Geriatrics & Gerontology INTERNATIONAL Medline Indexed

Geriatr Gerontol Int 2015; **15:** *220–228*

METHODOLOGICAL REPORT: EPIDEMIOLOGY, CLINICAL PRACTICE AND HEALTH

Reliability of the interRAI Long Term Care Facilities (LTCF) and interRAI Home Care (HC)

Hongsoo Kim,[1] Young-Il Jung,[2] Moonhee Sung,[3] Ji-Yoon Lee,[4] Ju-Young Yoon[5] and Jong-Lull Yoon[6]

[1]Graduate School of Public Health and Institute of Health and Environment, [2]Graduate School of Public Health, [3]Graduate School of Public Health, Seoul National University, Seoul, [6]Department of Family Medicine and Geriatrics, Hallym University Hospital, Gyeonggi-Do, South Korea, [4]Department of Nursing, Kangwon National University, Chuncheon, South Korea; and [5]School of Nursing, University of Wisconsin-Madison, Madison, Wisconsin, USA

Aim: Sharing clinical information across care settings is a cornerstone to providing quality care to older people with complex conditions. The purpose of the present study was to examine the reliability of the interRAI Long Term Care Facilities (interRAI LTCF) and the interRAI Home Care (interRAI HC), comprehensive and integrated assessment instruments with common core items, in Korea, an Asian county where comprehensive geriatric assessment is not widely used in long-term care.

Methods: The Korean version of the instruments was developed through field tests, as well as multiple iterations of translations, back-translations and expert reviews. For the reliability test, a random sample of 908 older people in 27 long-term care hospitals or nursing homes, or at home with home care, were assessed by regular staff, among which a subsample of 534 people were dually assessed. The Cronbach's alphas of seven major composite scales in the instruments were examined for internal consistency. Interrater reliability was tested using agreement, kappa coefficients and interclass correlation coefficients.

Results: The internal consistencies of all key measures were adequate (Cronbach's alpha ≥0.75). The overall mean kappa statistics of the items in the interRAI LTCF and those in the interRAI HC were 0.78 and 0.89, respectively. All key common items in the interRAI LTCF and the interRAI HC had almost perfect ($\kappa \geq 0.81$) or substantial ($0.61 \leq \kappa \leq 0.80$) interrater reliability.

Conclusions: The findings show the interRAI LTCF and the interRAI HC have adequate reliability for assessing the function and health of frail older adults across various long-term settings, which can promote continuity of care for the aged. **Geriatr Gerontol Int 2015; 15: 220–228.**

Keywords: continuity of care, geriatric assessment, long-term care, reliability.

Introduction

Because the population is aging rapidly, the need for long-term care (LTC) is increasing. A prerequisite for the provision of quality long-term care is identification of the complex care needs that each individual older person has, for which the usefulness of comprehensive geriatric assessment (CGA) is widely known. Studies have reported CGA improves quality and outcomes of care for frail older people in various care settings.[1,2] Recently, as acknowledgement of the importance of continuity of care has increased, there has also been greater recognition that not only comprehensiveness, but also continuity of relevant geriatric assessment information across care settings is critical.[3] Few instruments have the potential to support integrated as well as comprehensive assessment of older people with LTC needs across various care settings, except the interRAI suite.

Accepted for publication 2 May 2014.

Correspondence: Professor Hongsoo Kim PhD MPH RN, Graduate School of Public Health and Institute of Health and Environment, Seoul National University, Seoul, 151-742, South Korea. Email: hk65@snu.ac.kr

The interRAI suite is an integrated family of assessment instruments developed to be used for comparable assessments for a wide range of vulnerable populations across settings.[3,4] The first interRAI instrument was the Resident Assessment Instrument-Minimum Data Set (RAI-MDS),, a mandatory assessment tool for people in nursing homes in the USA in the 1980s.[5] Instruments for other settings (home care, acute care, community etc.) and populations (hospice, mental health etc.) were developed in the 1990s.[4] These instruments were originally developed to be independent, but the interRAI put effort into redeveloping them by 2008 into an integrated suite providing compatible assessment approaches with a common core set of approximately 70 items (e.g. activities of daily living, cognitive skills etc.) across all instruments and unique items specific to each instrument.[4] There is now a collective set of 19 instruments composing the integrated interRAI assessment system, among which nine instruments are specific to geriatric care. It allows the collection of comparable assessment information about vulnerable older people across various care settings. The assessment system has been used as a mandatory tool in Canada, New Zealand, Ireland and other countries, and is also widely used for within- and cross-country research in Europe.

Korea has an urgent need for adopting such comparable geriatric assessment tools. As a country with a rapidly aging population, Korea introduced public long-term care insurance in 2008 to meet the increasing need for LTC services.[6] LTC in Korea is provided in both institutional (LTC hospitals or facilities) and community (home care) settings, but there are no common CGA tools for LTC beneficiaries that can be used across the settings, which would enable the examination and comparison of resident conditions and services provided at the different settings.

Good psychometric properties of the RAI-MDS and the RAI-Home Care (RAI-HC), earlier versions of the interRAI Long Term Care Facilities (interRAI LTCF) and the interRAI Home Care (interRAI HC), respectively, were reported in several studies before the interRAI suite was introduced.[7–10] Only Hirdes *et al*. have examined the interrater reliability of five instruments of the interRAI suite including the interRAI LTCF and the interRAI HC.[4] They found an overall kappa mean value of 0.75 for the common items appearing on two or more instruments, and kappa ranged between 0.63 and 0.73 for the items that appeared on just one of the five instruments.[4] The psychometric properties of instruments can vary in different cultures and contexts. The present study aimed to evaluate the reliability of the common and unique items of interRAI LTCF and the interRAI HC across both institutional and home care settings in Korea, an Asian country where CGA is not widely used.

Methods

Study design and participants

This is a cross-sectional, psychometric study. A total of 908 participants were randomly selected from rosters provided by a total of 27 LTC service organizations, including 17 LTC institutions where the interRAI LTCF applied and 10 home care agencies where the interRAI HC applied for research purposes. All were located in metropolitan areas – Seoul, Gyeonggi and Incheon – as metropolitan areas are where the majority of LTC institutions in South Korea are located.[11] A subsample of 534 individuals randomly selected from the total sample was also dually assessed to examine interrater reliability. There was no difference between the subsample and the total in the characteristics listed in Table 1, except comorbidity in the LTCF sample: the proportion of dementia patients with or without stroke in the subsample (61.0%) was significantly higher than that of its counterpart (43.8%, not shown). The older adults receiving care in the three types of LTC settings represented a wide range of health and functional statuses, and potential for improvement.

Instruments

The interRAI LTCF and the interRAI HC (2009 version), tested in the present study, are the two most widely used instruments in the interRAI suite. The interRAI LTCF and the interRAI HC are independent comprehensive geriatric assessment systems whose items cover key domains for services provided as well as the health and function of older people, including activities of daily living, mood, psychosocial well-being, problem behaviors and health conditions.[3] More detailed explanations of the instruments are elsewhere.[12,13] These assessment instruments were developed following the guidance of certain key design principles.[4] The assessments were to be completed using all available sources of information, and judgments were to be made based on observable traits within a clear time frame for observations; specific operational definitions and coding instructions with inclusion and exclusion criteria were provided for each assessment item; assessments using these instruments were intended to be useful for multiple stakeholders by supporting care planning, outcome measurement, quality improvement and resource allocations, for which evidence-based decision-support applications based on items in the tools, such as scales, quality indicators, case-mix groups and so on, were also developed; and each instrument was to be independent for a specialized care population or setting, but at the same time part of a parallel and integrated assessment system.

The interRAI LTCF and the interRAI HC have subscales, among which the following six core scales,

Table 1 General characteristics of participants

	interRAI LTCF (n = 621) n (%)	interRAI HC (n = 287) n (%)
Age (years)		
65–74	150 (24.2)	78 (27.2)
75–84	285 (45.9)	133 (46.3)
≥85	186 (30.0)	76 (26.5)
Mean ± SD	80.0 ± 7.5	79.3 ± 7.5
Sex		
Male	176 (28.3)	115 (40.1)
Female	445 (71.7)	172 (59.9)
Marital status†		
Married	188 (30.3)	144 (50.4)
Widowed	402 (64.8)	140 (49.0)
Other	30 (4.8)	2 (0.7)
Chronic conditions†		
Stroke	163 (26.3)	89 (31.1)
Dementia	231 (37.3)	49 (17.1)
Stroke and dementia	115 (18.6)	39 (13.6)
None of the above	111 (17.9)	109 (38.1)
Institution type		
LTC hospital	298 (48.0)	
LTC facility	323 (52.0)	

†These variables have one case with missing values. InterRAI HC, interRAI Home Care, interRAI LTCF, interRAI Long Term Care Facilities; LTC, long-term care.

known key descriptors of frail older people in LTC and geriatric literature,[14] were selected: activities of daily living (ADL), instrumental activities of daily living (IADL; performance and capacity), cognition, depression, communication and pain. The ADL hierarchy scale is a seven-level measure of functional performance based on four items: personal hygiene, locomotion, toilet use and eating.[15] Using an algorithm, severity patterns of the variables categorize people from 0 (independent) to 6 (total dependence). Two IADL scales are embedded in the interRAI HC: the IADL performance scale measures what a person actually does, whereas the IADL capacity scale measures what an individual is capable of doing.[8] The scales are based on eight items – meal preparation, housework, finance management, medication management, phone use, stairs, shopping and transportation – in the interRAI HC, and higher scores indicate more difficulties in performance or capacity.

The depression rating scale (DRS) describes the mood status of an individual. Using seven mood-related items common in the two interRAI tools, the DRS ranges from 0 to 14.[16] The communication scale ranges from 0 to 8, and encompasses two four-point items: make self understood and ability to understand others.[17] A higher communication score reflects poorer communication. Finally, the five-level pain scale summarizes pain frequency and intensity: a higher pain score indicates higher pain.[18] More detailed information on the scales is provided in existing literature.

Translation

According to the standardized procedure of interRAI for developing translated versions of the instruments, the first author made an official contract with interRAI for instrument development. Guided by the official World Health Organization process of translation and adaptation of research instruments,[19] the Korean versions of the interRAI LTCF and the interRAI HC were developed through iterations of the following steps. An initial forward translation from the English version of the interRAI LTCF and the interRAI HC to a Korean version was carried out by a translator who had knowledge of geriatric and long-term care, and was also fluent in both Korean and English. Second, the translated versions were reviewed and commented on by a bilingual panel consisting of a geriatrician, three nurses, a social worker, and two health services researchers who had expertise in the topic and instrument development, and also by another expert panel of experienced LTC practitioners. In this process, some items and wording

were modified in consideration of their relevance to the culture, language and LTC context of Korea. Third, using the same approach as in the aforementioned forward translation, a backward translation of the instruments from Korean to English was independently carried out by a team of two Korean Americans who had a medical and social work background, respectively, and were versed in Korean and English, but who had no knowledge of the instrument. Fourth, the developed instruments were pretested and debriefed, based on which items were revised for clarity. The total numbers of items in the Korean version of the interRAI LTCF and the interRAI HC were 274 in 19 sections and 267 in 20 sections, respectively, among which approximately three-quarters of the items (n = 204) were common.

Procedure

The assessments were carried out by regular staff at each institution, who were trained in the assessment procedure by the research team in a 3- to 4-h session, with a training package including the assessment manual, standardized PowerPoint presentations and survey protocols. The assessors were registered nurses and/or social workers with general knowledge of resident assessment, and were familiar with the conditions of the older people for whom they provided care. Assessors were trained and encouraged to use various information sources including resident interviews, direct observation, family or close relatives or friends, coworkers and chart review; the final ratings were based on their own clinical judgments. Dual assessments for interrater reliability were required to be completed independently within 3 days.[4] Dual assessors were prohibited from discussing the case with each other, asked not to exchange information and were blinded to the results of others. The median time for completing an assessment was approximately 40 min. Medication data were mostly provided on a separate copy of the medication order sheet. Informed consent was sought from participants or their proxies. The study was approved by the institutional review board for human subject research at the institution with which the corresponding author who carried out the data collection was affiliated.

Analysis

The internal consistencies of the six key subscales in the newly developed tools were evaluated using Cronbach's alpha statistics: over 0.70 is excellent reliability. Interrater reliability of the items was estimated, and the average kappa statistics by section and the percentage of items whose kappa statistics were 0.40 or higher in each section were summarized. The reliability of key common individual or groups of items, which are crucial for describing the function of individuals with long-term care needs, were also examined using agreement, unweighted Cohen's kappa (for nominal variables), weighted Cohen's kappa and the intraclass correlation coefficient (ICC) with a two-way random model for absolute agreement (for ordinal variables). Following Landis and Koch,[20] the extent of agreement using kappa values was evaluated: slight from 0 to 0.20; fair from 0.21 to 0.40; moderate from 0.41 to 0.60; substantial from 0.61 to 0.80; and almost perfect from 0.81 to 1.0. An ICC of 0.40 or higher was considered adequate reliability; an ICC or 0.70 or higher, excellent reliability.[21] We report reliability of interRAI LTCF from a combined sample of LTC hospitals and LTC facilities, because when they were separated, all reliability statistics only differed at the second decimal level, except a few items whose reliability was still substantial in both settings. Some items were not relevant for the reliability test (e.g. assessor's signature) and others had little or no variance, so kappa statistics were not estimable.

Results

The mean age of the institutionalized sample was 80.0, and the majority were female (71.7%), widowed (64.8%), and had stroke and/or dementia that had been diagnosed by physicians (82.2%; Table 1). Compared with the institutionalized sample, the home care participants were more likely to be men, married and have stroke. The internal consistency of the core subscales of the interRAI LTCF and the interRAI HC examined in the present study was excellent (Table 2). Among the institutionalized older people, the communication scale had the highest Cronbach's alpha statistics (0.96), and the depression scale had the lowest (0.82). The home care recipients had a similar pattern.

The grand mean (and median) kappa of common and unique items in the interRAI LTCF were 0.78 (0.79), and those in the interRAI HC were 0.89 (0.91; Table 3). Almost the same mean and median kappas of the common items only were observed in the interRAI LTCF (0.79 and 0.80) and the interRAI HC (0.89 and 0.91). The kappa statistics of all tested items in the interRAI LTCF were 0.41 or higher, except two items. The average kappa by section ranged between 0.64 (Responsibility) and 0.95 (Continence). For the interRAI HC, all tested items except one in the Health Condition section attained kappas of 0.41 or higher, ranging by section from 0.81 (Environmental Assessment) to 0.98 (Identification Information). According to Landis and Koch, the reliability of the 221 items tested in the interRAI LTCF was almost perfect for 40.3%, substantial for 53.8%, moderate for 5.0% and fair for 0.9%; for the 205 items of the interRAI HC, it was almost perfect for 89.5%, substantial for 12.2%, moderate for 0.98% and fair for 0.49% (not shown).[20]

Table 2 Internal consistency of core measures in the interRAI LTCF and interRAI HC

		interRAI LTCF (n = 621)	interRAI HC (n = 287)	All
	No. items			
ADL	4	0.91	0.94	0.92
IADL performance	8		0.93	0.93
IADL capacity	8		0.94	0.94
Depression	7	0.82	0.76	0.80
Communication	2	0.96	0.96	0.96
Pain	2	0.88	0.93	0.90

Internal consistency was measured by Cronbach's alpha. ADL, activities of daily living; IADL, instrumental activities of daily living; interRAI HC, interRAI Home Care, interRAI LTCF, interRAI Long Term Care Facilities.

Table 4 shows the average agreement and reliability statistics of key common items in the interRAI LTCF and the interRAI HC. First, the mean agreement statistics of the items ranged from 0.79 to 1.0 in the institutionalized sample, and were 0.84 or higher in the home care sample. All key single or groups of items listed in the table had substantial or almost perfect kappas in institutional and home care settings. The ICC for ordinal variables were all aligned with the kappa statistics, but were somewhat higher for most items. ICC were generally higher in home care settings than in institutional, as were the kappa statistics.

Discussion

It can be challenging to obtain reliable data on the functional and mental status of frail older adults. As comprehensive and integrated assessment systems for people with chronic care needs, the interRAI LTCF and the interRAI HC were designed to standardize data collection and assessment in order to increase the reliability of the data. The present study found that the interRAI LTCF and interRAI HC are reliable assessment instruments for older people with long-term care needs in both institutional and home care settings in Korea. The eight key composite measures observed all had excellent internal consistency among Korean older adults receiving LTC. Almost all items in the interRAI LTCF (93.8%) and in the interRAI HC (98.5%) achieved almost perfect or substantial reliability, which shows these instruments are adequate for practice and research purposes in Korean LTC settings. Hirdes *et al.* reported the average kappas of the common items in the interRAI LTCF and the interRAI HC to be 0.74 and 0.69, respectively; the comparable statistics in the present study were 0.79 for the interRAI LTCF and 0.89 for the interRAI HC.[4] Interrater reliability was higher in the home care setting than in the institutional setting in the present study, which could be due to several causes. Institutionalized older people are more likely to be impaired in cognition and communication skills; assessing them is more challenging. Some home care cases were assessed simultaneously by the paired independent raters because of the limited visitation schedule during the observation period for the assessment. More than half of the home care participants had been receiving visits from the assessors for more than 1 year; the assessors knew very well the health conditions and care needs of the participants.

As for individual items, the kappas of all the items except two in the interRAI LTCF and one item in the interRAI HC were moderate or better (≥0.41), which are similar findings to that of Hirdes *et al.*, reporting just four items that had kappas below 0.40 across the five instruments they observed.[4] The two items with relatively low kappa values in the institutional setting were gastro-intestine or gastro-urethral bleeding (agreement = 0.99; kappa = 0.33) and blood pressure measure in last year (agreement = 0.95; kappa = 0.37). The former had a much lower prevalence (prevalence index [PI] = –0.99; bias index [BI] = 0.00), and the latter had quite a high prevalence (PI = 0.93; BI = –0.02), which contributed to high agreement and lower kappa statistics, so-called "kappa paradoxes";[22] the prevalence-adjusted, bias-adjusted kappas[23]of the two items were excellent at 0.98 and 0.91, respectively. In the home care setting, the only item whose kappa value was below 0.41 was vomiting (agreement = 0.99; kappa = 0.40). This is a five-level ordinal variable, so we cannot calculate PI and BI; but fewer than 5% of the patients had the symptom, which resulted in a relatively low kappa with high agreement. More attention to the coding of these items should be addressed in the training session.

Table 3 Interrater reliability of the interRAI LTCF and interRAI HC by section

Section	interRAI LTCF (n = 434)				interRAI HC (n = 100)			
		Items for which reliability estimates were available				Items for which reliability estimates were available		
	Total items	No. items	Percent that were reliable (kappa ≥.41)	Average reliability	Total items	No. items	Percent that were reliable (kappa ≥.41)	Average reliability
Identification information	16	9	100	0.91	21	9	100	0.98
Intake and initial history	15	10	100	0.90	9	2	100	0.90
Cognition	10	10	100	0.78	9	9	100	0.91
Communication and vision	6	6	100	0.82	4	4	100	0.88
Mood and behavior	20	20	100	0.71	20	20	100	0.88
Psychosocial well-being	19	19	100	0.74	10	10	100	0.92
Functional status	19	18	100	0.83	37	34	100	0.89
Continence	4	4	100	0.95	4	4	100	0.90
Disease diagnoses	22	21	100	0.75	22	19	100	0.91
Health conditions	36	35	97.1	0.74	36	34	97.1	0.87
Oral and nutritional status	14	11	100	0.82	11	8	100	0.84
Skin condition	7	7	100	0.80	7	5	100	0.92
Activity pursuit (LTC only)	18	18	100	0.73	0	0		
Medications	2	1	100	0.91	3	2	100	0.96
Treatments and procedures	47	23	95.7	0.81	41	18	100	0.90
Responsibility	10	5	100	0.64	1	0		
Social supports (HC only)	0	0			13	12	100	0.94
Environmental assessment (HC only)	0	0			10	10	100	0.81
Discharge potential	4	4	100	0.74	5	5	100	0.93
Discharge	3	0			2	0		
Assessment information	2	0			2	0		
Mean (median) kappa for total items	274	221		0.78 (0.79)	267	205		0.89 (0.91)
Mean (median) kappa for common items	204	168		0.79 (0.80)	204	153		0.89 (0.91)

HC, home care; interRAI HC, interRAI Home Care, interRAI LTCF, interRAI Long Term Care Facilities; LTC, long term care.

Table 4 Average interrater reliability of key common items in the interRAI LTCF and interRAI HC

Key section and item	interRAI LTCF No. Items†	Average percentage agreement	Average kappa	Average intraclass correlation coefficient	interRAI HC No. items†	Average percentage agreement	Average kappa	Average intraclass correlation coefficients
Identification Information								
Sex	1	1.00	1.00		1	1.00	1.00	
Marital status	1	0.95	0.89		1	0.98	0.96	
Cognition								
Cognitive skills	1	0.84	0.86	0.92	1	0.89	0.93	0.96
Memory/recall ability	4	0.90	0.77		3	0.95	0.90	
Periodic thinking disorder	3	0.89	0.77		3	0.95	0.88	
Acute change in mental status	1	0.94	0.82		1	0.98	0.94	
Communication and Vision								
Understanding/Communication	2	0.79	0.84	0.91	2	0.84	0.87	0.92
Hearing	1	0.79	0.78	0.84	1	0.90	0.90	0.93
Vision	1	0.79	0.74	0.82	1	0.93	0.90	0.92
Mood and Behavior								
Mood distress	11	0.86	0.67	0.68	11	0.93	0.86	0.85
Self-reported mood	3	0.84	0.72	0.74	3	0.91	0.87	0.85
Behavior symptoms	6	0.93	0.80	0.83	6	0.99	0.90	0.88
Psychosocial Well-Being								
Social relationships	3	0.88	0.76	0.75	6	0.97	0.93	0.93
Major life stressors	1	0.96	0.78		1	0.99	0.97	
Functional Status								
ADL self-performance	10	0.80	0.85	0.92	10	0.87	0.92	0.95
Continence								
Bladder continence	1	0.86	0.88	0.91	1	0.90	0.89	0.88
Bowel continence	1	0.89	0.91	0.95	1	0.92	0.91	0.93
Disease Diagnoses	21 (20)	0.96	0.75		21 (18)	0.98	0.90	
Health Conditions								
Falls	1	0.99	0.92	0.93	1	1.00	1.00	1.00
Pain symptoms‡	5	0.88	0.78	0.82	4	0.92	0.90	0.94
Instability of conditions	3	0.92	0.78		3	0.97	0.93	
Self-reported health	1	0.79	0.67	0.68	1	0.93	0.81	0.73
Oral and Nutritional Status								
Height	1			0.99	1			0.99
Weight	1			0.97	1			0.99
Nutritional status	4 (3)	0.97	0.83		3	0.96	0.70	
Mode of nutritional status	1	0.89	0.91	0.93	1	0.93	0.93	0.91
Skin Condition								
Pressure ulcer stage	1	0.97	0.83	0.87	1	0.98	0.94	0.88
Medications								
Allergy to any drug	1	1.00	0.91		1	1.00	1.00	
Treatments and Procedures								
Prevention	8	0.96	0.77		7	0.98	0.93	
Treatments	11 (9)	0.99	0.87	0.87	11 (7)	0.98	0.90	0.88
Programs	3	0.97	0.79	0.85	3	0.96	0.86	0.83
Hospital use	1			0.73	1			1.00
ED use	1			0.83	1			0.99
Physical restraint use	3	0.90	0.79	0.80	1	0.95	0.67	

†A few domains have items with too low prevalence or no variation, so reliability estimation was not possible. Thus, for such domains, the actual numbers of items whose reliability was tested in this study are provided in the brackets. ‡One of the five items measuring pain symptoms was binary, so the intraclass correlation coefficient statistic presented is the average of the intraclass correlation coefficient statistics for the other four ordinary items.
ADL, activities of daily living; ED, emergency department; interRAI HC, interRAI Home Care, interRAI LTCF, interRAI Long Term Care Facilities.

The key common items of the interRAI LTCF and the interRAI HC for measuring the function of frail older people had excellent reliability. Mood distress and self-reported mood had relatively lower reliability than others in the institutional setting, which was a consistent finding with that of existing studies.[4,24] This might be because mood can change easily even in one day, and it could also be hard to measure in older people with cognitive impairment and difficulties in communication. Environments where older people can report their mood status candidly, as well as more careful assessment, are necessary. In the training session, skill building and strategy sharing for assessing challenging cases would be useful, encouraging the use of objective observations and information gathering from other relevant sources. Further studies are also suggested on the reliability and validity of items to measure mood.

This is the first study to develop and test an integrated health and function assessment system, using the so-called third generation assessment instruments, for older people in Korea.[3] The study found that the interRAI LTCF and the interRAI HC are reliable tools for assessing the health and function of frail older people in a LTC context where CGA has not been widely used. Still another strength of the present study was that the reliability test was carried out by regular staff in the actual practice setting, rather than by researchers with more advanced assessment skills. We also drew samples from both institutional and home care settings, and compared the reliabilities of key common items in the tools tested in both settings. The study also had limitations. Some items in our sample had no or little variance, so their reliability could not be tested; and others had very low prevalence, so their kappa estimates could be unstable. Further studies on the validity of the key scales are required.

Studies in Western countries have reported positive effects of implementation of interRAI tools, including their applications for care planning and patient placement, on length of stay, quality of care and/or patient outcomes.[2,25,26] Studies evaluating the impacts of adopting the newly developed, evidence-based, integrated CGA tools on the quality of care and quality of life of frail older people in Korea are recommended. Cross-national studies on quality, safety, and outcomes of long-term care would be also beneficial.

Acknowledgments

This research was supported by the National Research Foundation of Korea(NRF) Grant funded by the Korean Government(MSIP)(Grant number: 2010-0002802), and also a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute(KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea(Grant number: HI13C2250). The funding source had no role in the study design; in the collection, analysis and interpretation of data; in the writing of the manuscript; or in the decision to submit the manuscript for publication.

Disclosure statement

The authors declare no conflict of interest.

References

1 Ellis G, Whitehead MA, O'Neill D, Langhorne P, Robinson D. Comprehensive geriatric assessment for older adults admitted to hospital. *Cochrane Database Syst Rev* 2011; (7) CD006211.
2 Mor V, Intrator O, Fries BE *et al*. Changes in hospitalization associated with introducing the Resident Assessment Instrument. *J Am Geriatr Soc* 1997; **45**: 1002–1010.
3 Gray LC, Berg K, Fries BE *et al*. Sharing clinical information across care settings: the birth of an integrated assessment system. *BMC Health Serv Res* 2009; **9**: 71.
4 Hirdes JP, Ljunggren G, Morris JN *et al*. Reliability of the interRAI suite of assessment instruments: a 12-country study of an integrated health information system. *BMC Health Serv Res* 2008; **8**: 277.
5 Morris JN, Hawes C, Fries BE *et al*. Designing the national resident assessment instrument for nursing homes. *Gerontologist* 1990; **30**: 293–307.
6 Seok JE. Public long-term care insurance for the elderly in Korea: design, characteristics, and tasks. *Soc Work Public Health* 2010; **25**: 185–209.
7 Hawes C, Morris JN, Phillips CD, Mor V, Fries BE, Nonemaker S. Reliability estimates for the Minimum Data Set for nursing home resident assessment and care screening (MDS). *Gerontologist* 1995; **35**: 172–178.
8 Landi F, Tua E, Onder G *et al*. Minimum data set for home care: a valid instrument to assess frail older people living in the community. *Med Care* 2000; **38**: 1184–1190.
9 Mor V, Angelelli J, Jones R, Roy J, Moore T, Morris J. Inter-rater reliability of nursing home quality indicators in the US. *BMC Health Serv Res* 2003; **3** (1): 20.
10 Morris JN, Fries BE, Steel K *et al*. Comprehensive clinical assessment in community setting: applicability of the MDS-HC. *J Am Geriatr Soc* 1997; **45**: 1017–1024.
11 Korea National Health Insurance Services. *Long-Term Care Insurance Statistical Yearbook 2012*. Seoul: NHIS, 2013.
12 Mor V, Finne-Soveri H, Hirdes J, Gilgen R, Dupasquier J. Long term care quality monitoring using the interRAI common clinical assessment language. In: Smith PC, Mossialos E, Papanicolas I, Leatherman S, eds. *Performance Measurement for Health System Improvement: Experience, Challenges, and Prospects*, 1st edn. Cambridge: Cambridge University Press, 2009; 472–505.
13 OECD /European Commission. *A Good Life in Old Age? Monitoring and Improving Quality in Long-term Care – OECD Health Policy Studies*. Paris: OECD Publishing, 2013.
14 Casten R, Lawton MP, Parmelee PA, Kleban MH. Psychometric characteristics of the minimum data set I: confirmatory factor analysis. *J Am Geriatr Soc* 1998; **46**: 726–735.
15 Morris JN, Fries BE, Morris SA. Scaling ADLs within the MDS. *J Gerontol A Biol Sci Med Sci* 1999; **54** (11): M546–M553.
16 Burrows AB, Morris JN, Simon SE, Hirdes JP, Phillips C. Development of a minimum data set-based depression

rating scale for use in nursing homes. *Age Ageing* 2000; **29**: 165–172.
17 Frederiksen K, Tariot P, De Jonghe E. Minimum Data Set Plus (MDS+) scores compared with scores from five rating scales. *J Am Geriatr Soc* 1996; **44**: 305–309.
18 Fries BE, Simon SE, Morris JN, Flodstrom C, Bookstein FL. Pain in US Nursing Homes Validating a Pain Scale for the Minimum Data Set. *Gerontologist* 2001; **41**: 173–179.
19 World Health Organization. Management of substance abuse: process of translation and adaptation of instruments, 2014 [Cited 4 June 2014.] Available from URL: http://www.who.int/substance_abuse/research_tools/en/.
20 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–174.
21 Fless JL. *The Design and Analysis of Clinical Experiment*. New York: John Wiley and Sons, 1986.
22 Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. *J Clin Epidemiol* 1990; **43**: 543–549.
23 Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol* 1993; **46**: 423–429.
24 Sgadari A, Morris JN, Fries BE *et al*. Efforts to establish the reliability of the Resident Assessment Instrument. *Age Ageing* 1997; **26** (Suppl 2): 27–30.
25 Boorsma M, Frijters DH, Knol DL, Ribbe ME, Nijpels G, van Hout HP. Effects of multidisciplinary integrated care on quality of care in residential care facilities for elderly people: a cluster randomized trial. *CMAJ* 2011; **183** (11): E724–E732.
26 Hirdes JP, Poss JW, Caldarelli H *et al*. An evaluation of data quality in Canada's Continuing Care Reporting System (CCRS): secondary analyses of Ontario data submitted between 1996 and 2011. *BMC Med Inform Decis Mak* 2013; **13** (27). doi: 10/1186/1472-69-47-13-27