Development and Validation of the Standard Chinese Version of the CARE Item Set (CARE-C) for Stroke Patients

Ke-Yin Chang, MD, PhD, Chen-Yu Hung, MD, Chien-Wei Kao, MS, Fuk-Tang Tan, MD, Barbara Gage, PhD, Ching-Lin Hsieh, PhD, Tyng-Guey Wang, MD, and Der-Sheng Han, MD, PhD

Abstract: The Continuity Assessment Record and Evaluation (CARE) item set is a standardized, integrative scale for evaluation of functional status across acute and postacute care (PAC) providers. The aim of this study was to develop a Chinese version of the CARE (CARE-C) item set and to examine its reliability and validity for assessment of functional outcomes among stroke patients.

The CARE-C was administered in two samples. Sample 1 included 30 stroke patients in the outpatient clinic setting for the purpose of examining interrater and test–retest reliabilities and internal consistency. Sample 2 included 138 stroke patients admitted to rehabilitation units for the purpose of investigating interrater and test–retest reliabilities and internal consistency. Six of the 11 subscales met acceptable levels of internal consistency (Cronbach’s alpha > 0.7). The criterion-related validity of the CARE-C showed moderate to high correlations of its subscales of cognition and basic and instrumental activities of daily living with the Barthel index, IADL scale, and MMSE.

The CARE-C is a useful instrument for evaluating functional quality metrics in the Chinese stroke population. The development of the CARE-C also facilitates the assessment of the PAC program in Taiwan and future research is warranted for validating the capability of CARE-C to identify patients’ functional change over time and its generalizability for nonstroke populations.

Abbreviations: BAdL = basic activities of daily living, CARE = the Continuity Assessment Record and Evaluation, CARE-C = the Chinese version of the Continuity Assessment Record and Evaluation, EQ-5D = EuroQOL five dimensions questionnaire, HHA = home health agencies, IADL scale = Lawton Instrumental Activities of Daily Living scale, ICC = intraclass correlation coefficient, IRF = inpatient rehabilitation facilities, LTCH = long-term care hospitals, MMSE = Mini-Mental State Examination, PAC = postacute care, SD = standard deviation, SNF = skilled nursing facilities.

INTRODUCTION

Postacute care (PAC), which refers to medical care services that support the individual’s recovery from illness or management of chronic disability, aims to enhance the functional status of patients discharged from acute hospitalization. This advance in the medical and care professions contributes to an increase in the number of PAC providers, which also relieves the heavy burden imposed on the acute care system. Medicare, a national social insurance program in the United States, offers health insurance for Americans aged 65 years or older and pays for PAC settings, including skilled nursing facilities (SNF), inpatient rehabilitation facilities (IRF), home health agencies (HHA), and long-term care hospitals (LTCH). However, the evaluation scales used in each setting differ by PAC providers. Each setting has its own instruments and own scales: IRFs use the Patient Assessment Instrument, SNFs use the Minimum Data Set, and HHAs use the Outcome and Assessment Information Set. The discrepancies between each assessment tool causes difficulty in comparing patients’ outcomes across PAC providers and potentially results in inconsistent payment standards. Therefore, a single, standardized set of assessment items, the Continuity Assessment Record and Evaluation (CARE) item set, was developed by the Centers for Medicare and Medicaid Services for evaluation of functional status across the various acute and postacute services.

The CARE item set contains four domains, including medical status клинический комплекс, functional status, cognitive status, and social support factors and is released in 4 versions: institutional admission, home health admission, discharge, and expired. Stroke is the third leading cause of death in Taiwan. It
is estimated that 80,000 new or recurrent stroke cases per year spent US$375 million in medical cost in Taiwan. Because the Administration of National Health Insurance in Taiwan implemented a PAC program for stroke participants, the largest population with disability needing continued medical, nursing, and rehabilitation services, we found it necessary to develop a Chinese-based, integrative, easily performed scale to assess functional outcome at different admission stages and facilities. Hence, the aim of this study was to develop a Chinese version of the CARE item set (CARE-C) and examine whether it is reliable and valid for evaluating functional outcomes among stroke patients across different medical care providers.

**METHODS**

**Translation and adaptation of the CARE questionnaire**

Five board-certificated rehabilitation physicians, who were fluent in both Chinese and English, translated the English version of the CARE questionnaire into Chinese. The primary investigator reviewed and revised the first translated Chinese prototype and convened 2 expert committees to formulate an expert-consensus version. The members in the expert meetings, including rehabilitation physicians, physical therapists, occupational therapists, speech therapists, nursing professionals, and public health specialists, discussed which items would be retained or modified for adaptation to differences in cultures and medical systems. The expert-consensus copy was finalized as the CARE-C after approval by all the members of the expert committee and was then pilot tested on a sample of stroke patients.

**Sampling**

The interrater and test–retest reliability and internal consistency were examined in Sample 1 which included 30 adult stroke patients recruited from a community hospital. Those participants underwent regular follow-up in the outpatient clinic and were chronically stable with the most recent cerebral vascular insult recorded more than 2 years previously. Stroke was diagnosed by neurologists based on clinical criteria—the rapid development of signs of a focal disturbance of cerebral function lasting over 24 hours without an apparent nonvascular cause. The stroke type, ischemic or hemorrhagic, was recorded from neuroimage reports, including computer tomography (CT) or magnetic resonance imaging (MRI). Interrater reliability was tested independently between 1 nurse and 1 physiotherapist, both of whom had more than 5 years of working experience and were familiar with the test instructions for the CARE-C. The test–retest reliability was assessed by the nurse with an interval of 1 week between the first and second visits.

The validity of the CARE-C was evaluated in Sample 2 which included 138 stroke patients, who were recruited from the rehabilitation wards of 7 representative hospitals, including 5 medical centers and 2 regional hospitals that were located in the north, central, south, and east districts of Taiwan. The inclusion criteria consisted of adults aged more than 20 years and a recent cerebral vascular insult within 3 months before recruitment. After the patients were admitted to the target hospital, a group of case managers who owned nursing profession and had received an 1-day course of CARE-C assessment instruction, collected the data from the CARE-C, Barthel index, Lawton Instrumental Activities of Daily Living (IADL) scale, EuroQOL five dimensions questionnaire (EQ-5D), and Mini-Mental State Examination (MMSE) within 3 days after admission. The study was approved by the Research Ethics Committee in National Taiwan University Hospital (No. 201406020RINB), and written informed consent was obtained from all of the participants.

**Instrument for Criterion-Related Validity**

**Barthel Index of Activities of Daily Living**

The Barthel index is an ordinal scale used to evaluate 10 items of activities of daily living: feeding, bathing, grooming, dressing, bowel control, bladder control, toileting, chair transfer, ambulation, and stair climbing. The scores range from 0 to 100; a higher score indicates a better degree of independence in activities of daily living.

**Lawton Instrumental Activities of Daily Living (IADL) Scale**

The IADL scale assesses 8 items of instrumental activities of daily living: the ability to use a telephone, shopping, food preparation, housekeeping, laundry, mode of transportation, responsibility for own medications, and ability to handle finance. Historically, women are scored on all 8 categories of function, whereas the areas of food preparation, housekeeping, laundering are usually omitted for men. Due to evolution of views on gender equality, current public opinion favors that men should be scored like women. Therefore, we assessed all 8 items for both women and men. The summary score ranges from 0 (the lowest function) to 8 (the highest function).

**EQ-5D**

The EQ-5D is a widely used questionnaire for evaluation of health status in 5 dimensions: mobility, self-care, usual activities, pain/discomfort, and anxiety/depression. Each item is responded to in 3 levels: 1 for no problem, 2 for moderate problem, and 3 for severe problem. The summarized responses are then transformed into a preference-based score where 1 stands for the best possible health status and 0 indicates death.

**MMSE**

The MMSE is the most commonly used instrument to assess cognitive status in adults. It tests orientation, attention, memory, language, and visual–spatial skills. The maximal score is 30 points and a score below 27 is considered to indicate cognitive impairment.

**Statistical Analysis**

The size of Sample 1 was estimated based on an interrater reliability of 0.9 expressed by intraclass correlation coefficient (ICC) between the 2 evaluators for the original English version of the CARE questionnaire. The calculation of the sample size for an interrater reliability derived from the desired width of the confidence interval around the ICC statistic, rather than from rejection of a null hypothesis that the ICC was different from zero. Another step that influenced the sample size estimation is the numbers of raters. In the present study, to achieve a confidence interval of 0.2 for the interrater reliability, at least 15 participants were required. In Sample 2, to detect a true difference in correlation coefficients between the null and alternative hypothesis levels, the minimum number of pairs of subjects was computed by using an expected power and 2-sided type I error probability. We assumed the correlation between the CARE-C and any of the 4 validated scales to be
0.5, and 115 participants were needed for a predesignated alpha value of 0.05 and a power of 0.8.14 Continuous and categorical variables were expressed by mean with standard deviation (SD) and proportion, respectively. The internal consistency of the subscale was assessed by Cronbach alpha coefficients; values exceeding 0.70 were considered acceptable.15 We also reported the index at the item level, “Cronbach alpha if item deleted” to demonstrate the increment or drop in the alpha’s value if eliminating a scale component. The ICC was used to estimate the test–retest and interrater reliabilities. An ICC value between 0.4 and 0.8 indicates fair to good agreement, whereas a value greater than 0.8 is regarded as excellent.16 Spearman correlations between the subscale scores of the CARE-C and the total scores of the Barthel index, IADL scale, EQ-5D, and MMSE were used to determine criterion-related validity. A correlation coefficient between 0.4 and 0.8 is satisfactory.17 All the statistical analyses were performed using SPSS software, version 16.0 (Chicago, IL) and a P-value <0.05 indicated statistical significance.

RESULTS

The Development of the CARE-C

The original English version of the CARE questionnaire includes 282 items, distributed in 11 chapters. We extracted 110 items relevant to the social, cultural, and medical background of Taiwan society after receiving input from 2 expert committees. We further discarded 58 items which are questions with multiple choices or with skip logic and conditional branching. After discussion with the original author, the final version of the CARE-C contained 52 items which were categorized into 11 subscales: cognition, delirium, disordered behavior, depression, pain, continence, perception, impairment, basic activity of daily living, transfer, and instrumental activity of daily living. The principle of scoring is that higher scores represent higher quality. For example, participants with a higher depression subscale score are more depressed, whereas participants achieving a greater score in the cognition subscale have better cognitive function (Fig. 1).

Characteristics of the Sample

For Sample 1, we recruited 30 participants from the outpatient clinic, 40% of whom were female (n = 12). The average age was 68.8 years (SD = 16.4) and the mean duration from their first-ever stroke to being recruited into the study was 5.4 years (SD = 4.2). Ischemic stroke accounted for 53% of Sample 1. For Sample 2, 138 patients were enrolled, including 41% females, the mean age was 64.4 years (SD = 13.4) years, and 28% of the strokes were hemorrhagic. The average time between stroke onset and the first assessment was 29.6 days (SD = 23.6).

Interrater and Test–Retest Reliability

Among the 52 items, 1 in the delirium subscale and 3 in the disordered behavior subscale could not generate the value of ICC for interrater and test–retest reliability because their
between-subject and within-subject variances were zero (see Supplemental Table 1, http://links.lww.com/MD/A472, which detailed the Cronbach α, interrater reliability, test–retest reliability, and corrected item-total correlation of CARE-C for each item). The interrater reliability ranged from 0.72 to 0.99 at the subscale level and from 0.13 to 1 at the item level. The ICC of the 10 subscales (excluding the disordered behavior subscale) were all larger than 0.7, whereas only 1 of the 49 items (excluding 4 items without ICC values) had an ICC of less than 0.4 (Table 1).

Regarding test–retest reliability, the value varied from 0.60 to 1 at the subscale level and from 0.46 to 1 at the item level. Most subscales (excluding the disordered behavior subscale) had ICCs of more than 0.7 and only one subscale, the depression subscale, demonstrated an ICC of 0.6. The ICCs of all items (excluding 4 items without ICC values) were higher than 0.4.

### Internal Consistency and Item-to-Subscale Correlation

The Cronbach alpha coefficients of each subscale of the CARE-C are listed in Table 2. Six of the 11 subscales met acceptable levels of reliability (Cronbach alpha >0.7). One (the disordered behavior subscale) could not generate the value of Cronbach alpha coefficient and four (the depression, pain, continence, and impairment subscales) had Cronbach alpha coefficients of less than 0.7. Spearman rank correlation coefficients between the items and subscales are shown in Supplemental Table 1, http://links.lww.com/MD/A472. The absolute value of correlation coefficients ranged from 0.2 to 1 (excluding 5 items where correlation coefficients could not be generated due to no between-item variance), and 85.1% of all item-to-subscale correlations (40/47) met the minimal acceptable item convergent validity criteria (Spearman rank correlation >0.40).

### Criterion-Related Validity

In Table 2, the correlations between each subscale of the CARE-C and the instruments for criterion-related validity (Barthel index, IADL scale, EQ-5D, and MMSE) are shown. The Barthel index was positively associated with the cognition subscale (r = 0.52) but inversely associated with the delirium subscale (r = -0.24). The IADL scale was positively correlated with the subscales of cognition (r = 0.51), continence (r = 0.31), perception (r = 0.44), basic activity of daily living (BADL) (r = 0.66), transfer (r = 0.64), and IADL (r = 0.81) but inversely correlated with the disordered behavior (r = -0.25) and impairment (r = -0.18) subscales. The EQ-5D was positively correlated with the subscales of cognition (r = 0.26), continence (r = 0.38), perception (r = 0.33), BADL (r = 0.63), transfer (r = 0.58), and IADL (r = 0.57) but inversely correlated with the depression (r = -0.27), pain (r = -0.30), and impairment (r = -0.41) subscales. The MMSE was positively correlated with the cognition (r = 0.911), depression (r = 0.202), pain (r = 0.187), continence (r = 0.376), perception (r = 0.787), BADL (r = 0.599), transfer (r = 0.411), and IADL (r = 0.577) subscales but correlated negatively with the delirium subscale (-0.411).

### DISCUSSION

The operation of a sustainable PAC program required an integrative tool capable of multidimensionally evaluating patients in different settings, including the acute, postacute, and long-term care settings. The development of the CARE-C facilitated the assessment of integrity in a PAC service in the Chinese population, especially for stroke participants. The findings reported here suggest that the CARE-C is a reliable measurement tool for stroke patients, with the interrater and test–retest reliability (represented by ICC) exceeding 0.7 in 82% of the subscales (9/11). Most subscales in the CARE-C had...
TABLE 2. Criteria-Related Validity of the Chinese version of the Continuity Assessment Record With the Barthel Index, Lawton Instrumental Activities of Daily Living Scale, EQ-5D Five Dimensions Questionnaire, and Mini-Mental State Examination

| Subscale                  | Mean | Standard Deviation | Range   | % of Ceiling (Number) | % of Flooring (Number) | Barthel Index (Correlation Coefficient) | IADL Scale (Correlation Coefficient) | EQ-5D (Correlation Coefficient) | MMSE (Correlation Coefficient) |
|---------------------------|------|--------------------|---------|-----------------------|------------------------|----------------------------------------|-------------------------------------|-------------------------------|-------------------------------|
| Cognition                 | 9.6  | 1.7                | 1–16    | 18.1 (25)             | 23.1 (32)              | 0.522†                                 | 0.505†                             | 0.258†                         | 0.911†                         |
| Delirium                  | 0.7  | 1.4                | 0–8     | 0.7 (1)               | 74.6 (103)             | -0.241†                                | -0.252†                            | -0.072                         | -0.411†                        |
| Disordered behavior       | 0.1  | 420                | 0–3     | 1.4 (2)               | 95.6 (132)             | -0.114                                 | -0.019                             | -0.043                         | -0.134                         |
| Depression                | 0.5  | 0.7                | 0–2     | 10.1 (14)             | 63.7 (88)              | -0.088                                 | -0.042                             | -0.274†                        | 0.202†                         |
| Pain                      | 0.3  | 0.7                | 0–2     | 10.8 (15)             | 82.6 (114)             | -0.048                                 | 0.003                              | -0.302†                        | 0.187†                         |
| Continence                | 11.1 | 2.9                | 3–15    | 5.8 (8)               | 0 (0)                  | 0.537†                                 | 0.311†                             | 0.377†                         | 0.376†                         |
| Perception                | 11.7 | 2.9                | 4–14    | 48.5 (67)             | 4.3 (6)                | 0.466†                                 | 0.440†                             | 0.325†                         | 0.787†                         |
| Impairment                | 1.9  | 0.5                | 0–3     | 0 (0)                 | 0.7 (1)                | -0.222†                                | -0.177†                            | -0.409†                        | -0.100                         |
| Basic activity of daily   | 14.9 | 7.5                | 4–30    | 11.5 (16)             | 0 (0)                  | 0.811†                                 | 0.664†                             | 0.625†                         | 0.599†                         |
| living                    |      |                    |         |                       |                        |                                       |                                     |                               |                               |
| Transfer                  | 10.0 | 6.0                | 0–24    | 5.8 (8)               | 1.4 (2)                | 0.752†                                 | 0.637†                             | 0.578†                         | 0.411†                         |
| Instrumental activity of  | 26.1 | 18.12              | 4–84    | 3.6 (5)               | 0 (0)                  | 0.782†                                 | 0.809†                             | 0.573†                         | 0.577†                         |
| daily living              |      |                    |         |                       |                        |                                       |                                     |                               |                               |

EQ-5D = EuroQOL five dimensions questionnaire, IADL scale = Lawton Instrumental Activities of Daily Living scale, MMSE = Mini-Mental State Examination.

† P < 0.05.

Cronbach alpha coefficients higher than 0.7, except the subscales of disordered behavior, depression, pain, continence, and impairment. The correlation analysis with the Barthel index, IADL scale, EQ-5D, and MMSE revealed that the CARE-C is a valid measurement for evaluating the aspect of basic ADL, instrumental ADL, and cognitive function.

Interrater reliability between 2 evaluators was good (ICC > 0.7) in most of the subscales except the disordered behavior subscale. This subscale evaluated the ability of physical or verbal assaults, which no participant in Sample 1 carried out. The finding was reasonable because the group of patients was chronically stable and could comply with the rehabilitation program in the outpatient setting. Therefore, the value of ICC could not be generated because the overall variance in this sample was zero. At the item level, only one item evaluating the ability for memorizing the last object among three given nouns had an ICC of less than 0.4 (Supplemental Table 1, http://links.lww.com/MD/A472). We speculate that the result might derive from inconsistency of the inquiring skills regarding registration recall, which suggests that training should be emphasized for future raters. Likewise, the test–retest reliability reached 0.7 of ICC in most of the subscales, except the depression subscale. Since the depression subscale only included 2 items, each composed of 2 options, any small change in both items at an 1-week interval might lead its ICC to have a lower value compared with other subscales. Compared with the original version of CARE item set, interator reliability was very good for items in impairment, basic ADL, as all overall statistics indicates substantial or almost perfect agreements among raters in inpatient rehabilitation facilities in the United States. The items in cognition, delirium, and behavior also had similar lower agreements due to a small sample size in these items.18

Most of the Cronbach alpha coefficients at the subscale level appeared excellent, except for the depression, pain, continence, and impairment subscales. Cronbach alpha coefficient was positively correlated with the number of items in a subscale, which explains why the depression and pain subscales, both of which included only 2 items, showed low values in internal consistency. Regarding the continence subscale, the inclusion of swallowing assessment along with bowel and bladder assessment was the most likely cause for poor agreement between items. The above reason could also be applied to the impairment subscale, which evaluates four relatively heterogeneous qualities—weight bearing restriction, a decrease in the grasping ability, impaired respiratory status, and endurance ability—in a subscale.

The criterion-related validity of the CARE-C was considered satisfactory, showing moderate to high correlations of its subscale with the total score in the 4 validated scales. The high correlations between the BADL subscale and Barthel index (r = 0.811), between the IADL subscale and IADL scale (r = 0.809), and between the cognition subscale and MMSE, indicated that the CARE-C is capable of measuring the three most commonly involved dimensions in stroke patients. We did not identify a subscale of the Care-C with a high association (r > 0.7) with the EQ-5D. The result was understandable because the EQ-5D is a generic measure of health status that is not designed to evaluate a specific quality. However, the moderate correlation between the EQ-5D and the subscales of the ADL-related subscales (ADL and IADL) might imply that the ability to perform basic and instrumental ADL played a crucial part in well-being of health in the stroke population.

The present study has several limitations worth noting. First, we only chose stroke patients for our validation of the CARE-C, which limited generalizability to other populations. Therefore, further research is warranted to investigate the performance of the CARE-C in nonstroke participants. Second, because aphasia is present in a substantial portion of stroke...
victims, its impact on their comprehension and expression might affect the validity regarding psychometrical-related subscales, such as the cognition and depression subscales. Third, we were unable to obtain an ICC value regarding the reliability of the disordered behavior subscale due to absence of variation in the responses to the questionnaire. The finding is understandable because the participants in Sample 1 were recruited in out-patient settings and were assumed to be chronically stable without aggressiveness. However, because the CARE-C is designed for a continuous measurement of patients admitted without aggressiveness. However, because the CARE-C is designed for a continuous measurement of patients admitted

**CONCLUSIONS**

The CARE-C is useful for the evaluation of functional quality metrics in the Chinese stroke population. The scale can be used for acute and PAC providers and includes subscales of memory, problem solving, communication, functional status, mobility, and self-care ability. The development of the CARE-C also facilitates an assessment of a PAC providing program and further contributes to improvement in its payment system. Future research is warranted for validating the capability of the CARE-C for identifying patients’ functional change over time and its generalizability to nonstroke populations.

**REFERENCES**

1. Redberg RF. The role of post-acute care in variation in the Medicare program. *JAMA Intern Med.* 2015;175:1058.
2. Burke RE, Juarez-Colunga E, Levy C, et al. Rise of post-acute care facilities as a discharge destination of US hospitalizations. *JAMA Intern Med.* 2015;175:295–296.
3. Centers for Medicare and Medicaid Services (CMS). Medicare and Medicaid programs: reporting outcome and assessment information set data as part of the conditions of participation for home health agencies. *Final Rule Fed Reg.* 2005;70:76199–76208.
4. Mallinson T, Leland NE, Chan TH. The need for uniform quality reporting across post-acute care rehabilitation settings: an examination of accidental falls. *J Am Geriatr Soc.* 2015;63:195–197.
5. Deutsch A, Kline T, Kelleher C, et al. Continuity Assessment Record and Evaluation (CARE) Item Set: Video Reliability Testing NW Washington, DC: The Brooking Institution; 2012.
6. Hsieh F-I, Lien L-M, Chen S-T, et al. Get With The Guidelines-Stroke performance indicators: surveillance of stroke care in the Taiwan Stroke Registry: Get With The Guidelines-Stroke in Taiwan. *Circulation.* 2010;122:1116–1123.
7. Li L, Wang HM, Shen Y. Chinese SF-36 Health Survey: translation, cultural adaptation, validation, and normalisation. *J Epidemiol Community Health.* 2003;57:259–263.
8. Thorvaldsen P, Asplund K, Kuulasmaa K, et al. Stroke incidence, case fatality, and mortality in the WHO MONICA project. World Health Organization Monitoring Trends and Determinants in Cardiovascular Disease. *Stroke.* 1995;26:361–367.
9. Hsieh I-P, Lin J-H, Jing J-S, et al. Comparison of the psychometric characteristics of the functional independence measure, 5 item Barthel index, and 10 item Barthel index in patients with stroke. *J Neurol Neurosurg Psychiatry.* 2002;73:188–190.
10. Vergara I, Bilbao A, Orive M, et al. Validation of the Spanish version of the Lawton IADL Scale for its application in elderly people. *Health Qual Life Outcomes.* 2012;10:130.
11. Graf C. Hartford Institute for Geriatric Nursing. The Lawton Instrumental Activities of Daily Living (IADL) scale. *Med Surg Nurs.* 2008;17:343–344.
12. Lee HY, Hung MC, Hu FC, et al. Estimating quality weights for EQ-5D (EuroQol-5 dimensions) health states with the time trade-off method in Taiwan. *J Formos Med Assoc.* 2013;112:699–706.
13. Arevalo-Rodriguez I, Smailagic N, Roque IFM, et al. Mini-Mental State Examination (MMSE) for the detection of Alzheimer’s disease and other dementias in people with mild cognitive impairment (MCI). *Cochrane Database Syst Rev.* 2015;3:CD010783.
14. Shieh G. Sample size requirements for the design of reliability studies: precision consideration. *Behav Res Methods.* 2014;46:808–822.
15. Tsang KL, Chi I, Ho SL, et al. Translation and validation of the standard Chinese version of PDQ-39: a quality-of-life measure for patients with Parkinson’s disease. *Mov Disord.* 2002;17:1036–1040.
16. Peeters MJ, Beltyukova SA, Martin BA. Educational testing and validity of conclusions in the scholarship of teaching and learning. *Am J Pharm Educ.* 2013;77:186.
17. Fu X, Xu J, Song L, et al. Validation of the Chinese version of the quality of nursing work life scale. *PLos ONE.* 2015;10:e0121150.
18. Smith L, Gage B, Deutsch A, et al. Continuity Assessment Record and Evaluation (CARE) Item Set: additional provider-type specific interrater reliability analyses. Centers for Medicare & Medicaid Services Contract No. HHSM-500-2005-00291. 2012.