RESEARCH METHODOLOGY

Herth Hope Index: psychometric testing of the Chinese version

Keung Sum Chan, Ho Cheung William Li, Sally Wai-chi Chan & Violeta Lopez

Accepted for publication 15 October 2011

Abstract

Aim. This article is a report on psychometric testing of the Chinese version of the Herth Hope Index.

Background. The availability of a valid and reliable instrument that accurately measures the level of hope in patients with heart failure is crucial before any hope-enhancing interventions can be appropriately planned and evaluated. There is no such instrument for Chinese people.

Methods. A test–retest, within-subjects design was used. A purposive sample of 120 Hong Kong Chinese patients with heart failure between the ages of 60 and 80 years admitted to two medical wards was recruited during an 8-month period in 2009. Participants were asked to respond to the Chinese version of the Herth Hope Index, Hamilton Depression Rating Scale and Rosenberg’s Self-Esteem Scale. The internal consistency, content validity and construct validity and test–retest reliability of the Chinese version of the Herth Hope Index were assessed.

Results. The newly translated scale demonstrated adequate internal consistency, good content validity and appropriate convergent and discriminant validity. Confirmatory factor analysis added further evidence of the construct validity of the scale.

Conclusion. Results suggest that the newly translated scale can be used as a self-report assessment tool in assessing the level of hope in Hong Kong Chinese patients with heart failure.

Keywords: heart failure, Herth Hope Index – Chinese version, nursing, psychometric testing

Introduction

Heart failure is a major disease among older people, which is associated with significant morbidity and mortality, and has turned into a major public health concern. There is increasing recognition that hope influences the adaptation to illness and promotes wellness (Herth 1992). Lai et al. (2003) postulated that hope is a prerequisite to effective coping and decision-making, and has a protective function against the physiological and psychological stress of illness. Therefore, through the implementation of hope-enhancing strategies, the physical health and psychological well-being of patients with heart failure can be improved as a result of increasing levels of hope. First, however, assessing the level of hope is essential for the evaluation of the effectiveness of enhancing strategies. To produce a sound assessment, it is vital that there are valid...
and reliable instruments that accurately measure the level of hope in patients with heart failure.

One promising scale for measuring hope in adults is the Herth Hope Index (HHI), which was developed by Herth (1991). The generation of the HHI items was based on Dufault and Martocchio’s (1985) multi-dimensional concept of hope, which includes philosophical, theological, sociological, psychological and nursing perspectives of hope. The HHI was developed with three subscales, each representing one of the three combined domains of the conceptual model, defined as follows: (a) inner sense of temporality and future (cognitive-temporal), defined as the perception that a positive, desired outcome is realistically probable in the near or distant future; (b) inner positive readiness and expectancy (affective-behavioural), defined as a feeling of confidence with initiation of plans to affect the desired outcome; and (c) inter-connectedness with self and others (affiliative-contextual), defined as the recognition of the interdependence and interconnectedness between self and others and between self and spirit (Herth 1992).

Background

The HHI has been widely used in clinical settings and research (Phillips-Salimi et al. 2007). The psychometric properties of the HHI were initially tested by Herth (1992). The results demonstrated adequate internal consistency and test–retest reliabilities, and appropriate concurrent criterion-related and divergent validity. The factorial structure of the HHI was examined, and the results showed that there were three underlying factors, which is congruent with the hypothesized configuration of the factor structure proposed by Herth (1991). Since then, a number of studies have been carried out to evaluate the psychometric properties of the HHI in samples of adults and adolescents (Benzein & Berg 2003, Wahl et al. 2004, Phillips-Salimi et al. 2007). All these studies have generally supported the HHI as a suitable tool for investigating the level of hope in the various population studies. However, the factorial structure of the HHI needs further exploration. In addition, the HHI has not been translated into the Chinese language or used with Hong Kong Chinese people. As a result of cultural differences, some concepts or items in the original instrument may be inappropriate for people in another culture (Ferraz 1997); therefore, the findings may be inaccurate (Flaherty et al. 1988). Given these issues, before using a translated version of the scale in Chinese society, it was crucial to evaluate both its linguistic and cultural equivalence. In addition, the psychometric properties of the Chinese version of the HHI needed to be empirically tested.

The study

Aim

The aim of the study was to test the psychometric properties of the Chinese version of the Herth Hope Index.

Methodology

The study was divided into two stages: translating the HHI into Chinese language and testing the psychometric properties of the Chinese version of the HHI.

Stage 1: Translation of the HHI

The term ‘hope’ has been frequently used by Chinese to describe belief in a positive outcome or fulfilment related to the circumstance of one’s life, which is similar to the definition suggested in the Western literature. Nevertheless, before starting translation and using an instrument with a new cultural group, it is crucial to ensure the conceptual congruence of a construct. A bilingual clinical psychologist working with Chinese children was consulted to review the original scale for the Hong Kong Chinese context, and the items were found to reflect behaviours and emotions that are also typical for Chinese children.

The instrument was then translated and back-translated following the technique described by Bracken and Barona (1991). The items were first translated from the English language to the Chinese language by the researcher. Another translator, blinded to original items, completed the back-translation. The retranslated English version and the original English version were compared to see if the meaning of each item had been maintained.

Stage 2: Testing the psychometric properties of the Chinese version of the HHI Semantic and content equivalence

A panel of experts was set up to evaluate the translated scale. All the experts were bilingual, with experience in translation and validation of instruments. The panel of experts was asked to rate the equivalence of translation of each item using a 4-point rating scale (from 1 = not equivalent to 4 = most equivalent). Based on their responses, an equivalence rate was calculated. Any item deemed not equivalent by receiving a rating 1 or 2 by more than 20% of respondents would be amended. For the content equivalence, the expert panel was asked to rate each item on a 4-point scale (from 1 = not relevant to 4 = very relevant). The Content Validity Index (CVI) is the percentage of the total items rated as either 3 or 4. A CVI score of 80% or higher is generally considered to indicate good content validity (Polit & Beck 2008).
**Construct validity testing**
Convergent validity was established by finding correlations between scores on the Chinese version of the HHI and Rosenberg’s Self-Esteem Scale (RSES) using the Pearson product–moment correlation coefficient. A previous study indicated that people who perceived higher levels of self-esteem would entertain greater hope (Ritchie 2001). We hypothesized that there would be a positive correlation between the Chinese version of the HHI and RSES scores.

Discriminant validity was estimated by finding a correlation between scores on the Chinese version of the HHI and the Hamilton Depression Rating Scale (HDRS). There is some evidence that people with higher levels of hope will report fewer depressive symptoms (Davidson et al. 2007). We hypothesized that there would be a negative correlation between the Chinese version of the HHI and RSES scores.

To allow more precise testing of the configuration of the factor structures of the Chinese version of the HHI, confirmatory factor analysis (CFA) was carried out using LISREL version 8.8 for Windows (Scientific Software International Inc., Lincolnwood, IL, USA). The overall fit of the data model with the scale was examined by goodness of fit indices, including the chi-square/degrees of freedom ratio ($\chi^2$/d.f. ratio), Root Mean Square Error of Approximation (RMSEA), Comparative Fix Index (CFI) and Goodness of Fit Index (GFI). The $\chi^2$/d.f. ratio is a measure of global fit. A $\chi^2$/d.f. value between 1 and 5 indicates good fit (Byrne 1998). The RMSEA is an indication of model fit, and is based on the population discrepancy function, which is a standardized measure of error of approximation (Li et al. 2008). Browne and Cudeck (1993) claim that RMSEA values of up to 0.08 suggest a reasonable fit of the model in the population. The GFI is a measure of global fit between a theoretical model and the data, where a value of 0.9 or higher is considered to indicate a good model-data fit (Floyd & Widaman 1995). The CFI is the indicator of how much better the model fits compared with an independence model, where a value of 0.95 or higher indicates good fit (Hu & Bentler 1999).

Four models were tested based on previous factor analysis on the instrument (Herth 1992, Benzein & Berg 2003, Wahl et al. 2004, Phillips-Salimi et al. 2007).

**Reliability testing**
Internal consistency reliability of the translated instrument was assessed by calculating Cronbach’s alpha. The intraclass correlation coefficient (ICC) was used to estimate the test–retest reliability coefficient. These results indicated a reliability of 0.80 or higher, which is acceptable for the instrument to be used in research (Nunally & Bernstein 1994).

**Participants**
Hong Kong Chinese patients with heart failure admitted to two medical wards at two acute public hospitals were invited to participate. The inclusion criteria for the study were: (a) all participants should be between the ages of 60 and 80 years, and (b) they should be able to speak Cantonese and read Chinese. We excluded patients with psychiatric problems identified in their medical records.

Dixon (2005) recommends a 10:1 ratio, that is ten cases for each item to be factor analysed. Therefore, a purposive sample of 120 Hong Kong Chinese patients with heart failure was recruited during an 8-month period in 2009.

**Instruments**

**Herth Hope Index**
The HHI comprises 12 items. Participants are asked to rate each item on a 4-point Likert scale, and are scored from 1 to 4 ($1$ = strongly disagree, $2$ = disagree, $3$ = agree, $4$ = strongly agree); total possible scores thus range from 12 to 48, with higher scores indicating greater hopefulness. The psychometric properties of the HHI have been empirically tested (Herth 1992), showing appropriate concurrent criterion-related and divergent validity, adequate construct validity and good internal consistency and test–retest reliabilities.

**Rosenberg’s Self-Esteem Scale**
The RSES was designed to measure self-esteem as a global disposition (Rosenberg 1965). The RSES consists of ten items. Responses are rated using a 4-point Likert scale ranging from 1 (strongly disagree) to 4 (strongly agree), with total possible scores ranging from 10 to 40. Higher scores indicate higher levels of self-esteem.

The Chinese version of the RSES has been used with adults (Chou & Chi 2001, Ngai et al. 2007). The findings demonstrated acceptable psychometric properties for use with Chinese people.

**Hamilton Depression Rating Scale**
The HDRS contains 17 items pertaining to symptoms of depression experienced over the past week (Hamilton 1960). Nine of the items are scored on a 5-point scale, ranging from 0 to 4. The remaining eight items are scored on a 3-point scale, from 0 to 2. Scores can range from 0 to 54, with higher scores indicating greater number of symptoms.

The Chinese version of the HDRS has been used with adults (Zheng et al. 1988). The findings support the validity and reliability of the HDRS for use with Chinese people.
Data collection
Demographical data was collected on the day of recruitment. Participants were asked to respond to the Chinese version of the HHI, HDRS and RSES. To examine the stability of the Chinese version of the HHI, 20% of the participants \((n = 24)\) were randomly selected to respond to the Chinese version of the HHI again after 2 weeks.

Ethical considerations
The study was approved by the appropriate ethics committee. Written consent was obtained and all participants were assured of their confidentiality.

Results
Participant demographics
The demographical data are shown in Table 1. There were similar numbers of male and female heart failure patients in the sample.

Validity
Semantic and content equivalence
The equivalence rate was 95%, indicating that each item of the Chinese version of the HHI remained conceptually and idiomatically the same as in the English version. The CVI was 94%, indicating the valid content of the Chinese version of the HHI.

<table>
<thead>
<tr>
<th>Demographical</th>
<th>(n) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>68 (56.7)</td>
</tr>
<tr>
<td>Female</td>
<td>52 (43.3)</td>
</tr>
<tr>
<td>Age (years)</td>
<td></td>
</tr>
<tr>
<td>60–64</td>
<td>22 (18.3)</td>
</tr>
<tr>
<td>65–69</td>
<td>35 (29.2)</td>
</tr>
<tr>
<td>70–74</td>
<td>38 (31.7)</td>
</tr>
<tr>
<td>75–80</td>
<td>25 (20.8)</td>
</tr>
<tr>
<td>Education level</td>
<td></td>
</tr>
<tr>
<td>Primary</td>
<td>59 (49.2)</td>
</tr>
<tr>
<td>Lower secondary</td>
<td>47 (39.2)</td>
</tr>
<tr>
<td>Upper secondary</td>
<td>10 (8.3)</td>
</tr>
<tr>
<td>Tertiary</td>
<td>4 (3.3)</td>
</tr>
<tr>
<td>Religion</td>
<td></td>
</tr>
<tr>
<td>No religion</td>
<td>36 (30.0)</td>
</tr>
<tr>
<td>Buddhism</td>
<td>23 (19.2)</td>
</tr>
<tr>
<td>Taoism</td>
<td>34 (28.3)</td>
</tr>
<tr>
<td>Christian</td>
<td>27 (22.5)</td>
</tr>
</tbody>
</table>

Construct validity
Convergent validity was established by finding correlations between scores on the Chinese version of the HHI and RSES. There was a positive correlation between scores on the Chinese version of the RSES and HHI \((r = 0.40, n = 120, P < 0.01)\), indicating that patients with higher levels of self-esteem entertained greater hope. The results support the convergent validity of the HHI.

Discriminant validity was estimated by finding a correlation between scores on the Chinese version of the HHI and the HDRS. There was a negative correlation between scores on the Chinese version of the HHI and HDRS \((r = -0.40, n = 120, P < 0.01)\), indicating that high levels of hope were associated with fewer self-reported depressive symptoms. The results support the discriminant validity of the HHI.

Confirmatory factor analysis
The results of the four factor models are summarized in Table 2. The results revealed that the 12-item three-factor model developed by Herth (1992) was the best fit across all model-fit indices. The parameter estimates of this three-factor model are shown in Figure 1.

Reliability
The internal consistencies of the Chinese version of the HHI for the test and retest were found to have alpha coefficients of 0.89 and 0.80 respectively. Cronbach’s alpha coefficients for each of the four sub-scales are shown in Table 3. High item-total correlations, ranging from 0.51 to 0.79, were also found for responses to most items in the Chinese version of the HHI. The test–retest reliability coefficient at a 2-week interval was 0.86 (ICC value).

Discussion
Study limitations
Non-probability sampling limits the ability to generalize the results. In addition, only moderate correlations between scores on the Chinese version of the HHI and HDRS, and HHI and RSES were found. Further testing of the discriminant and convergent validity of the Chinese version of the HHI is recommended.

Psychometric properties of the Chinese version of the HHI
The discriminant validity of the Chinese version was supported. The data revealed that there was a moderate, negative correlation between scores on the Chinese version of
the HHI and those on the HDRS. The finding is consistent with a previous study (Davidson et al. 2007), showing that low levels of hope in people are associated with a higher risk of depression. The convergent validity of the newly developed scale was also supported. Our results revealed a moderate, positive correlation between the Chinese version of the HHI and RSES scores. The finding is consistent with a previous study (Ritchie 2001), showing that people with higher levels of self-esteem have greater hope.

To test the instrument’s factor structure more precisely, confirmatory factor analysis was performed to evaluate whether the proposed factor models in previous studies could adequately fit the data. Of all the factor models being evaluated, the unitary (single-factor) model was found to be the poorest fit to the data. These findings are consistent with the previous argument that hope is a multi-dimensional concept (Dufault & Martocchio 1985) and HHI was developed to capture this multi-dimensionality.

The results revealed that the 12-item three-factor model developed by Herth (1992) was the best fit across all model-fit indices. Overall, the results of the evaluation fit the HHI and those on the HDRS. The finding is consistent with a previous study (Davidson et al. 2007), showing that low levels of hope in people are associated with a higher risk of depression. The convergent validity of the newly developed scale was also supported. Our results revealed a moderate, positive correlation between the Chinese version of the HHI and RSES scores. The finding is consistent with a previous study (Ritchie 2001), showing that people with higher levels of self-esteem have greater hope.

To test the instrument’s factor structure more precisely, confirmatory factor analysis was performed to evaluate whether the proposed factor models in previous studies could adequately fit the data. Of all the factor models being evaluated, the unitary (single-factor) model was found to be the poorest fit to the data. These findings are consistent with the previous argument that hope is a multi-dimensional concept (Dufault & Martocchio 1985) and HHI was developed to capture this multi-dimensionality.

The results revealed that the 12-item three-factor model developed by Herth (1992) was the best fit across all model-fit indices. Overall, the results of the evaluation fit the HHI and those on the HDRS. The finding is consistent with a previous study (Davidson et al. 2007), showing that low levels of hope in people are associated with a higher risk of depression. The convergent validity of the newly developed scale was also supported. Our results revealed a moderate, positive correlation between the Chinese version of the HHI and RSES scores. The finding is consistent with a previous study (Ritchie 2001), showing that people with higher levels of self-esteem have greater hope.

To test the instrument’s factor structure more precisely, confirmatory factor analysis was performed to evaluate whether the proposed factor models in previous studies could adequately fit the data. Of all the factor models being evaluated, the unitary (single-factor) model was found to be the poorest fit to the data. These findings are consistent with the previous argument that hope is a multi-dimensional concept (Dufault & Martocchio 1985) and HHI was developed to capture this multi-dimensionality.

The results revealed that the 12-item three-factor model developed by Herth (1992) was the best fit across all model-fit indices. Overall, the results of the evaluation fit
What is already known about this topic
- Heart failure is associated with significant morbidity and mortality, and it has turned into a major public health concern.
- The Herth Hope Index is a promising scale for measuring hope in adults, which was designed for use in clinical settings and has been widely used in previous research studies.
- The Herth Hope Index has not been translated into Chinese or used with Hong Kong Chinese people.

What this paper adds
- The newly translated scale demonstrated adequate internal consistency, good content validity and appropriate convergent and discriminant validity. 
- Confirmatory factor analysis added further evidence of the construct validity of the scale.

Implications for practice and/or policy
- The newly translated scale can be used as a screening tool by nurses working in clinical settings to assess the levels of hope in Hong Kong Chinese patients with heart failure.
- Based on the level of hope in patients, hope-enhancing interventions to promote physical and psychological wellness can be appropriately planned, implemented and evaluated.

were convincing, and confirmed that there was a very good fit between the hypothesized model (Herth 1991) and the data.

The study results showed that the coefficients for both the internal consistencies and test–retest reliability of the new form were high. In addition, item-total correlation data indicated that each item was highly correlated with the total score. This suggests that the items in the scale are relatively homogeneous and are measuring the same psychological construct. Consistent with previous findings (Herth 1992, Benzein & Berg 2003, Wahl et al. 2004, Phillips-Salimi et al. 2007), this study provides empirical evidence of the reliability of the Chinese version of the HHI.

Conclusion
Despite some potential limitations, our results suggest that the Chinese version of the HHI can be used as a valid and reliable tool to assess the level of hope in Hong Kong Chinese patients with heart failure.

Funding
This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Conflict of interest
No conflict of interest has been declared by the authors.

Author contributions
CW, KS, SC and VL were responsible for the study conception, design and drafting of the manuscript. KS performed the data collection. CW and KS performed the data analysis. CW made critical revisions to the paper for important intellectual content and provided statistical expertise. KS, SC and VL supervised the study.

References


---

The *Journal of Advanced Nursing* (JAN) is an international, peer-reviewed, scientific journal. JAN contributes to the advancement of evidence-based nursing, midwifery and health care by disseminating high quality research and scholarship of contemporary relevance and with potential to advance knowledge for practice, education, management or policy. JAN publishes research reviews, original research reports and methodological and theoretical papers.

For further information, please visit JAN on the Wiley Online Library website: www.wileyonlinelibrary.com/journal/jan

**Reasons to publish your work in JAN:**
- **High-impact forum:** the world’s most cited nursing journal and with an Impact Factor of 1.540 – ranked 9th of 85 in the 2010 Thomson Reuters Journal Citation Report (Social Science – Nursing). JAN has been in the top ten every year for a decade.
- **Most read nursing journal in the world:** over 3 million articles downloaded online per year and accessible in over 10,000 libraries worldwide (including over 6,000 in developing countries with free or low cost access).
- **Fast and easy online submission:** online submission at http://mc.manuscriptcentral.com/jan.
- **Positive publishing experience:** rapid double-blind peer review with constructive feedback.
- **Early View:** rapid online publication (with doi for referencing) for accepted articles in final form, and fully citable.
- **Faster print publication than most competitor journals:** as quickly as four months after acceptance, rarely longer than seven months.
- **Online Open:** the option to pay to make your article freely and openly accessible to non-subscribers upon publication on Wiley Online Library, as well as the option to deposit the article in your own or your funding agency’s preferred archive (e.g. PubMed).

© 2011 Blackwell Publishing Ltd