Measurement Invariance of 10-Item Resilience Scale Specific to Cancer in Americans and Chinese: A Propensity Score-based Multidimensional Item Response Theory Analysis

Mu Zi LIANG, MD, PhD, Peng CHEN, PhD, Alex Molassiotis, PhD, Sangchoon Jeon, PhD, Ying TANG, PhD, Guang Yun HU, PhD, Yun Fei Zhu, PhD, Zhe SUN, PhD, Yuan Liang YU, MS, M. Tish Knobf, PhD, Zeng Jie YE, PhD

PII: S2347-5625(22)00229-3
DOI: https://doi.org/10.1016/j.apjon.2022.100171
Reference: APJON 100171

To appear in: Asia-Pacific Journal of Oncology Nursing

Received Date: 7 October 2022
Revised Date: 21 November 2022
Accepted Date: 22 November 2022


This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 The Author(s). Published by Elsevier Inc. on behalf of Asian Oncology Nursing Society.
Manuscript title: Measurement Invariance of 10-Item Resilience Scale in Americans and Chinese with Cancer: A Propensity Score-based Multidimensional Item Response Theory Analysis

Corresponding author: Zeng Jie Ye

Article type: Original Article

Author Agreement Statement:
We the undersigned declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere. We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We understand that the Corresponding Author is the sole contact for the Editorial process. He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs.

Author Contributions:
Mu Zi Liang: Conceptualization, Methodology, Writing- Original draft preparation
Peng Chen: Conceptualization, Methodology, Writing- Original draft preparation
Alex Molassiotis: Methodology, Software
Sangchoon Jeon: Methodology, Software
Ying Tang: Visualization, Investigation
Guang Yun Hu: Visualization, Investigation
Yun Fei Zhu: Investigation
Zhe Sun: Investigation
Yuan Liang Yu: Investigation
M. Tish Knobf: Conceptualization, Methodology, Writing- Reviewing and Editing
Zeng Jie Ye: Conceptualization, Methodology, Writing- Reviewing and Editing

Role of the Funding Source:
This research was funded by grants from National Natural Science Foundation of China (No.72274043, 71904033), Young Elite Scientists Sponsorship Program by CACM (No.2021-QNRC2-B08), Humanity and Social Science Foundation of Department of Education of Guangdong Province (No.2020WTSCX009), Humanity and Social Science Foundation of Guangzhou (No.2021GZGJ57), Guangdong Research Center for TCM Service and Industrial Development, Guangzhou University of Chinese Medicine (2022ZDA03) and Humanity and Social Science Foundation of Guangzhou University of Chinese Medicine (No. 2021SKYB07).

Declaration of interests:
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Measurement Invariance of 10-Item Resilience Scale Specific to Cancer in Americans and Chinese: A Propensity Score-based Multidimensional Item Response Theory Analysis

Mu Zi LIANG a, MD, PhD; Peng CHEN b, PhD; Alex Molassiotis c, PhD; Sangchoon Jeon d, PhD; Ying TANG e, PhD; Guang Yun HU f, PhD; Yun Fei Zhu g, PhD; Zhe SUN h, PhD; Yuan Liang YU i, MS; M. Tish Knobf j ***, PhD; Zeng Jie YE k **, PhD

* These authors contributed equally to this work and should be considered co-first authors
** These authors should be considered co-correspondence

Affiliations: a Guangdong Academy of Population Development, Guangzhou, Guangdong Province, China, 510600; b Basic Medical School, Guizhou University of Traditional Chinese Medicine, Guiyang, China, 550025; c School of Nursing, The Hong Kong Polytechnic University, Hong Kong, Hong Kong SAR; d School of Nursing, Yale University, Orange, CT 06477, USA; e Institute of Tumor, Guangzhou University of Chinese Medicine, Guangzhou, China, 510006; f Army Medical University, Chongqing Municipality, China, 400038; g Shenzhen People’s Hospital, Shenzhen, China, Guangdong Province, China, 518000; h The First Affiliated Hospital, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China, 510405; i South China University of Technology, Guangzhou, Guangdong Province, China, 510641; j School of Nursing, Yale University, Orange, CT 06477, USA; k Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China, 510006

Address correspondence to:
M. Tish Knobf, School of Nursing, Yale University, Orange, CT 06477, USA; Email: tish.knobf@yale.edu
Zeng Jie Ye, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China, 510006; E-mail: zengjiye@qq.com

Author Note: This study was not preregistered. The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Conflict of interest statement:
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
Measurement Invariance of 10-Item Resilience Scale in Americans and Chinese with Cancer: A Propensity Score-based Multidimensional Item Response Theory Analysis

Mu Zi Liang a *, MD, PhD; Peng Chen b *, PhD; Alex Molassiotis c *, PhD; Sangchoon Jeon d, PhD; Ying Tang e, PhD; Guang Yun Hu f, PhD; Yun Fei Zhu g, PhD; Zhe Sun h, PhD; Yuan Liang Yu i, PhD; M. Tish Knobf j **, PhD; Zeng Jie Ye k **, PhD

* These authors contributed equally to this work and should be considered co-first authors
** These authors should be considered co-correspondence

Affiliations: a Guangdong Academy of Population Development, Guangzhou, Guangdong Province, China, 510600; b Basic Medical School, Guizhou University of Traditional Chinese Medicine, Guiyang, China, 550025; c School of Nursing, The Hong Kong Polytechnic University, Hong Kong, Hong Kong SAR; d School of Nursing, Yale University, Orange, CT 06477, USA; e Institute of Tumor, Guangzhou University of Chinese Medicine, Guangzhou, China, 510006; f Army Medical University, Chongqing Municipality, China, 400038; g Shenzhen People's Hospital, Shenzhen, China, Guangdong Province, China, 518000; h The First Affiliated Hospital, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China, 510405; i South China University of Technology, Guangzhou, Guangdong Province, China, 510641; j School of Nursing, Yale University, Orange, CT 06477, USA; k Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China, 510006

Address correspondence to:
M. Tish Knobf, School of Nursing, Yale University, Orange, CT 06477, USA; Email: tish.knobf@yale.edu
Zeng Jie Ye, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong Province, China, 510006; E-mail: zengjieye@qq.com

Received: October 7, 2022, Accepted: November 21, 2022, Published: ****


ABSTRACT

Objective: Little is known about measurement invariance (MI) of resilience instruments in cancer care. This study was designed to examine MI of 10-Item Resilience Scale (RS-SC-10) in Americans and Chinese with cancer using propensity score-based multidimensional item response theory (MIRT) analysis.

Methods: A sample of 924 patients were enrolled from the Be Resilient to Cancer (BRTC) trial involving one hospital in America and three hospitals in China. Data were collected from the RS-SC-10 and Hospital Anxiety and Depression Scale (HADS). Propensity score matching (PSM) and MIRT were performed to evaluate Differential Item Function (DIF). Integrated Discrimination Improvement (IDI) and Net Reclassification Improvement (NRI) were used to indirectly estimate the MI through incremental prediction ability of MIRT-based score over total score.
**Results:** RS-SC-10 retained 10 items with monotonous thresholds and its original two-factor structure. Non-uniform DIF was recognized in Item 4 (P= 0.0011, Δ%β1= 4.15%) and Item 8 (P= 0.0017, Δ%β1= 5.99%). NRI ranged from 9.04% to 35.01% and IDI ranged from 8.82% to 20.60%.

**Conclusion:** Although partial measurement invariance has been identified between Americans and Chinese, RS-SC-10 remains a critical indicator to emotional distress in cancer care.

**Keywords:** Measurement invariance, resilience, RS-SC-10, Americans, Chinese, Propensity score matching, Multidimensional Item Response Theory

**INTRODUCTION**

Despite significant improvements in 5-year survival [1-2], psychosocial distress, (e.g., depression, anxiety, fear of cancer recurrence) persists in cancer survivors which is associated with a poorer Quality of Life (QoL) [3-4]. Resilience, defined as the ability to bounce back after a traumatic event, has become an important concept in research across disciplines [5]. For example, cancer survivors with higher resilience levels were reported to have lower psychological distress as well as better QoL [6]. Thus, resilience may be a critical indicator of QoL and data support use in descriptive and intervention studies. A new resilience instrument, the Resilience Scale Specific to Cancer (RS-SC), was developed based on Shift-Persist theory and Resilience Model to Breast Cancer [7-8]. Subsequent to the original 25 item scale, a shortened version
was created, the 10-item RS-SC (RS-SC-10) based on Item Response Theory (IRT) analysis [9-10], and validated in our Be Resilient to Breast Cancer (BRBC) trial [11-12]. RS-SC-10 is a psychometrically established resilience instrument for cancer survivor. We wanted to determine if RS-SC-10 had the property of measurement invariance (MI) in different ethnic groups (i.e., Americans Vs. Chinese), which is defined as the instrument captures the same underlying construct across distinct groups or time period [13]. Patients can vary in their interpretations, understanding and conceptualization of some items due to differences in language or in cultural assumptions [14]. For example, due to the difference in cultural background, Americans and Chinese may read, interpret and answer “item 9: I believe that good fortune will come after surviving a disaster, derived from RS-SC-10” quite differently. If we compare Americans and Chinese on a same scale, the “actual difference” in scoring may not be attributed to a “true difference”. These differences can be sources of bias resulting in nonequivalent constructs across ethnically diverse groups. In addition, confounders such as unbalanced demographics in most MI-based research are often neglected resulting in biased parameter estimation in MI [15-16]. In the current study, Propensity Score Matching (PSM) was performed for a confounder-balanced comparison and then followed by a multidimensional item response theory (MIRT) -based Differential Item Function (DIF) analysis. In addition, exploratory indicators including NRI (Net Reclassification Improvement) and IDI (Integrated Discrimination Improvement), were utilized in the Resilience-Distressed prediction models (MIRT-based Score Vs. Total Score) to provide
additional information about MI between Americans and Chinese. MI could be confirmed if similar NRI and IDI were identified between Americans and Chinese, indicating that MIRT-based score had higher net benefits and better predictable abilities to Distressed outcome than resilience total score. We hypothesized that: (1) RS-SC-10 would retain its original two-factor structure in Americans and Chinese; (2) uniform/non-uniform DIF would be identified in some items between Americans and Chinese; (3) Similar NRI and IDI would be identified between Americans and Chinese and MIRT-based score could offer incremental predictive value above total score.

METHOD

Language equivalence

The cross-cultural translation of RS-SC-10 was performed according to WHO’s guideline. Two independent Chinese-English native-speaking researchers translated the Chinese version into English [17]. The two translated scales were compared and consensus was reached by a panel of experts in the field of cancer care. The consensus version was then back-translated by a third bilingual English-Chinese researcher and the back translation English version was reviewed against the original Chinese version. No significant discrepancies were identified and the final version was approved and sent to 15 Americans with cancer for final validation.

Patients and data collection
One hospital in America and three hospitals in China participated in the current study. A potential sample of 1052 patients were initially approached between March 2018 and May 2021. Patients with missing demographic or instrument data were excluded, resulting in a final sample of 924 (87.8%). The inclusion criteria were: (1) confirmed diagnosis of cancer, (2) aged >18 years and (3) receiving active treatment. The exclusion criteria were: (1) linguistic or intellectual difficulties, (2) had a currently active Axis I psychiatric disorder, (3) unwilling to participate in the study. The patients were approached by trained research nurses and instruments were administered following informed consent.

Ethics approval

The current study was part of the Be Resilient to Cancer (BRTC) and details of ethic approval have been described [18-20].

Instruments

10-item Resilience Scale Specific to Cancer (RS-SC-10)

RS-SC-10 is derived from the original 25-item RS-SC, which has five domains of Generic Element, Benefit Finding, Support and Coping, Hope for The Future, and Meaning for Existence [7]. RS-SC-10 has two domains: Generic Elements and Shift-Persist [10, 21]. It is a 5-Likert scale with higher scores indicating higher resilience levels (total score ranges from 10-50). RS-SC-10 were attached in the appendix file.

Hospital Anxiety and Depression Scale (HADS)
HADS contains 7 items for anxiety and 7 items for depression, respectively, with higher scores indicating higher levels of anxiety and depression [22]. In this current study, a cut-off of 7 was applied to select patients with high anxiety or depression [23]. Patients indicated by anxiety or depression ≥ 7 were defined as Distressed (outcome=1) while others were defined as Non-distressed (outcome=0).

**Statistical analysis**

First, the propensity score was calculated by multivariable logistic regressions and demographics including gender, education, cancer type, stage of cancer and time since diagnosis associated with resilience were matched in PSM. To maximize the ratio while maintain a 90% match rate in the BRTC group, we used a 1-to-2 (Americans to Chinese) matching design using the nearest neighbor method within a caliper of 0.1 [24]. Absolute standardized difference < 0.1 indicates a good balance. PSM methods were successfully performed in our previous research [19, 25]. Second, local independence was examined by item-pair residual correlations and less than 0.2 is recommended [26]. Third, Confirmatory Factor Analysis-based and Bifactor-based MIRT models were explored in the current study [27]. A compensatory logistic multidimensional grade response model (MGRM-C) was performed to estimate the item parameters, which was detailed as the equation below:

\[
P_{ijk} = \frac{\exp(a_i \theta_j + d_{jk})}{1 + \exp(a_i \theta_j + d_{jk})}
\]

MGRM-C is a logistic probability model \((P_{ijk})\) that examinee \((j)\) will respond with
category $k$ (and above) of item $i$ as a function of the item-category threshold (or easiness parameter, $d_{ik}$), item discrimination parameter vector ($a_i$), and examinee ability parameter vector ($\theta_j$). Log-likelihood (LL), Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Sample-adjusted BIC (SABIC) were compared between Confirmatory Factor Analysis-based and Bifactor-based MIRT models. When the optimal model was chosen, Multidimensional Discrimination (MDISC, $<0.5$ indicates poor; 0.5-1.0, moderate; 1.0-1.5, good; $>1.5$, excellent) and Multidimensional Difficulty (MDIFF, monotonous distribution indicates good fitting) were utilized as the fitting indicators. Additionally, item trace and item information surface were also visualized. Fourth, an iterative hybrid ordinal logistic approach with Monte Carlo Simulation was used to evaluate Differential Item Functions (DIF) between Americans and Chinese [28]. $U_i$ is defined as a discrete random variable representing the ordered item response to item $i$, while $u_i (= 0, 1, \ldots, m_i-1)$ is defined as the actual response to item $i$ with $m_i$ ordered response categories. Based on the parallel regression assumption, regression coefficients are estimated for all cumulative logits with varying intercepts ($\alpha_k$). For each item, an intercept-only (null) model and three nested models are formed as follows:

**Model 0**: \[
\text{logit } P(u_i \geq k) = \alpha_k
\]

**Model 1**: \[
\text{logit } P(u_i \geq k) = \alpha_k + \beta_1 \times \text{ability}
\]

**Model 2**: \[
\text{logit } P(u_i \geq k) = \alpha_k + \beta_1 \times \text{ability} + \beta_2 \times \text{group}
\]

**Model 3**: \[
\text{logit } P(u_i \geq k) = \alpha_k + \beta_1 \times \text{ability} + \beta_2 \times \text{group} + \beta_2 \times \text{ability} \times \text{group}
\]

where $P(u_i \geq k)$ is defined as the cumulative probabilities that the actual item response $u_i$ falls in category $k$ or higher. Ability represents the latent trait measured by the test.
Uniform DIF is evaluated by comparing the log likelihood values for Models 1 and 2 while non-uniform DIF by Models 2 and 3. An overall test of total DIF is estimated by comparing Models 1 and 3. Bonferroni corrected $P$ values were set at $0.01/4=0.0025$ for domain of Generic Elements and $0.01/6=0.0017$ for domain of Shift-Persist. Fifth, two models including, Model 1: total score Vs. Model 2: MIRT-based score, were compared based on AUC, NRI (Net Reclassification Improvement), IDI (Integrated Discrimination Improvement), calibration curves (estimated by Brier score), Clinical Impact Curve and Decision Curve Analysis [29-31]. NRI focuses on reclassification tables constructed separately for participants with and without events and quantifies the correct movement in categories, while IDI focuses on differences between sensitivities and ‘one minus specificities’ for models with and without the new marker. All statistical analyses were performed by IRTPRO software and ‘lordif’, ‘mirt’, ‘riskRegression’, ‘nricens’, ‘PredictABEL’, ‘rms’ and ‘rmda’ packages in R software.

RESULTS

Demographics in the unmatched and matched data

In the unmatched dataset (N=924), mixed, breast and lung cancer diagnoses accounted for 30.6%, 29.1%, 25.8%, respectively. In Figure 1A, before PSM, significant demographic differences were noted between Americans and Chinese in gender ($P=0.0004$), education ($P=0.0234$), cancer type ($P=0.0003$), stage of cancer ($P<0.0001$) and time since diagnosis ($P<0.0001$). After PSM, 186 Americans (90.7%)
and 372 Chinese (51.7%) were included for further MIRT analysis and the demographics were well balanced by small standardized differences (ranged from 0.01 to 0.08), which was presented in Figure 1A.

**Item distribution and local independence**

The item distribution for Americans and Chinese were summarized in Figure 1B and 1C, respectively. In addition, item-pair local independence was visualized in Figure 1D and 1E, indicating local independence was satisfied.

**Confirmatory Factor Analysis-based Vs. Bifactor-based MIRT models**

Although better fitting indicators (P < 0.0001 for both Americans and Chinese) were identified in the Bifactor-based MIRT model (Model 2, Figure 2B) compared with Confirmatory Factor Analysis-based model (Model 1, Figure 2A), negative Slope values (S1 and S2) were recognized in Model 2 (i.e., S1: -0.43 and -0.18 for item 1 and 2 in Americans, respectively; S2: -0.09 and -0.73 for item 6 and 7 in Chinese, respectively). In addition, disordered thresholds were identified in Americans (item 6, MDIFF4) and Chinese (item 2, MDIFF4), indicating the information overextraction. Thus, Confirmatory Factor Analysis-based MIRT model (Model 1) was chosen as the optimal model according to the parsimonious model guideline. In Figure 2C and 2D, 10 item traces were visualized for Americans and Chinese respectively, and monotonous distribution of theta values were confirmed. Finally, Expected Total Score, Test Information and Test Standard Errors were summarized for Americans and Chinese in Figure 3 (A, B, and C) and Figure 3(D, E, and F), respectively.
Differential Item Functions (DIF)

For the domain of Generic Elements, Americans showed higher resilience compared with Chinese although the difference was not statistically significant (Figure 4A). At the individual level (Figure 4B), all items except for item 4 showed no DIFs. According to Monte Carlo simulations-based thresholds in Figure 4C, only non-uniform DIF was recognized in Item 4 (P= 0.0011, Δ\%β1= 4.15%). As for the domain of Shift-Persist, Americans showed lower resilience compared with Chinese although the difference was not statistically significant (Figure 4D). At the individual level (Figure 4E), all items except for item 8 showed no DIFs. According to Monte Carlo simulations-based thresholds in Figure 4F, only non-uniform DIF was recognized in Item 8 (P= 0.0017, Δ\%β1= 5.99%).

The Comparison of Different Prediction Models (Total Score Vs. MIRT-based Score)

Two models were developed to construct the prediction model for Distressed outcome, including Model 1: total score and Model 2: MIRT-based score. Compared with Model 1, AUC in Model 2 increased from 75.0%-86.5 to 89.2%-90.1% in four samples (Figure 5A). NRI ranged from 9.04% to 35.01% and IDI ranged from 8.82% to 20.60% (Figure 5A). In Figure 5B, brier scores in Model 2 ranged from 9.3 to 13.4, which were significantly less than those in Model 1 (ranged from 14.5 to 18.8). In Figure 5C, DCA indicated that Model 2 showed higher net benefits compared with Model 1. Thus, Model 2 had better predictable ability to Distressed outcome than
Model 1, and CICs about Model 2’ clinical utilization in different samples were
detailed in Figure 5D.

DISCUSSION

In the current study, an ethnically diverse sample of Americans and Chinese was used
to evaluate MI comparing two languages: English and Mandarin. Before MI
estimation, PSM was performed to help control the effect of potential confounders
resulting in a balanced demographic-comparison, which was not well addressed in
previous research [32]. The Confirmatory Factor Analysis-based MIRT model
(between-item multidimensional theory framework) was better than the Bifactor-
based MIRT model (within-item multidimensional theory framework) in
consideration of fitting indicators and information overextraction. It indicated that the
underlying resilience constructs were conceptualized comparable in both the
Americans and Chinese, and RS-SC-10 retained 10 items with monotonous thresholds
and its original two-factor structure. In addition, a 5-Likert scaling was suitable for
RS-SC-10 in consideration of non-disordered thresholds identified in the current
study. Although the local independence assumption was partly compromised owing to
several high (> 0.2) item-pair residuals associations (i.e., items 2 and 6, item 5 and 9,
etc), the problematic item-pair proportions were small and could be ignored. In
addition, Test Information revealed that RS-SC-10 could effectively distinguish
patients with lower-middle or upper-middle resilience, confirming Hypothesis 1.
As for Hypothesis 2, non-uniform DIF were identified in Item 4 and 8, indicating that item-means equivalence was partially compromised in the level of resilience construct. For example, non-uniform DIF was recognized in Item 4 and we concluded that Americans had higher responses than Chinese when the patients had low resilience (from -4 to -2), while Chinese had higher responses than Americans when the patients had moderate-high resilience (from -2 to 4). Similar results could be derived from Item 8. Based on these findings, item 4 and 8 might not be suitable outcomes for cross-culturally resilience-related intervention between Americans and Chinese because only invariant items could be administered to compare language groups [33]. In addition, more item parameters were estimated in a MIRT model than a unidimensional IRT model such as Rasch model, and a larger sample size should be warranted in the future research to replicate these findings especially for Americans in consideration of only 186 patients in the current study [34].

As for Hypothesis 3, MIRT-based Score Vs. Total Score were compared and anchored against Distressed outcome based on a cross-sectional design. Similar NRI and IDI were identified between Americans and Chinese and MIRT-based score had higher net benefits and better predictable abilities to Distressed outcome than resilience total score. Thus, MI could be indirectly confirmed by the Resilience-Distressed prediction model and MIRT-based score instead of total score should be utilized in clinical settings to have a more precise resilience screening as well as a potential tool for predicting emotional distress. However, PRO-related anchors had subjective bias and external validations, for instance, diagnoses elicited from trained physician, objective
indicators of cortisol, C-reactive protein, systolic pressure, etc., could be incorporated into MIRT validation research, which would provide more confidence to the interpretation of results [35-36].

LIMITATION

Several limitations should be considered. First, although PSM was performed to help control the effect of potential confounders, the sample size especially for Americans is inevitably reduced resulting in a decreased statistical power in MIRT analysis [37-38]. Therefore, the findings in the current study should be replicated in future research especially with a larger sample of Americans. Second, the sample is not balanced based on the cancer stage and more than half (53%) are patients with advanced stage cancer (IV). Thus, these findings derived from the current study might not be generalized to patients with early cancer stage and should be further validated. Third, compensatory logistic multidimensional grade response model (MGRM-C) was performed in this study, indicating that a higher response in the domain of Generic could compensate a lower response in the domain of Shift-Persist (linear accumulation), and this model might not be consistent with resilience construct resulting in biased parameter evaluations [39]. A non-compensatory MGRM, which means the probability is the product of probabilities derived from Generic and Shift-Persist, is recommended to test the robustness of findings in future research [39]. Fourth, in the current study, MI was estimated by Item Response Theory (IRT)-based methods and these findings should be further confirmed and compared by Classic Test Theory (CTT)-based Factor Model in future research [40].
CONCLUSION

Although partial measurement invariance has been identified between Americans and Chinese, RS-SC-10 remains a critical indicator to emotional distress in cancer care.

ACKNOWLEDGEMENTS

The authors would like to thank the participants with cancer who participated in this study.

AUTHORS’ CONTRIBUTION

Conceived and designed the analysis: M. Tish Knobf, Zeng Jie Ye

Collected the data: Mu Zi Liang, Peng Chen, Ying Tang, Guang Yun Hu, Yun Fei Zhu, Zhe Sun, Yuan Liang Yu

Contributed data or analysis tools: Alex Molassiotis, Sangchoon Jeon

Performed the analysis: Mu Zi Liang, Peng Chen

Wrote the paper: Mu Zi Liang, Peng Chen, M. Tish Knobf, Zeng Jie Ye

FUNDING

This research was funded by grants from National Natural Science Foundation of China (No.72274043, 71904033), Young Elite Scientists Sponsorship Program by CACM (No.2021-QNRC2-B08), Humanity and Social Science Foundation of Department of Education of Guangdong Province (No.2020WTSCX009), Humanity and Social Science Foundation of Guangzhou (No.2021GZGJ57), Guangdong Research Center for
TCM Service and Industrial Development, Guangzhou University of Chinese Medicine (2022ZDA03) and Humanity and Social Science Foundation of Guangzhou University of Chinese Medicine (No. 2021SKYB07).

DECLARATION OF COMPETING INTEREST

None declared.

ETHICS STATEMENT

This study was approved by the Clinical Research Ethics Board of Guangzhou University of Chinese Medicine (Approval No. 2016KYTD08).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Reference:


Figure caption:

Figure 1. Demographics, Item Distribution and Local Independence

Figure 2. Confirmatory Factor Analysis-based Vs. Bifactor-based MIRT Models

Figure 3. Expected Total Score, Test Information and Test Standard Errors

Figure 4. Differential Item Functioning Using Iterative Hybrid Ordinal Logistic Regression/Item Response Theory and Monte Carlo Simulations

Figure 5. The Comparison of Different Prediction Models (Total Score Vs. MIRT-based Score)
### Table

<table>
<thead>
<tr>
<th>Characteristics (%)</th>
<th>Before propensity score matching</th>
<th>After propensity score matching</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total (N=204)</td>
<td>American (N=100)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>403(45.0)</td>
<td>67(67.0)</td>
</tr>
<tr>
<td>Female</td>
<td>571(55.0)</td>
<td>33(33.0)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school or less</td>
<td>379(45.0)</td>
<td>70(70.0)</td>
</tr>
<tr>
<td>College or higher</td>
<td>540(55.0)</td>
<td>30(30.0)</td>
</tr>
<tr>
<td>Cancer stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast</td>
<td>289(42.9)</td>
<td>42(42.0)</td>
</tr>
<tr>
<td>Colon-rectal</td>
<td>485.3(55.0)</td>
<td>48.5(55.0)</td>
</tr>
<tr>
<td>Lung</td>
<td>238(21.5)</td>
<td>37(37.0)</td>
</tr>
<tr>
<td>Lymphoma</td>
<td>304(29.4)</td>
<td>30.4(29.4)</td>
</tr>
<tr>
<td>Prostate</td>
<td>540(55.0)</td>
<td>30(30.0)</td>
</tr>
<tr>
<td>Melanoma</td>
<td>205(29.8)</td>
<td>20.5(29.8)</td>
</tr>
<tr>
<td>Stage of cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>170(18.4)</td>
<td>17(17.0)</td>
</tr>
<tr>
<td>II</td>
<td>239(25.9)</td>
<td>31(31.0)</td>
</tr>
<tr>
<td>III</td>
<td>144(15.0)</td>
<td>14(14.0)</td>
</tr>
<tr>
<td>IV</td>
<td>371(40.0)</td>
<td>37(37.0)</td>
</tr>
<tr>
<td>Time since diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;6 months</td>
<td>170(18.4)</td>
<td>17(17.0)</td>
</tr>
<tr>
<td>6-12 months</td>
<td>239(25.9)</td>
<td>31(31.0)</td>
</tr>
<tr>
<td>&gt;12 months</td>
<td>144(15.0)</td>
<td>14(14.0)</td>
</tr>
</tbody>
</table>

### Figures

#### Figure A

- **Title**: Disease impact on quality of life
- **Details**: Comparison of disease impact on quality of life between American and Chinese patients before and after propensity score matching.

#### Figure B

- **Title**: Time since diagnosis impact on disease impact
- **Details**: Heatmap showing the impact of time since diagnosis on disease impact.

#### Figure C

- **Title**: Cancer stage impact on disease impact
- **Details**: Heatmap showing the impact of cancer stage on disease impact.

#### Figure D

- **Title**: Treatment effectiveness
- **Details**: Heatmap showing the treatment effectiveness.

#### Figure E

- **Title**: Overall disease impact
- **Details**: Heatmap showing the overall disease impact.