BMJ Open

# Chinesisation, adaptation and validation of the Chelsea Critical Care Physical Assessment Tool in critically ill patients: a cross-sectional observational study

Zhigang Zhang,[1,2] Guoqiang Wang,[3] Yuchen Wu ,[3] Jin Guo,[4] Nannan Ding,[5] Biantong Jiang,[6] Huaping Wei,[3] Bin Li,[3] Weigang Yue,[3] Jinhui Tian[7]

**To cite:** Zhang Z, Wang G, Wu Y, *et al*. Chinesisation, adaptation and validation of the Chelsea Critical Care Physical Assessment Tool in critically ill patients: a cross-sectional observational study. *BMJ Open* 2021;**11**:e045550. doi:10.1136/bmjopen-2020-045550

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2020-045550).

ZZ, GW, JG and HW are joint first authors.

Received 07 October 2020
Revised 19 February 2021
Accepted 17 March 2021



For numbered affiliations see end of article.

**Correspondence to**
Dr Yuchen Wu;
yuchen0723@126.com

## ABSTRACT

**Purpose** To translate and adapt the Chelsea Critical Care Physical Assessment Tool (CPAx) into Chinese version ('CPAx-Chi'), test the reliability and validity of CPAx-Chi, and verify the cut-off point for the diagnosis of intensive care unit-acquired weakness (ICU-AW).

**Study design** Cross-sectional observational study.

**Methods** Forward and back translation, cross-cultural adaptation and pretesting of CPAx into CPAx-Chi were based on the Brislin model. Participants were recruited from the general ICU of five third-grade class-A hospitals in western China. Two hundred critically ill adult patients (median age: 53 years; 64% men) with duration of ICU stay ≥48 hours and Glasgow Coma Scale ≥11 were included in this study. Two researchers simultaneously and independently assessed eligible patients using the Medical Research Council Muscle Score (MRC-Score) and CPAx-Chi.

**Results** The content validity index of items was 0.889. The content validity index of scale was 0.955. Taking the MRC-Score scale as standard, the criterion validity of CPAx-Chi was r=0.758 (p<0.001) for researcher A, and r=0.65 (p<0.001) for researcher B. Cronbach's α was 0.939. The inter-rater reliability was 0.902 (p<0.001). The area under the receiver operating characteristic curves of CPAx-Chi for diagnosing ICU-AW based on MRC-Score ≤48 were 0.899 (95% CI 0.862 to 1.025) and 0.874 (95% CI 0.824 to 0.925) for researcher B. The best cut-off point for CPAx-Chi for the diagnosis of ICU-AW was 31.5. The sensitivity was 87% and specificity was 77% for researcher A, whereas it was 0.621, 31.5, 75% and 87% for researcher B, respectively. The consistency was high when taking CPAx-Chi ≤31 and MRC-Score ≤48 as the cut-off points for the diagnosis of ICU-AW. Cohen's kappa=0.845 (p=0.02) in researcher A and 0.839 (p=0.04) for researcher B.

**Conclusions** CPAx-Chi demonstrated content validity, criterion-related validity and reliability. CPAx-Chi showed the best accuracy in assessment of patients at risk of ICU-AW with good sensitivity and specificity at a recommended cut-off of 31.

## Strengths and limitations of this study

- Two researchers assessed and collected data independently, which improved the reference value of the validation data.
- We took Medical Research Council Muscle Score ≤48 as the criterion to demonstrate the best cut-off point for the diagnosis of intensive care unit-acquired weakness using the Chinese version of Chelsea Critical Care Physical Assessment Tool (CPAx-Chi).
- This is a non-randomised pool of participants chosen primarily by their availability during the study period.
- There were specific exclusion criteria that may have stopped the potential 'ceiling and floor' effects of CPAx-Chi to be tested.

## INTRODUCTION

Intensive care unit-acquired weakness (ICU-AW) is a severe and debilitating complication in critically ill patients. The prevalence of ICU-AW in patients receiving mechanical ventilation for more than 4–7 days has been reported to be 38%–86%.[1–5] The prevalence of ICU-AW in patients with sepsis is 86%.[1 2 5] Early identification, assessment and active prevention are crucial to reduce ICU-AW risk because the pathophysiological mechanism of ICU-AW is not clear, and efficacious pharmacotherapy is lacking.[1 6]

A gold standard for the diagnosis of ICU-AW is not available, and the Medical Research Council Muscle Score (MRC-Score) is the most widely used diagnostic tool for ICU-AW.[7] Other tests are also frequently used to test for ICU-AW but there is no uniform cut-off point. The MRC-Score evaluates the strength subjectively in three muscle groups of all four limbs according to the Oxford Muscle Strength Grading Scale. The latter is not only affected by several factors, it also cannot evaluate respiratory function. Several studies have shown that diaphragmatic dysfunction is correlated

significantly with ICU-AW,[8–10] and that the function of respiratory muscles may be related to the occurrence and development of ICU-AW.

The Chelsea Critical Care Physical Assessment Tool (CPAx) could be an optimal tool for predicting and evaluating ICU-AW. CPAx can be used to measure physical function, mobility, grip strength, respiratory function and cough ability.[11–13] CPAx has been translated into several languages for use in the UK, Sweden, Denmark and other countries.[14 15] However, a Chinese version of CPAx (CPAx-Chi), or the cut-off point of CPAx for the diagnosis of ICU-AW is lacking. Therefore, the aim of this study is to translate and adapt the CPAx into 'CPAx-Chi', test the reliability and validity of CPAx-Chi, and verify the cut-off point for the diagnosis of ICU-AW.

## MATERIALS AND METHODS

### Translation, cross-cultural adaptation and pretesting

The translation of the original CPAx tool into Chinese was completed with the consent and assistance of the primary original author (EJ Corner).[11–13] Translation, cross-cultural adaptation and pretesting were done based on the model described by Brislin.[16–18]

### Translation

Three bilingual authors with Chinese as their native language undertook the forward translation of CPAx from English to Chinese. One was a physician experienced within the specialty of critical illness; one was a nurse experienced within the specialty of critical illness; one was a graduate student in nursing with College English Test 6 certification unfamiliar with clinical medicine. A seminar was conducted to discuss and synthesise the results of the three translators. Different opinions were resolved through group consultation, and then integrated into CPAx-Chi, which was named 'CPAx-Chi-Forward'.

### Back translation

Three bilingual translators with English as their native language translated CPAx-Chi-Forward back into English. One was a doctoral student in nursing based in the UK; one was a doctoral student in physiotherapy based in Canada; one was a certified English linguist. They were unfamiliar with and blinded to the original CPAx version. A seminar was conducted to discuss and compare CPAx-Chi-Forward with the original CPAx.[11 12] Discrepancies between the three translations were discussed until consensus was reached, and then the final synthesised back-translated English version was named 'CPAx-Eng-Back'. The researchers provided a final report that included the annotations from translators about their *rationale* for translation, choices and linguistic considerations to the author of the original CPAx.

### Cross-cultural adaptation

Nine experts revised the items of CPAx-Chi-Forward based on their theoretical knowledge, practical experience, subjective feelings and expression in the Chinese language. Two were specialists in critical care medicine, five were nursing specialists in critical care, one was a respiratory therapist and one was a physiotherapist. During the process, some words were rephrased or adjusted due to linguistic, grammatical, terminological or cultural differences between English and Chinese. Changes from the original CPAx version to the synthesised back-translated English version were discussed and accepted by the original author.[11]

### Pretesting and verifying cultural adaptation

Forty ICU nurses from the First Hospital of Lanzhou University applied CPAx-Chi-Forward and manual dynamometer (WCS-100) to assess ICU patients. Meanwhile, a 5-step Likert scale method was used to assess if the written expression in CPAx-Chi-Forward was readily comprehensible, well described and conform to Chinese grammar, and suggestions could be noted. The result showed that there were no significant differences regarding the assessments of 'readily comprehensive', 'well described', 'conform to Chinese grammar' in nurses with varied sex, nationality, professional title or time working in the ICU ($p>0.05$) (table 1) from culture adaptation. Adjustments were not deemed necessary and CPAx-Chi-Forward had good cross-cultural adaptation. Therefore, the final CPAx-Chi was accepted.

### Verification of CPAx-Chi

#### Study design

This was a cross-sectional observational study, and the flow chart is shown in figure 1. We took the MRC-Score which is the most widely used diagnostic tool for ICU-AW as the comparator to test CPAx-Chi, and a manual dynamometer was used to assess grip strength (WCS-100). Meanwhile, researcher A and researcher B simultaneously and independently assessed eligible patients using the MRC-Score and CPAx-Chi.

#### Participants

Adult critically ill patients were recruited from the general ICU of five third-grade class-A hospitals in western China from September 2019 to June 2020. The recruiter explained the purpose and significance of the study to participants who meet the eligibility criteria, and then serial numbered.

The inclusion criteria were the following: (1) critically ill patients eligible for ICU admission; (2) age ≥18 years; (3) duration of ICU stay ≥48 hours; (4) Glasgow Coma Scale (GCS) score ≥11; (5) volunteered to participate in our study.

Patients were excluded if they may be misdiagnosed as ICU-AW just like: (1) unstable fracture, limb deformity or limb dysfunction; (2) myasthenia gravis or Guillain-Barre syndrome.

#### Sample size

The sample size was calculated by the principles of scale development. In general, the sample size was 10–15

**Table 1** Verifying cultural adaptation differences (n=40)

| Items | Readily comprehensible | | | Well described | | | Conform to Chinese grammar | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\bar{\chi} \pm s$ | F | P value | $\bar{\chi} \pm s$ | F | P value | $\bar{\chi} \pm s$ | F | P value |
| Sex | | | | | | | | | |
| Male | 47.25±1.36 | 054 | 0.47 | 47.13±1.48 | 3.11 | 0.08 | 47.42±1.56 | 0.63 | 0.43 |
| Female | 46.94±1.24 | | | 47.94±1.34 | | | 46.94±2.27 | | |
| Nationality | | | | | | | | | |
| Han nationality | 47.18±1.34 | 0.35 | 0.56 | 47.38±1.51 | 0.48 | 0.49 | 47.24±1.96 | 0.07 | 0.94 |
| Minority | 46.83±1.17 | | | 47.83±1.17 | | | 47.17±0.98 | | |
| Professional title | | | | | | | | | |
| Nurse | 47.40±1.35 | 0.76 | 0.48 | 47.60±2.01 | 0.31 | 0.73 | 48.20±0.63 | 2.47 | 0.09 |
| Senior nurse | 47.22±1.39 | | | 47.56±1.20 | | | 47.17±1.82 | | |
| Supervisor nurse and above | 46.75±1.31 | | | 47.17±1.40 | | | 46.50±2.32 | | |
| Time working in the ICU | | | | | | | | | |
| ≤3 years | 46.83±1.47 | 0.82 | 0.49 | 47.50±1.76 | 0.88 | 0.46 | 47.67±1.51 | 0.67 | 0.58 |
| 3–5 years | 47.25±1.29 | | | 47.83±1.34 | | | 47.67±1.96 | | |
| 5–10 years | 47.14±1.35 | | | 47.50±1.35 | | | 47.00±1.80 | | |
| ≥10 years | 47.13±1.36 | | | 46.75±1.67 | | | 46.63±2.13 | | |

ICU, intensive care unit.

times the number of scale items and add taking into 20% account loss to follow-up and participant attrition.[19] The sample size of this study was 120–180. However, there are some studies that introduced that a sample size of 200 is reasonably good for ordinary factor-analytical work with 40 or fewer variables.[19–21] Therefore, we finally took 200 cases as the sample size of this study.

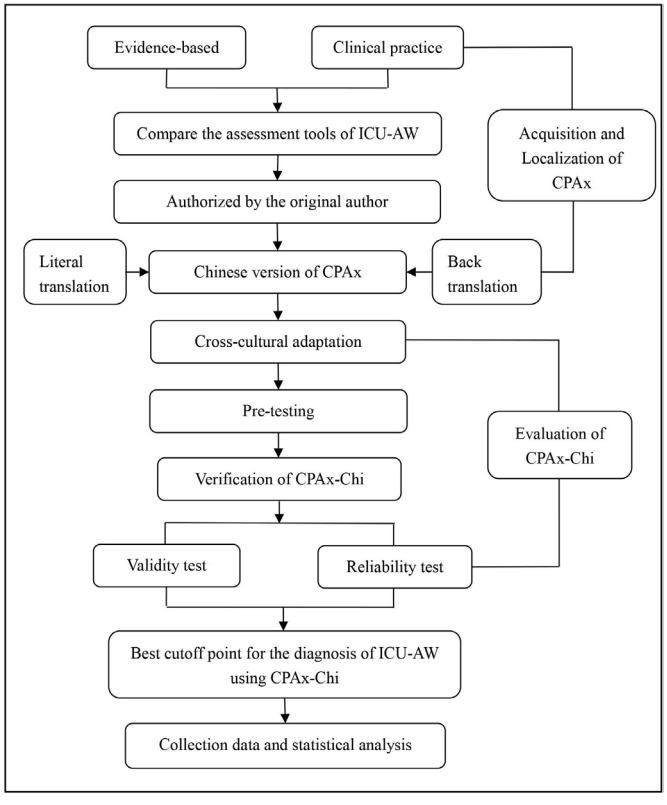


**Figure 1** The flow diagram of the research. CPAx, Chelsea Critical Care Physical Assessment Tool; CPAx-Chi, Chinese version of Chelsea Critical Care Physical Assessment Tool; ICU-AW, intensive care unit-acquired weakness.

### Ethical approval and consent to participate

In the pretesting study, we found that there were significant differences between patients with informed consent and the patients with no informed consent, especially in the items respiratory function, grip strength and transferring from bed to a chair. In order to ensure data validity and quality, the Ethics Committee provided a waiver of informed consent which was uploaded as supplemental information (online supplemental material). In addition, it was routine work that researchers assessed eligible patients using the MRC-Score and CPAx-Chi.

## Patient and public involvement

The study was designed to test the CPAx-Chi, and verified the cut-off point of CPAx-Chi to diagnose ICU-AW. However, patients were not involved in the design of the survey instrument, recruitment or conduct of the study. Patients who participated did so anonymously, and therefore the study team will be unable to disseminate the results to study participants.

## Statistical analyses

SPSS V.22.0 (IBM) was employed for statistical analyses. Frequency and percentages were used for dichotomous variables. The mean±SD was used for continuous variables. Content validity (CV) and criterion-related validity were employed to test the validity of CPAx-Chi. CV index (CVI) included the scale-level CVI (S-CVI) and item-level

CVI (I-CVI) which is the most widely used index in scale evaluation. The expert authority coefficient and Kendall synergy coefficient were used to calculate expert evaluation results, and the more Kendall synergy coefficient, the more consistent the results are. Cronbach's α coefficient and inter-rater reliability were used to test the reliability of CPAx-Chi. The MRC-Score was taken as the standard to calculate the receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) of CPAx-Chi. The cut-off point of CPAx-Chi was determined by the maximum value of the Youden Index (YI). The kappa test was used to test the consistency of the MRC-Score and CPAx-Chi. P<0.05 was considered significant.

## RESULTS

### Characteristics of participants

Nine experts adjusted the cultural adaptability of CPAx-Chi, and evaluated the importance and relevance of each item in the scale. Two (22.22%) were specialists in critical care medicine, five (55.56%) were nursing specialists in critical care, one (11.11%) was a respiratory therapist and one (11.11%) was a physiotherapist. The median age of specialists was 38 (IQR 33–50) years. The median time the specialists had been working in the ICU was 13 (IQR 6–23) years. There were nine specialists that included one (11.11%) undergraduate, four (44.44%) masters and four (44.44%) doctors; four (44.44%) intermediate titles and five (55.55%) senior titles.

Two-hundred critically ill patients participated in this study (128 (64%) men and 72 (36%) women; mean age: 53.24±15.06 years). The Acute Physiology and Chronic Health Evaluation score was 15.04±6.70. The mean duration of ICU stay was 9.04±6.15 days. The mean duration of hospital stay was 20.79±11.84 days. The duration of mechanical ventilation was 3.55±5.19 days. The principal diagnoses of participants were: craniocerebral injury (16, 8%), respiratory failure (22, 11%), surgical complications (68, 34%), hepatobiliary disease (42, 21%), cardiovascular disease (20, 10%), shock (14, 7%) and other (18, 9%). Also, 190 (95%) patients were transferred to other departments, 2 (1%) patients were discharged from ICU, 2 (1%) patients were transferred to other hospitals and 6 (3%) patients died.

### Validity

#### Content validity

The I-CVI was from 0.889 to 1. The S-CVI, which is the average of I-CVI, was 0.955. The median expert authority coefficient was 0.85 (IQR 0.75–0.95). The Kendall synergy coefficient was 0.61 (p=0.842), and a significant difference was not detected in the degree of expert coordination.

#### Criterion validity

The correlation coefficient for ICU-AW assessment by researcher A between the MRC-Score and CPAx-Chi was 0.60 (p<0.001). The correlation coefficient for ICU-AW assessment by researcher B between the MRC-Score and CPAx-Chi was 0.65 (p<0.001) (table 2).

**Table 2** Criterion validity (n=200)

| Researcher | Criterion | Mean±SD | r | P value |
|---|---|---|---|---|
| A | CPAx-Chi | 32.46±8.83 | 0.60 | 0.000 |
| | MRC-Score | 50.15±10.42 | | |
| B | CPAx-Chi | 33.43±9.08 | 0.65 | 0.000 |
| | MRC-Score | 50.81±10.50 | | |

CPAx-Chi, Chinese version of Chelsea Critical Care Physical Assessment Tool; MRC-Score, Medical Research Council Muscle Score.

### Reliability

The internal consistency of CPAx-Chi was acceptable (Cronbach's α=0.939). The correlation coefficient between researcher A and researcher B in the items of CPAx-Chi was between 0.668 and 0.992 (p<0.001). The correlation coefficient between researcher A and researcher B in CPAx-Chi total score was 0.902 (p<0.001) (table 3).

### Best cut-off point for the diagnosis of ICU-AW using CPAx-Chi

The ROC curve for ICU-AW diagnosis with CPAx-Chi was drawn taking MRC-Score ≤48 as the standard for the diagnosis of ICU-AW. An MRC-Score ranging from 0 to 48 was termed '1' (ICU-AW group). An MRC-Score >48 was termed '0' (non-ICU-AW group).

The AUC for researcher A was 0.899 (95% CI 0.862 to 1.025) (figure 2). The AUC for researcher B was 0.874 (95% CI 0.824 to 0.925) (figure 3). The best cut-off point was determined by the maximum value of the YI. The maximum YI for researcher A was 0.643, the cut-off point was 31.5, the sensitivity was 87% and specificity was 77%. The maximum YI for researcher B was 0.621, the cut-off point was 31.5, the sensitivity was 75% and specificity was 87%.

### MRC-Score and CPAx-Chi were consistent for the diagnosis of ICU-AW

We calculated 31 as the best cut-off point to diagnose ICU-AW using CPAx-Chi. Hence, if the total score of CPAx-Chi ranged from 0 to 31, it was marked as 1 (ICU-AW group), and if the total score of CPAx-Chi ranged from 32 to 50, it was marked as 0 (non-ICU-AW group). We found no significant difference in the total score of the ICU-AW group and non-ICU-AW group for researcher A (F=4.53, p=0.035) or researcher B (F=6.51, p=0.011). The test for consistency suggested that accepting CPAx-Chi ≤31.5 and MRC-Score ≤48 as the best cut-off points for the diagnosis of ICU-AW, then kappa was 0.845 (p=0.02) for researcher A, and kappa=0.839 (p=0.04) for researcher B (table 4).

## DISCUSSION

### Translation

The present study is the first to translate CPAx from English to Chinese using the Brislin model to guarantee sufficient equivalency.[16–18] Our study was strengthened by

Table 3 Inter-rater reliability (n=200)

| Physical parameter | Researcher | Mean±SD | r | P value |
|---|---|---|---|---|
| Respiratory function | A | 3.62±0.79 | 0.965 | <0.001 |
| | B | 3.64±0.76 | | |
| Cough | A | 4.10±0.99 | 0.715 | <0.001 |
| | B | 4.21±0.77 | | |
| Moving within a bed (eg, rolling) | A | 4.06±1.02 | 0.798 | <0.001 |
| | B | 4.08±1.02 | | |
| Supine to sitting on the bed edge | A | 3.13±1.06 | 0.766 | <0.001 |
| | B | 3.24±1.26 | | |
| Dynamic sitting | A | 3.69±1.17 | 0.701 | <0.001 |
| | B | 3.66±1.08 | | |
| Standing balance | A | 2.87±1.15 | 0.766 | <0.001 |
| | B | 1.97±1.21 | | |
| Sit to stand (starting position: ≤90° hip flexion) | A | 2.76±1.11 | 0.763 | <0.001 |
| | B | 2.67±1.20 | | |
| Transferring from bed to chair | A | 2.61±1.08 | 0.853 | <0.001 |
| | B | 2.94±1.16 | | |
| Stepping | A | 1.95±1.21 | 0.775 | <0.001 |
| | B | 2.33±1.47 | | |
| Grip strength | A | 3.76±1.35 | 0.992 | <0.001 |
| | B | 3.76±1.36 | | |
| CPAx-Chi score | A | 32.46±8.83 | 0.902 | <0.001 |
| | B | 33.59±9.44 | | |

CPAx-Chi, Chinese version of Chelsea Critical Care Physical Assessment Tool.

including a multidisciplinary team to remedy content variance, and included two Chinese nurses with English certifications studying, respectively, in the UK and Canada. We undertook tests for criterion validity and reliability for the completed translation.

### Validity of CPAx-Chi

Validity is the degree that a measured result reflects the measured content. The more consistent the measured result is with the measured content, the higher is the validity.[21 22] According to the handbook of scale development, when the number of experts is more than five, the good standard of I-CVI is more than 0.78, and the experts must be authoritative and coordinated.[22 23]

The present study involved nine ICU multidisciplinary experts with deep theoretical knowledge and clinical experience. The expert authority coefficient ranged from 0.75 to 0.95. The Kendall synergy coefficient was 0.61 (p=0.842) and I-CVI ranged from 0.889 to 1. Therefore, CPAx-Chi had good content validity.[24 25]

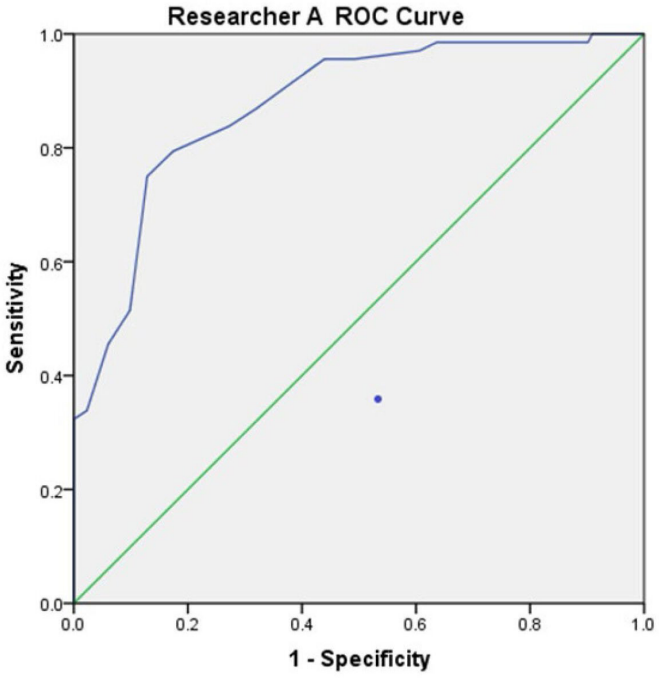


Figure 2 The ROC curve and area under the ROC curve of researcher A. ROC, receiver operating characteristic curve.

Corner *et al* demonstrated that the CVI of CPAx was 1 (p<0.05).[11 12] They also showed that CPAx has good predictive validity, and that the CPAx score could be used as an alternative indicator of functional prognosis in critically ill patients by analysing the relationship between the CPAx score and patient outcomes.[13] Other colleagues demonstrated the criterion validity of CPAx taking the scores for the MRC, Short Form (SF)-36, Sequential Organ Failure

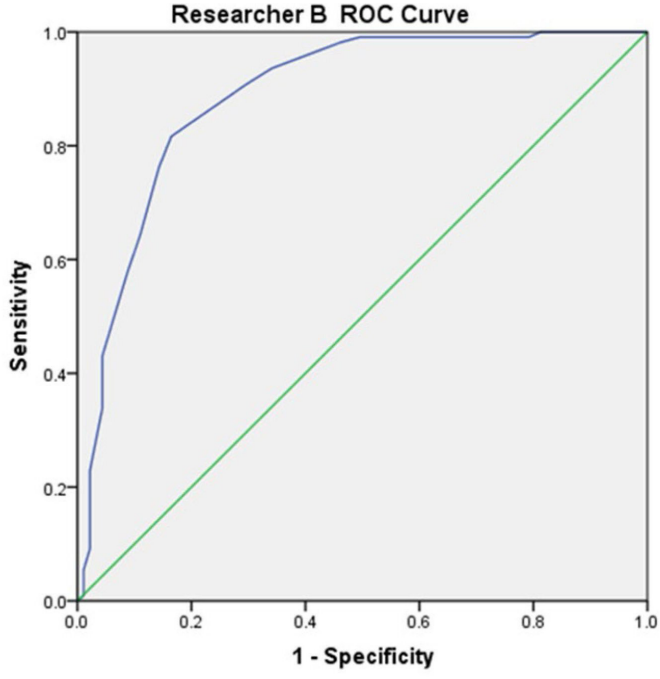


Figure 3 The ROC curve and area under the ROC curve of researcher B. ROC, receiver operating characteristic curve.

Table 4 MRC-Score and CPAx-Chi were consistent for the diagnosis of ICU-AW (n=200)

| Researcher | Scale | Group | MRC-score | | Kappa | P value |
|---|---|---|---|---|---|---|
| | | | ICU-AW (n) | Non-ICU-AW (n) | | |
| A | CPAx-Chi | ICU-AW | 74 | 12 | 0.845 | 0.038 |
| | | Non-ICU-AW | 3 | 111 | | |
| B | | ICU-AW | 66 | 6 | 0.839 | 0.04 |
| | | Non-ICU-AW | 9 | 119 | | |

CPAx-Chi, Chinese version of Chelsea Critical Care Physical Assessment Tool; ICU-AW, intensive care unit-acquired weakness; MRC-Score, Medical Research Council Muscle Score.

Assessment (SOFA) and GCS as a standard.[26] They found that the correlation coefficient between the CPAx score and MRC-Score was 0.65 (p<0.001). The correlation coefficients for the right upper limb, left upper limb, right lower limb and left lower limb with the CPAx score were, respectively, 0.69, 0.64, 0.69 and 0.67. The correlation coefficient between the CPAx score and SOFA score was 0.68 (p<0.001). The correlation coefficient between the CPAx score and GCS was 0.74 (p<0.001). The correlation coefficient between the physical-function item of SF-36 and the CPAx score was 0.72 (p=0.013). The correlation coefficient between the mental-function component of SF-36 and the CPAx score was 0.024 (p=0.95). In the present study, the correlation coefficient between the CPAx-Chi score and the items of the MRC-Score ranged from 0.60 to 0.65 (p<0.001). Therefore, CPAx-Chi had good validity.

### Reliability of CPAx-Chi

Cronbach's α mainly reflects the internal consistency of a scale.[21 22] In general, Cronbach's α should be >0.7; a value <0.6 indicates that the items of scale must be revised. From the perspective of psychometrics, the 'ideal' Cronbach's α should be >0.8.[27–29] The inter-rater reliability mainly demonstrates the consistency of evaluation results among different evaluators, and the stability of scales used among different evaluators.[30 31] An inter-rater correlation coefficient >0.7 indicates that the inter-rater reliability is good. The inter-rater correlation coefficient ranging from 0.8 to 0.9 indicates that the inter-rater reliability is high.[14 21 22 31] In the present study, Cronbach's α for CPAx-Chi was 0.939, and the inter-rater reliability of the CPAx-Chi score was 0.902 (p<0.001). The inter-rater correlation coefficient was >0.8 for the items of respiratory function, transfer from bed to chair and grip strength. The inter-rater correlation coefficient of other items of CPAx-Chi was all >0.7. Therefore, CPAx-Chi had good reliability.

### Best cut-off point, sensitivity and specificity of CPAx-Chi

Typically, evaluation of diagnostic performance is based on the ROC curve and AUC. If the AUC of a certain scale is 1, then it is considered to be a 'perfect' diagnostic tool, but the perfect tool does not exist in the real world. Hence, if the AUC of one scale ranges from 0.85 to 0.95, then the measurement effect of the scale is very good. If the AUC of one scale ranges from 0.5 to 0.7, then the measurement effect of the scale is considered to be undesirable. If the AUC of one scale is 0.5, then the measurement effect of the scale is barely functional.[32–34] Our experts regarded an MRC-Score ≤48 as the standard to diagnose ICU-AW. First, some studies have demonstrated the value of diagnostic ICU-AW using the Barthel Index,[35] grip strength,[36] ICU Mobility Scale,[37] de Morton Mobility Index[38] and the Physical Function Intensive Care Test[39] using MRC-Score ≤48 as the standard. Second, the best cut-off point, sensitivity and specificity of neuromuscular ultrasound, electrophysiological recordings, electromyography and other objective diagnostic methods used to diagnose ICU-AW have been verified using MRC-Score ≤48 as the criterion.[26 40–42] Third, scholars have constructed several models of early prediction of ICU-AW by taking MRC-Score ≤48 as a diagnostic criterion.[43–45] In the present study, the best cut-off point for the diagnosis of ICU-AW with CPAx-Chi was 31 points. This was verified by taking MRC-Score ≤48 as the criterion, and the sensitivity and specificity were good.

The kappa statistic quantifies inter-rater reliability for ordinal and nominal measures. In general, a kappa value between 0.40 and 0.60 indicates 'moderate' agreement, between 0.61 and 0.80 denotes 'substantial' agreement, and >0.81 reflects 'excellent' agreement; a negative value for kappa represents disagreement.[46 47] The concordance of the kappa value was high when taking the MRC-Score ≤48 and CPAx-Chi ≤31 as the best cut-off points to diagnose ICU-AW for researcher A and researcher B.

### Strengths and limitations

This is the first study about Chinesisation, adaptation and validation of the CPAx in critically ill patients. Second, there were two researchers assessed and collected data independently, which improved the reference value of the validation data. Third, this is the first study that demonstrated the best cut-off point for the diagnosis of ICU-AW using CPAx-Chi. However, there are some limitations in the study. First, this is a non-randomised pool of participants chosen primarily by their availability during the study period. Second, there were specific exclusion criteria that may have stopped the potential 'ceiling and floor' effects of CPAx-Chi to be tested. Third, there is no 'gold standard' for diagnosis of ICU-AW. but we took

MRC-Score ≤48 which was the most widely used 'standard' for ICU-AW as the criterion to demonstrate the best cut-off point for the diagnosis of ICU-AW using CPAx-Chi. In the future, we need to take other tools just like the maximum inspiratory pressure, MRC-Score, electromyography and neuromuscular ultrasound as the criterion to further demonstrate the best cut-off point for the diagnosis of ICU-AW using CPAx-Chi.

## CONCLUSIONS

We have demonstrated that CPAx-Chi had high criterion validity and reliability for assessing ICU-AW in adult patients in the ICU. CPAx-Chi showed high sensitivity and specificity in assessment of patients at risk of ICU-AW at a recommended cut-off of 31 points. To further confirm the clinical value of CPAx in assessing and diagnosing ICU-AW, it must be applied together with the MRC-Score, ultrasound, electrophysiology and electromyography. Also, multicentre, large-sample and randomised trials are needed to verify the best cut-off point for CPAx.

**Author affiliations**
[1]Intensive Care Units, Lanzhou University First Affiliated Hospital, Lanzhou, China
[2]School of Nursing, Lanzhou University, Lanzhou, China
[3]Intensive care uinits, Lanzhou University First Affiliated Hospital, Lanzhou, China
[4]Nursing Department, First People's Hospital of Lanzhou City, Lanzhou, China
[5]Nursing Department, Henan Provincial People's Hospital, Zhengzhou, China
[6]Nursing Department, Sichuan University West China Hospital, Chengdu, China
[7]Evidence-Based Medicine Center, Lanzhou University, Lanzhou, China

**Acknowledgements** The authors gratefully acknowledge all the participants. The authors gratefully acknowledge Charlesworth Author Services which had helped us to revise the language in the manuscript.

**Contributors** ZZ and YW conceived of the study, participated in its design and coordination, and helped to draft the manuscript. GW, HW and BJ performed the experiment and investigation. ND and JT participated in the design of the study and performed the statistical analysis. BL and JG participated in the project administration. WY and JT participated in the manuscript editing and review. All authors read and approved the final manuscript.

**Funding** This study was based on projects of: (1) the Health Commission of Gan Su Province (project no. GSWSHL2020-11) and (2) the First Hospital of Lanzhou University (no. ldyyyn2019-104).

**Disclaimer** The funding source had no involvement in data collection/analyses or manuscript preparation.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** The study was reviewed by the Ethics Committee of the First Hospital of Lanzhou University (LDYYLL2019-232) in Lanzhou, China.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data may be obtained from a third party and are not publicly available. Data can be requested from the Ethics Committee (contact via the First Hospital of Lanzhou University, Lanzhou, Gansu, China; email: ldyylwh@126.com) for researchers who meet the criteria for access to confidential data.



**ORCID iD**
Yuchen Wu http://orcid.org/0000-0002-1278-1698

## REFERENCES

1 Piva S, Fagoni N, Latronico N. Intensive care unit-acquired weakness: unanswered questions and targets for future research. *F1000Res* 2019;8:508. doi:10.12688/f1000research.17376.1
2 Tepper M, Rakic S, Haas JA, *et al*. Incidence and onset of critical illness polyneuropathy in patients with septic shock. *Neth J Med* 2000;56:211–4.
3 Qiu Y, Jiang L, Xi X. [Early incidence and prognosis of ICU-acquired weakness in mechanical ventilation patients]. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue* 2019;31:821–6.
4 Diaz Ballve LP, Dargains N, Urrutia Inchaustegui JG, *et al*. Weakness acquired in the intensive care unit. incidence, risk factors and their association with inspiratory weakness. observational cohort study. *Rev Bras Ter Intensiva* 2017;29:466–75.
5 Appleton RT, Kinsella J, Quasim T. The incidence of intensive care unit-acquired weakness syndromes: a systematic review. *J Intensive Care Soc* 2015;16:126–36.
6 Shepherd SJ, Newman R, Brett SJ, *et al*. Pharmacological therapy for the prevention and treatment of weakness after critical illness: a systematic review. *Crit Care Med* 2016;44:1198–205.
7 Fan E, Cheek F, Chlan L, *et al*. An official American thoracic Society clinical practice guideline: the diagnosis of intensive care unit-acquired weakness in adults. *Am J Respir Crit Care Med* 2014;190:1437–46.
8 Saccheri C, Morawiec E, Delemazure J, *et al*. ICU-acquired weakness, diaphragm dysfunction and long-term outcomes of critically ill patients. *Ann Intensive Care* 2020;10:1.
9 Jung B, Moury PH, Mahul M, *et al*. Diaphragmatic dysfunction in patients with ICU-acquired weakness and its impact on extubation failure. *Intensive Care Med* 2016;42:853–61.
10 Supinski GS, Morris PE, Dhar S, *et al*. Diaphragm dysfunction in critical illness. *Chest* 2018;153:1040–51.
11 Corner EJ, Wood H, Englebretsen C, *et al*. The chelsea critical care physical assessment tool (CPAx): validation of an innovative new tool to measure physical morbidity in the general adult critical care population; an observational proof-of-concept pilot study. *Physiotherapy* 2013;99:33–41.
12 Corner EJ, Soni N, Handy JM, *et al*. Construct validity of the chelsea critical care physical assessment tool: an observational study of recovery from critical illness. *Crit Care* 2014;18:R55.
13 Corner EJ, Hichens LV, Attrill KM, *et al*. The responsiveness of the chelsea critical care physical assessment tool in measuring functional recovery in the burns critical care population: an observational study. *Burns* 2015;41:241–7.
14 Holdar U, Eriksson F, Siesage K, *et al*. Cross-cultural adaptation and inter-rater reliability of the Swedish version of the chelsea critical care assessment tool (CPAX-Swe) in critically ill patients. *Disabil Rehabil* 2019:1–5.
15 Astrup K, Corner EJ, Hansen MG. Translation and cross-cultural adaptation of the chelsea critical care physical assessment tool into danish. *Physiother Theory Pract* 2018:1–8.
16 Maneesriwongul W, Dixon JK. Instrument translation process: a methods review. *J Adv Nurs* 2004;48:175–86.
17 Sousa VD, Rojjanasrirat W, Valmi DS. Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: a clear and user-friendly guideline. *J Eval Clin Pract* 2011;17:268–74.
18 Brislin RW. Back-translation for cross-cultural research. *J Cross Cult Psychol* 1970;1:185–216.
19 Comrey AL. Factor-analytic methods of scale development in personality and clinical psychology. *J Consult Clin Psychol* 1988;56:754–61.
20 Wu M. *The statistical analysis practice of questionnaire: SPSS operation and application*. Chongqing, 2010.

21 PRICE LR. *Psychometric methods:theory into practice*. Washington: Guilford Publications, 2016.
22 Souza AC, Alexandre NMC, Guirardello EB. Psychometric properties in instruments evaluation of reliability and validity. Propriedades psicométricas Na avaliação de instrumentos: avaliação dA confiabilidade E dA validade. *Epidemiol Serv Saude* 2017;26:649–59.
23 Ren Y. *Reliability and validity of Chinese version of PURPOSE-T and the preliminary application in hospitalized patients*. Tianjin Medical University, 2019.
24 Bentley TGK, Cohen JT, Elkin EB, *et al*. Validity and reliability of value assessment frameworks for new cancer drugs. *Value Health* 2017;20:200–5.
25 Chen AC-Y, Lee MS, Chen W-J, *et al*. Assessment in orthopedic training-an analysis of rating consistency by using an objective structured examination video. *J Surg Educ* 2013;70:189–92.
26 Wu Y, Ding N, Jiang B, *et al*. [Diagnostic tools of intensive care unit acquired weakness: a systematic review]. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue* 2018;30:1154–60.
27 Tavakol M. The reliability of assessments: the Bayesian Cronbach's alpha. *Med Teach* 2017;39:561.
28 Trizano-Hermosilla I, Alvarado JM. Best alternatives to Cronbach's alpha reliability in realistic conditions: congeneric and asymmetrical measurements. *Front Psychol* 2016;7:769.
29 de Vet HCW, Mokkink LB, Mosmuller DG, *et al*. Spearman-Brown prophecy formula and Cronbach's alpha: different faces of reliability and opportunities for new applications. *J Clin Epidemiol* 2017;85:45–9.
30 Dupépé EB, Davis M, Elsayed GA, *et al*. Inter-rater reliability of the modified medical Research Council scale in patients with chronic incomplete spinal cord injury. *J Neurosurg Spine* 2019:1–5.
31 Maidan I, Freedman T, Tzemah R, *et al*. Introducing a new definition of a near fall: intra-rater and inter-rater reliability. *Gait Posture* 2014;39:645–7.
32 Bowers AJ, Zhou X. Receiver operating characteristic (ROC) area under the curve (AUC): a diagnostic measure for evaluating the accuracy of predictors of education outcomes. *J Educ Students Placed at Risk* 2019;24:20–46.
33 Janssens ACJW, Martens FK, Cecile JW. Reflection on modern methods: revisiting the area under the ROC curve. *Int J Epidemiol* 2020;49:dyz274
34 Ke B-S, Chiang AJ, Chang Y-CI. Influence analysis for the area under the receiver operating characteristic curve. *J Biopharm Stat* 2018;28:722–34.
35 Kawahara K, Suzuki T, Yasaka T, *et al*. Evaluation of the site specificity of acute disuse muscle atrophy developed during a relatively short period in critically ill patients according to the activities of daily living level: a prospective observational study. *Aust Crit Care* 2017;30:29–36.
36 Cottereau G, Dres M, Avenel A, *et al*. Handgrip strength predicts difficult weaning but not extubation failure in mechanically ventilated subjects. *Respir Care* 2015;60:1097–104.
37 Tipping CJ, Bailey MJ, Bellomo R, *et al*. The ICU mobility scale has construct and predictive validity and is responsive. A multicenter observational study. *Ann Am Thorac Soc* 2016;13:887–93.
38 Sommers J, Vredeveld T, Lindeboom R, *et al*. De Morton mobility index is feasible, reliable, and valid in patients with critical illness. *Phys Ther* 2016;96:1658–66.
39 Nordon-Craft A, Schenkman M, Edbrooke L, *et al*. The physical function intensive care test: implementation in survivors of critical illness. *Phys Ther* 2014;94:1499–507.
40 Witteveen E, Sommers J, Wieske L, *et al*. Diagnostic accuracy of quantitative neuromuscular ultrasound for the diagnosis of intensive care unit-acquired weakness: a cross-sectional observational study. *Ann Intensive Care* 2017;7:40.
41 Mitobe Y, Morishita S, Ohashi K, *et al*. Skeletal muscle index at intensive care unit admission is a predictor of intensive care unit-acquired weakness in patients with sepsis. *J Clin Med Res* 2019;11:834–41.
42 Wieske L, Verhamme C, Witteveen E, *et al*. Feasibility and diagnostic accuracy of early electrophysiological recordings for ICU-acquired weakness: an observational cohort study. *Neurocrit Care* 2015;22:385–94.
43 Tzanis G, Vasileiadis I, Zervakis D, *et al*. Maximum inspiratory pressure, a surrogate parameter for the assessment of ICU-acquired weakness. *BMC Anesthesiol* 2011;11:14.
44 Peñuelas O, Muriel A, Frutos-Vivar F, *et al*. Prediction and outcome of intensive care Unit-Acquired paresis. *J Intensive Care Med* 2018;33:16–28.
45 Witteveen E, Wieske L, Sommers J, *et al*. Early prediction of intensive care unit-acquired weakness: a multicenter external validation study. *J Intensive Care Med* 2020;35:595–605.
46 Guo Y, Guo W, Qin Y. Consistency test based on Kappa coefficient and its software implementation. *Chinese J Health Stat* 2016;v.33:169–70.
47 Vanbelle S. A new interpretation of the weighted kappa coefficients. *Psychometrika* 2016;81:399–410.

# Approval letter of Ethics Committee of First Hospital of Lanzhou University

| | |
|---|---|
| **License Number**: LDYYLL2019-232 | **Date of Review:** 2019.09.17 |
| **Name of study protocol:** Chinesization, adaptation and validation of the Chelsea Critical Care Physical Assessment Tool in critically ill patients: A cross-sectional observational study | |
| **SFDA Number: -----** | **Category:** Scientific research |
| **Major researchers/ Units: Yuchen WU/ First Hospital of Lanzhou University** | |

**Review file (version number)**

| | | |
|---|---|---|
| Investigator's brochure | Yes ☐ | No ☑ |
| Study protocol | Yes ☑ | No ☐ |
| SFDA approvals | Yes ☐ | No ☑ |
| Drug inspection report | Yes ☐ | No ☑ |
| Technology contents | Yes ☐ | No ☑ |
| Informed consent | Yes ☐ | No ☑ |

**Results of the review by the Ethics Committee:**

1. At the committee conference, we have reviewed the technical proposal and the informed consent of the subjects. The results of the Committee's review of the proposal are as follows:

☑ **Approval** ☐ **Approval after revise** ☐ **Disapproval** ☐ **Suspension/ discontinue**

2. The study will be subject to ongoing review by the Ethics Committee during the process.

☑ **Yes** ☐ **No**

**3.** The frequency of review is every 12 months from the date of proposal approval. According to actual progress, the Ethics Committee has the right to change the frequency of continuous review.

**Signature by the Chairman: Xianlin Guo**

**Ethics Committee of the First Hospital of Lanzhou University**

**Date: 2019 /09/ 17**

## Attention

1. The study of "Approval" should follow the implementation of the program approved by the Ethics Committee and should conform to the principles of SFDA/GCP and the Helsinki Declaration.
2. Before submitting the “Approval after revise” protocol, the research plan should be marked item by item according to the evaluation opinions, which should be submitted to the Ethics Committee for review together with the preliminary opinion.
3. For “Disapproval” or “Suspension/ discontinue” protocol, the applicant and researcher may make a written complaint and state the reasons for the questions mentioned in the comments and recommendations of the Ethics Committee.
4. In the course of the research, the modification of the research protocol and the informed consent and other related documents should be examined and approved by the ethics committee before they can be implemented.
5. Serious or unexpected adverse events occurring in the Centre shall be reported to the SFDA and notified in writing to the Ethics Committee. The Committee shall have the power to make new decisions based on its assessment.
6. Regardless of whether the test starts or not, please submit an application for re-examination 1 month before the expiration of the continuing review date.

# 兰州大学第一医院伦理委员会批准函

| | | |
|---|---|---|
| 批件编号：LDYYLL2019-232 | 审查日期：2019.09.17 | |
| 研究方案名称:切尔西物理功能评估量表的汉化与临床应用 | | |
| 申办者/试验产品/SFDA 编号：---- | 类别 | 科学研究 |
| 主要研究者/单位：吴雨晨/兰州大学第一医院 | | |

审查文件（含版本号）

| | | |
|---|---|---|
| 研究者手册 | 有☐ | 无☑ |
| 研究方案 | 有☑ | 无☐ |
| SFDA 批件 | 有☐ | 无☑ |
| 药检报告 | 有☐ | 无☑ |
| 技术目录 | 有☐ | 无☑ |
| 知情同意书 | 有☐ | 无☑ |

伦理委员会审查结果：

1、委员会议对技术方案和受试者知情同意书进行了审查，委员会对该方案的审查决定如下：

☑ 同意 ☐ 修改后同意 ☐ 不同意 ☐ 终止或暂停

2、该研究进行过程中将接受伦理委员会的持续审查？ ☑ 是 ☐ 否

3、审查频率为方案批准之日起每 12 月一次。

伦理委员会有根据实际进展情况改变持续审查频率的权利。

主任委员签名：

兰州大学第一医院伦理委员会（盖章）：

日期：2019.9.17

注意：

1、"同意"的研究应遵循已经伦理委员会批准的方案执行，应符合 SFDA/GCP 和《赫尔辛基宣言》的原则。

2、"修改后同意"的研究方案在提交复审方案前，应按评审意见进行逐条修改并在修改处做出标记或说明，修改后的方案连同初审意见一并递交伦理委员会申请复审。

3、"不同意"和"暂停或终止"的研究方案，申办者和研究者可就伦理委员会的意见和建议中提及的问题作书面申诉，并陈述理由。

4、研究过程中对研究方案和知情同意书等相关文件所作的任何修改，均需得到伦理委员会审查同意后方可实施。

5、本中心发生的严重不良事件或意外不良事件需在向 SFDA 上报的同时向伦理委员会作出书面通报，伦理委员会有权根据对其评估做出新的决定。

6、无论试验开始与否，请在持续审查日到期前 1 个月提出再次审查的申请。