64
All	92	27.36	5.73

Table 4 Multiple comparisons among groups calculated by Turkey HSD post hoc test

Group	Group	Mean difference	Std. Error	Sig.	95% confidence Interval	
					Lower Bound	Upper Bound
Students	Nurses	-2.40	1.07	0.91	-5.06	0.29
	Panellists	-10.04*	1.58	0.00	-13.98	-6.09
Nurses	Students	2.38	1.07	0.91	-0.29	5.06
	Panellists	-7.65*	1.58	0.00	-11.59	-3.70
Panellists	Students	10.04*	1.58	0.00	6.09	13.98
	Nurses	7.65*	1.58	0.00	3.70	11.59

*The mean difference is significant at the 0.05 level

There was no statistically significant difference of mean score between students and RNs (p=0.06), displayed in Figure 2. The pass and fail score at -4 and -2 standard deviations (-4*SD* and -2*SD*) of the panellist mean score were 20.49 and 27.77. There were 35 (87.5%) students and 95% (n=38) nurses who had passed the score at the -4*SD* level while 27.5% (n=11) students and 42.5% (n=17) nurses accomplished at the -2*SD* level. The reliability, the Cronbach's alpha coefficient, of the 60-item and 47-item Geron-NCRS were 0.80 and 0.82, respectively. More specifically, individual items of the final scale were started from 0.807-0.826.



The mean difference is significant at the 0.05 level.

Figure 2 Comparison of mean scores among three groups.

Discussion

The assessment of cognitive functioning, especially clinical reasoning competency, is challenging. This study is one of the first to develop and assess the psychometric properties of the 47 items Geron-NCRS that has 20, 9, and 18 items in patient assessment, nursing diagnosis and nursing implementation respectively. It is reasonable to conclude that this tool, which was developed on the Script Theory, is valid and reliable. The content validity was at a satisfactory level of agreement. The internal reliability of this scale was higher than references with a Cronbach's alpha of 0.82 that is similar to other studies.[36] The construct validity was achieved by known group technique [25, 26]

This scale can be an alternative option to test clinical reasoning skills in gerontological nursing at different stage of clinical experiences. Although vignette

construction is time consuming,[37] this scale has some benefits over traditional test methods, especially those which have solving skills with ill-defined problems.[38] As hypothesized, the Geron-NCRS can be used to differentiate scores among panellists, RNs, and nursing students, except between students and RNs. There are some explanations that could clarify similarities of mean score between students and RNs. This phenomenon could be explained by the fact that the scoring grid, designed by experts, might not be suitable for the RNs in this study, who were mostly focused in an institutional setting, sub-specialty.[39, 40] In comparison, the student group had recently studied and was located in more variety of settings especially in the community field before taking the assessment. Some participants in the registered nurse group still had a low number of years (3 years) in their nursing career in a hospital-based setting. They might be perceived as lacking in experience.[41] The GPA in the student group might also be positively influenced by the clinical reasoning in this study.[15]

There are some validation concerns and administration techniques used with this scale that should be discussed. The 47-item Geron-NCRS has 14 vignettes which are lower than the recommendation of 15-25 vignettes. However, the numbers of items per vignettes were satisfactory at 2-4 items.[42] Even though generated and calculated by sufficient panellists, the scoring grid still has some limitations. The distribution of the panellists was of some concern. Eight out of twelve were recruited from a hospital-based background; and it could be argued that more community-based experts should be invited to this study. Besides, there were 30 items that maximum score was located at the +2 option[1] and the least response was more likely to be at the -2 option. Nouh and colleagues have suggested that embracing a 3-point Likert format could be an alternative.[40] In contrast, Wilson and colleagues proposed that a 5 or 6-point Likert format was more reliable than a 3-point Likert format.[43] Thirdly, despite its effectiveness and representation of uncertain situations, the Geron-NCRS still needs to keep its content updated. The agreement among experts might change over time due to the on-going development of knowledge, information, resources, regulation and clinical guidelines. This raises the question about whether the frequency of the vignettes, items revision and scoring re-affirmation need to be updated. Some limitations of this study should be cited. The sample sizes limit the generalizability of findings, and may lead to mismatching and possible variations. The panellists were recruited from different parts of the country, but the students and RNs were particularly based in the northern Thai region which might indicate inconsistency especially in geographical and cultural differences. The application of this scale should also be mentioned. An administration of this measurement can be completed in an hour. However, examinees, who are less familiar with this configuration of examination, might perceive the test format and make decisions based on insufficient information scenarios to be frustrating and confusing, and this might affect their performances and scores. Examinees also might feel the need to be prepared and have the opportunity to try out some example vignettes and items beforehand. This could take approximately 30 minutes for the introductory session.

Conclusion

Thailand is approaching the era that we called aging society. Because of this impending crisis, gerontology nurses are the front line of health care team for older adults in a wide variety of settings. The need for a valid and reliable scale to assess the clinical reasoning skill of nursing personnel in this growing field exists in order to promise the positive patient outcomes. The process of scale development and the findings in this study affirm the validity and reliability of the Geron-NCRS. It also offers a new approach to assess cognitive skills in nursing students. The processes were composed of clarifying definition and elaborating

of the clinical vignette/scenario, generating item pool, choosing the format of measurement, content validation, evaluation of vignettes and items, and examinees' testing and scoring assembly. Exploration into re-calculating the scoring grid is recommended. The ground-breaking approach by using the Script Concordance Test was introduced to Thai nursing professionals in this study. Implications for nursing practice, education, and research should also be mentioned. Nurse leaders or managers could use this scale to assess the RNs' clinical reasoning in providing care to seniors and to prove the impact of the nursing practice on patient safety and outcomes. Besides, nurse educators could routinely apply this scale to evaluate nursing students before they graduate which later can associate with the results of the national nursing license examination. For researchers, future studies are suggested in the use of this scale in summative assessment. This would involve using a different level of nursing students with more participants from various institutes. In this case, preparation of the test into an on-line format to obtain faster results is suggested. Panellists' decisions or responsiveness should be explored in qualitative data, especially if there are items with contradictory responses. Lastly, our scale needs to be tested with different population groups in different settings.

Acknowledgements

This project was funded by the Praboromarajchanok Institute. The authors also extend the warmest thanks to all participants who kindly took part in this study. We are very grateful to all content experts and reference panellists along with valuable comments and manuscript revision of colleagues.

References

1. Hofler L, Thomas K. Transition of new graduate nurses to the workforce: Challenges and solutions in the changing health care environment. North Carolina Medical Journal. 2016;77(2):133-136.
2. Banning M. The think aloud approach as an educational tool to develop and assess clinical reasoning in undergraduate students. Nurse Education Today. 2008;28.
3. Hunter S, Arthur C. Clinical reasoning of nursing students on clinical placement: Clinical educators' perceptions. Nurse Education in Practice. 2016;18:73-9.
4. Levett-Jones T, Hoffman K, Dempsey J, Jeong SY, Noble D, Norton CA, et al. The 'five rights' of clinical reasoning: An educational model to enhance nursing students' ability to identify and manage clinically 'at risk' patients. Nurse Education Today. 2010;30(6):515-20.
5. Simmons B. Clinical reasoning: Concept analysis. Journal of advanced nursing. 2010;66(5):1151-1158.
6. McCartney M. Nurses must be allowed to exercise professional judgment. BMJ. 2017;356:j1548.
7. Goudreau J, Boyer L, Letourneau D. Clinical nursing reasoning in nursing practice: A cognitive learning model based on a think aloud methodology. Quality Advancement in Nursing Education. 2014;1(1):1-18.
8. Stayt LC, Merriman C. A descriptive survey investigating pre-registration student nurses' perceptions of clinical skill development in clinical placements. Nurse Education Today. 2013;33(4):425-430.
9. Tapaneeyakorn V, Kosolchuenvijit J, Anonrath K, Wannasuntad S, Srikan P. Factors affecting clinical reasoning in nursing students of Boromarajonani College of Nursing Bangkok. Journal of Health Science Research. 2016;10(1):70-77.
10. Junod Perron N, Louis-Simonet M, Cerutti B, Pfarrwaller E, Sommer J, Nendaz M. Feedback in formative OSCEs: Comparison between direct observation and video-based formats. Medical Online. 2016;21:32160.
11. Halman S, Dudek N, Wood T, Pugh D, Touchie C, McAleer S, et al. Direct observation of clinical skills feedback scale: Development and validity evidence. Teaching and Learning in Medicine. 2016;28(4):385-394.
12. Orrock P, Grace S, Vaughan B, Coutts R. Developing a viva exam to assess clinical reasoning in pre-registration osteopathy students. BioMed Central Medical Education. 2014;14(1):193.
13. Kelly MA, Mitchell ML, Henderson A, Jeffrey CA, Groves M, Nulty DD, et al. OSCE best practice guidelines-applicability for nursing simulations. Advances in Simulation (London, England). 2016;1:10.

14. Zamani S, Amini M, Masoumi SZ, Delavari S, Namaki MJ, Kojuri J. The comparison of the key feature of clinical reasoning and multiple choice examinations in clinical decision makings ability. Biomedical Research. 2017;28(3):1115-1119.
15. Park WB, Kang SH, Lee YS, Myung SJ. Does objective structured clinical examinations score reflect the clinical reasoning ability of medical students? The American Journal of Medical Sciences. 2015;350(1):64-67.
16. Charlin B, Roy L, Brailovsky C, Goulet F, van der Vleuten C. The Script Concordance test: A tool to assess the reflective clinician. Teaching and Learning Medicine. 2000;12(4):189-195.
17. Lubarsky S, Dory V, Duggan P, Gagnon R, Charlin B. Script concordance testing: From theory to practice: AMEE guide No. 75. Medical Teacher. 2013;35(3):184-193.
18. Kazour F, Richa S, Zoghbi M, El-Hage W, Haddad FG. Using the Script Concordance Test to evaluate clinical reasoning skills in psychiatry. Academic Psychiatry. 2017;41(1):86-90.
19. Wan M. Using the script concordance test to assess clinical reasoning skills in undergraduate and postgraduate medicine. Hong Kong Medical Journal 2015;21(5):455-461.
20. Nseir S, Elkalioubie A, Deruelle P, Lacroix D, Gosset D. Accuracy of script concordance tests in fourth-year medical students. International Journal of Medical Education. 2017;8:63-9.
21. Duggan P, Charlin B. Introducing recent medical graduates as members of Script Concordance Test expert reference panels: What impact? MedEdPublish. 2016;5(2):32-44.
22. Cobb KA, Brown G, Hammond R, Mossop LH. Students' perceptions of the Script Concordance Test and its impact on their learning behavior: A mixed methods study. Journal of Veterinary Medical Education. 2015.
23. Goos M, Schubach F, Seifert G, Boeker M. Validation of undergraduate medical student script concordance test (SCT) scores on the clinical assessment of the acute abdomen. BMC surgery. 2016;16(1):57-.
24. Borges PFB. Script Concordance Test: A proposal for evaluation of the management of nursing reasoning in uncertain situations. Botucatu, Portugal: Universida de Estadual Paulista; 2016.
25. Dawson T. Can script concordance testing be utilized in nursing education to accurately assess clinical reasoning skills?. Cullowhee, NC: Western Carolina University, Cullowhee, NC; 2014.
26. Deschênes MF, Goudreau J. Addressing the development of both knowledge and clinical reasoning in nursing through the perspective of script concordance: An integrative literature review. Journal of Nursing Education and Practice. 2017;7(12):29-38.
27. Peeters C, de Groot-Bolluijt W, Gobbens R, van Brunschot M, editors. Digital Script Concordance Test for Clinical Reasoning. CAA: International Computer Assisted Assessment Conference; 2014; Zeist, The Netherlands: Springer, Cham.
28. Piovezan RD, Custodio O, Cendoroglo MS, Batista NA, Lubarsky S, Charlin B. Assessment of undergraduate clinical reasoning in geriatric medicine: Application of a script concordance test. Journal of American Geriatric Society. 2012;60(10):1946-1950.
29. Ruiz JG, Tunuguntla R, Charlin B, Ouslander JG, Symes SN, Gagnon R, et al. The Script Concordance Test as a measure of clinical reasoning skills in geriatric urinary incontinence. Journal of the American Geriatric Society. 2010;58(11):2178-2184.
30. Pantulawan S, Kunavikitikul W, Nantsupawat, R, Srisuphan W. Development and psychometric testing of the Nursing Performance for Patient Safety Scale. Pacific Rim International Journal of Nursing Research. 2016:20 (1):45-59.
31 Gagnon R, Charlin B, Coletti M, Sauve E, van der Vleuten C. Assessment in the context of uncertainty: How many members are needed on the panel of reference of a script concordance test? Medical Education. 2005;39(3):284-291.
32. Lineberry M, Kreiter CD, Bordage G. Threats to validity in the use and interpretation of script concordance test scores. Medical Education. 2013;47(12):1175-1183.
33. Fournier JP, Demeester A, Charlin B. Script Concordance Tests: Guidelines for construction. BioMed Central Medical Informatics and Decision Making. 2008;8(1):18.
34. Charlin B, Gagnon R, Lubarsky S, Lambert C, Meterissian S, Chalk C, et al. Assessment in the context of uncertainty using the script concordance test: More meaning for scores. Teaching and Learning in Medicine. 2010;22(3):180-186.
35. Duggan P, Charlin B. Summative assessment of 5th year medical students' clinical reasoning by Script Concordance Test: Requirements and challenges. BMC Meical Education. 2012;12(29).

36. Mathieu S, Couderc M, Glace B, Tournadre A, Malochet-Guinamand S, Pereira B, et al. Construction and utilization of a script concordance test as an assessment tool for dcem3 (5th year) medical students in rheumatology. BioMed Central Medical Education. 2013;13(1):166.
37. Boulouffe C, Doucet B, Muschart X, Charlin B, Vanpee D. Assessing clinical reasoning using a script concordance test with electrocardiogram in an emergency medicine clerkship rotation. Emergency Medicine Journal. 2014;31(4):313-316.
38. Humbert AJ, Besinger B, Miech EJ. Assessing clinical reasoning skills in scenarios of uncertainty: Convergent validity for a Script Concordance Test in an emergency medicine clerkship and residency. Academic Emergency Medicine. 2011;18(6):627-34.
39. Charlin B, Tardif J, Boshuizen HP. Scripts and medical diagnostic knowledge: Theory and applications for clinical reasoning instruction and research. Academic Medicine. 2000;75(2):182-190.
40. Nouh T, Boutros M, Gagnon R, Reid S, Leslie K, Pace D, et al. The Script Concordance Test as a measure of clinical reasoning: A national validation study. American Journal of Surgery. 2012;203(4):530-534.
41. Uy RC, Sarmiento RF, Gavino A, Fontelo P. Confidence and information access in clinical decision-making: An examination of the cognitive processes that affect the information-seeking behavior of physicians. AMIA Annual Symposium Proceedings. 2014;2014:1134-1140.
42. Gagnon R, Charlin B, Lambert C, Carriere B, Van der Vleuten C. Script Concordance Testing: More cases or more questions? . Advances in Health Science Education: Theory and practice. 2009;14(3):367-375.
43. Wilson AB, Pike GR, Humbert AJ. Analyzing script concordance test scoring methods and items by difficulty and type. Teaching and Learning in Medicine. 2014; 26(2):135-145.

การพัฒนาและตรวจสอบคุณภาพของแบบประเมินการให้เหตุผลทางคลินิกในการพยาบาลผู้สูงอายุ

ปรัศนี ศรีกัน สิริสุดา เตชะวิเศษ*

บทคัดย่อ: การประเมินทักษะการให้เหตุผลทางคลินิกในการพยาบาลผู้สูงอายุในนักศึกษาพยาบาลนั้นมีความจำเป็น ซึ่งจะส่งผลให้เกิดผลลัพธ์ทางการพยาบาลที่ได้มาตรฐาน แต่ยังขาดเครื่องมือในการประเมินทักษะการให้เหตุผลทางคลินิกในการพยาบาลผู้สูงอายุ วัตถุประสงค์ของการศึกษาครั้งนี้เป็นการพัฒนาและการตรวจสอบคุณภาพของเครื่องมือในการประเมินการให้เหตุผลทางคลินิกในการพยาบาลผู้สูงอายุของนักศึกษาพยาบาล โดยเป็นการศึกษาแบบภาคตัดขวาง (Cross-sectional study) **ในการประเมินผลการพัฒนาเครื่องมือที่สร้างขึ้นมาใหม่โดยมีการสร้างโจทย์สถานการณ์สมมุติ จำนวน 18 สถานการณ์** ซึ่งมีทั้งหมด 72 ข้อ ซึ่งเป็นแบบลิเคิร์ทสเกล (Likert scale) 5 ระดับ ในรูปแบบของการทดสอบความสอดคล้องของสคริปต์ (Script concordance test) โดยใช้ทฤษฎีสคริปต์ (The Script theory) และทฤษฎีสมมุติฐานการอนุมาน (The hypothetical-deductive theory) โดยมีกระบวนการพัฒนาและทดสอบเครื่องมือ 6 ขั้นตอน รวมถึงการตรวจสอบความตรงและความเที่ยงของเครื่องมือ หลังจากนั้นได้ทำการตรวจสอบเครื่องมือกับกลุ่มพยาบาลวิชาชีพและนักศึกษาพยาบาล จำนวน 80 คน แบ่งเป็นกลุ่มละเท่าๆ กัน ส่วนการหาความตรงเชิงโครงสร้างได้มีการทดลองใช้แบบประเมินเปรียบเทียบกับคะแนนของพยาบาลผู้เชี่ยวชาญขั้นสูงด้านการพยาบาลผู้สูงอายุจำนวน 12 คน ซึ่งทำให้ได้แบบประเมินที่มี 14 สถานการณ์ซึ่งมีทั้งหมด 47 ข้อ ผลการเปรียบเทียบคะแนนทักษะการให้เหตุผลทางคลินิกในการพยาบาลผู้สูงอายุ ระหว่างกลุ่มโดยใช้การวิเคราะห์ความแปรปรวนแบบทางเดียว และการเปรียบเทียบความแตกต่างรายคู่โดยวิธีของเชฟเฟ่ (Scheffe' s post hoc test) ผลการศึกษาพบว่าคะแนนเฉลี่ยของพยาบาลผู้เชี่ยวชาญขั้นสูงด้านการพยาบาลผู้สูงอายุ พยาบาลวิชาชีพ และนักศึกษาพยาบาลแตกต่างอย่างมีนัยสำคัญทางสถิติ กล่าวคือพยาบาลผู้เชี่ยวชาญขั้นสูงด้านการพยาบาลผู้สูงอายุมีคะแนนเฉลี่ยสูงสุดตามด้วยพยาบาลวิชาชีพและนักศึกษาพยาบาลตามลำดับ จึงกล่าวได้ว่าประเมินการให้เหตุผลทางคลินิกในการพยาบาลผู้สูงอายุสามารถจำแนกความแตกต่างระหว่างทักษะการให้เหตุผลทางคลินิกในกลุ่มตัวอย่างทั้งสามกลุ่มนี้ ซึ่งมีความตรงทางเนื้อหาและความน่าเชื่อถือสามารถนำไปประยุกต์ในด้านการศึกษาพยาบาล โดยนำแบบประเมินทักษะการให้เหตุผลทางคลินิกในวิชาการพยาบาลผู้สูงอายุสำหรับนักศึกษาพยาบาลได้

Pacific Rim Int J Nurs Res 2019; 23(3) 243-257

คำสำคัญ: การให้เหตุผลทางคลินิก นักศึกษาพยาบาล ผู้สูงอายุ การทดสอบความสอดคล้องของสคริปต์

ติดต่อที่: ปรัศนี ศรีกัน อาจารย์ วิทยาลัยพยาบาลบรมราชชนนี พะเยา*
***E-mail:** pratsanii@gmail.com*
***สิริสุดา เตชะวิเศษ** อาจารย์ วิทยาลัยพยาบาลบรมราชชนนี พะเยา*
***E-mail:** sirisuda.t@bcnpy.ac.th*