

RESEARCH ARTICLE

Construct validity and reliability of the Sinhala version of the Chalder fatigue questionnaire in a cohort following dengue infection in Sri Lanka

ND Perera^{1*}, SP Prathapan³, DA Wass³ and A Wijewickrama¹

Ministry of Health, No 231, De Saram Place, Colombo 10.

National Institute of Infections Diseases, Mandavila Road, Angoda.

Faculty of Medicine, University of Sri Jayawardenapura, Gangodawila, Nugegoda.

Submitted: 17 May 2020; Revised: 30 April 2021; Accepted: 09 July 2021

Abstract: The objective of the study was to culturally adapt, translate and assess the validity and reliability of the 11-item Chalder fatigue questionnaire (CFQ) among adults (18–60 years) confirmed with dengue infection admitted to a tertiary care hospital in Colombo District, Sri Lanka. Modified Delphi technique was used in cultural adaptation and assessing face, content and consensual validity. A descriptive cross-sectional validation study was conducted among 110 patients. CFQ was administered one month after having dengue fever for assessing post-infectious fatigue (PIF). CFQ-Sinhala version (CFQ-S) was assessed for construct validity and reliability. Construct validity of CFQ-S was described with hypothesised scale structure and with confirmatory factor analysis. The culturally adapted CFQ-S confirmed the original two-factor structure among adults after one month of having dengue infection. CFQ-S demonstrated satisfactory internal consistency of ≥ 0.7 . Cronbach's alpha coefficient was 0.85 for the overall scale. The test-retest reliability was assessed by calculating the intra-class correlation coefficient between the two assessments and reported as 0.89 on the overall scale. The study confirmed satisfactory levels of validity and the reliability of the CFQ-S, a valid tool to screen for PIF. The two-factor model described by the original author was confirmed as the best fitting model by triangulation of results.

Keywords: Chalder fatigue scale, construct validity, dengue fever, post-infectious fatigue, reliability

INTRODUCTION

The global incidence of dengue has amplified 30-fold over the past fifty years. It is endemic in many tropical

and sub-tropical countries and the reporting of the first outbreak is also on the rise making it a universal concern. Bhatt *et al.* (2012) have projected the global incidence of dengue to be approximately 390 million cases per year, almost three times higher than the estimate of the World Health Organization (WHO). In Sri Lanka, the total confirmed as dengue infected for the year 2018 was 32989, of which 21.7 % (n = 7174) was reported from Colombo District (Ministry of Health, 2019).

The post-infectious period following a dengue infection is a relatively less studied research area. Post-infectious fatigue (PIF) after dengue infection and other viral infections have been observed in several studies. Fatigue is a subjective sensation of tiredness, lack of energy and exhaustion. When fatigue becomes chronic and accompanied by a disability, it is considered as an illness. Many physicians and researchers have looked into an entity named chronic fatigue syndrome (Gelder *et al.*, 2009). Fukuda *et al.* (1994) proposed a conceptual framework to define and study fatigue. Fatigue lasting more than or equal to one month is termed prolonged fatigue (Fukuda *et al.*, 1994).

A review of fatigue measuring scales by Hjollund (2007) reveals that there are 252 different approaches to assess fatigue. There are ad-hoc approaches, assessing by single questions, using multi-system scales, or by 'fatigue specific' scales. The most commonly used fatigue

* Corresponding author (nadeekayasanth@gmail.com)  <https://orcid.org/0003-2725-3086>



This article is published under the Creative Commons CC-BY-ND License (<http://creativecommons.org/licenses/by-nd/4.0/>). This license permits use, distribution and reproduction, commercial and non-commercial, provided that the original work is properly cited and is not changed in anyway.

precise scales are the fatigue severity scale, fatigue questionnaire/Chalder fatigue scale, multi-dimensional fatigue inventory, Piper fatigue scale, fatigue impact scale, etc. The Chalder fatigue questionnaire, (CFQ) fatigue impact scale and the Piper fatigue scale have been used to assess fatigue after infection (Hjollund *et al.*, 2007).

Several studies have assessed PIF following an infection of dengue. One study carried out in Singapore assessing PIF using the fatigue questionnaire (FQ)/Chalder fatigue questionnaire (CFQ) two months following hospitalisation has described this clinical entity. CFQ, a tool validated in several settings assesses the physical and mental dimensions of fatigue. Feeling of the presence of exhaustion and lack of energy has been measured in the physical fatigue section and subjective feeling of being psychologically exhausted with consensus on concentration, recall and speech has encompassed in the mental fatigue section (Chalder *et al.*, 1993; Seet *et al.*, 2007; Cella & Chalder, 2010).

Chalder fatigue questionnaire (CFQ)

CFQ was originally developed as a 14-item scale to measure the severity of fatigue, and special care was taken to develop the tool as a generic measure. Symptoms that are directly related to fatigue are included and those symptoms only associated with chronic fatigue syndrome are excluded from the tool. The rating of items is on a four-point Likert scale as in the general health questionnaire (GHQ). Two methods for scoring have been described as a bimodal system (GHQ method) and a four-point system (four-point Likert score) (Chalder *et al.*, 1993).

Initially, 275 newly registered patients at a general practice setting completed the CFQ. Another 100 consecutive attendees completed the CFQ and the fatigue section of the revised clinical interview schedule (CIS-R). Exploratory factor analysis was conducted in both sets of samples by principal component analysis (PCA) and a two-factor model; physical fatigue dimension and mental fatigue dimension has emerged. Considering the results, an 11- item fatigue scale was developed and it has shown better validity and reliability over the 14- item scale. The internal consistency reliability was calculated for the revised version and the Cronbach's alpha for the 11-item version was 0.89 with the physical fatigue and the mental fatigue sub-scales having Cronbach's alpha values of 0.845 and 0.821, respectively. Criterion validity had been assessed in the sample of 100 attendees by computing a two by two table and a receiver operating characteristic

(ROC) curve. The cut off value was decided at 3/4 with a sensitivity of 75.5 % and specificity of 74.5 %. These results should be interpreted with caution, since the CIS-R is not a 100 % criterion measure and considering the small sample size of 100 (Chalder *et al.*, 1993). There is evidence for post-infectious fatigue/persistent fatigue following a dengue infection from international as well as local studies. The CFQ has been translated and cross-culturally adapted in different settings such as Brazil and China and in Hong-Kong among diverse study populations (Cho *et al.*, 2007; Won & Fielding, 2010; Fong *et al.*, 2015; Jing *et al.*, 2016;). The study conducted in Brazil tested psychometric measures by Cho *et al.* (2007) among primary care attendees. A pilot study (n = 204) and a proper validation study (n = 304) had been conducted. Study participants were assessed with the CFQ and the fatigue section of the CIS-R. The Brazilian version of the fatigue scale was shown to reproduce the two dimensional factor structure following PCA. The Cronbach's alpha was 0.88 confirming satisfactory reliability (Cho *et al.*, 2007).

Wong and Fielding (2010) reported findings from their study on the construct validity of a Chinese version (Cantonese version) of the CFQ. The study participants (n = 201) were assessed by the Chalder fatigue scale, short form health survey (SF-12) and hospital anxiety and depression scale (HADS). Confirmatory factor analysis (CFA) was tested for one factor, two factor and a three factor model. A two-factor correlated model showed model fit, which was quite similar to the original English version. Good internal consistency was demonstrated with a Cronbach's alpha of 0.863 (Wong & Fielding, 2010).

Considering the local studies, Ball *et al.* (2011) had used CFQ in assessing fatigue among the general population and all participants were assessed with the CFQ Bradford somatic inventory (BSI) and the short form 36 Health Survey questionnaire. There were a total of 37 items and a confirmatory factor analysis was conducted via M-plus. In their study, they have included 13 items of fatigue as one sub-scale with the other two scales, which is not exactly similar to the CFQ scale used in the current study, which has only 11 items (Ball *et al.*, 2011). Although there were several local studies assessing the prevalence of fatigue among dengue patients the researchers were unable to gather evidence of a proper validation study of a tool measuring fatigue among patients with dengue infection or any other infectious disease. Therefore, it is considered as a timely requirement to evaluate the validity and reliability of a suitable tool to assess post-infectious fatigue among patients suffering from a dengue infection.

METHODOLOGY

The current study adopted a systematic process in selecting a suitable tool to assess PIF, cultural adaptation, translation to Sinhala language and to evaluate judgemental validity, construct validity and reliability.

Operationalisation of the concept of PIF was done with an extensive literature survey and listing down of all the available definitions and finalised with a group of experts in Medicine, Psychiatry, Neurology, Community Medicine and Immunology. The definition of post-infectious fatigue, following dengue infection was operationalised as 'a subjective feeling of tiredness, lack of energy and exhaustion lasting for at least one-month duration following dengue infection' (Fukuda *et al.*, 1994; Seet *et al.*, 2007; Gelder *et al.*, 2009).

A tool was selected based on certain criteria; tools originally developed in English language, after the year 1980 and with an acceptable level of validity and reliability. After reviewing nearly twenty tools, five tools were selected based on the operationalised definition and the context of the study. Fatigue severity scale -FSS (Krupp *et al.*, 1989), Piper fatigue scale-revised -PFS (Piper *et al.*, 1989), Chalder fatigue questionnaire -CFQ /Fatigue questionnaire - FQ (Chalder *et al.*, 1993), multidimensional fatigue inventory -MFI (Smets *et al.*, 1995) and fatigue section of SF-36 (Ware & Sherbourne, 1992) were selected for further review. A Modified Delphi technique was carried out in selecting the most appropriate tool for the current research. All the experts unanimously selected the CFQ as the best tool to assess the post-infectious fatigue in the first round. The underlying reasons for the selection of the CFQ were; it follows the operationalised definition; it is a simple yet a multidimensional tool; easy to administer with an acceptable level of validity and reliability; it has been used to assess post-infectious fatigue following dengue infection in Singapore and in Sri Lanka previously in an unpublished study (Seet *et al.*, 2007). Although it has been developed to assess fatigue severity in general practice settings, it has also been used in various situations to assess fatigue (Cella & Chalder, 2010).

Permission was obtained from the author to use the CFQ following cultural adaptation. Further discussions were conducted with the author about the suitability of the tool to assess post-infectious fatigue among dengue patients. The tool has been originally developed as a self-administered tool, considering the different educational level of the participants in the current context, CFQ was used as an interviewer-administered tool with the permission of the author.

Several techniques have been discussed in the literature on the translation of technical instruments. In this research, the forward and backward translation method was selected (Tsang *et al.*, 2017; World Health Organization, 2017).

Cultural adaptation of the CFQ was done using a Modified Delphi technique with a team of experts (n=8) from the fields of clinical medicine, psychiatry, community medicine, neurology and psychology. This iterative procedure was conducted in two rounds. During the first round, the panel was provided with a concept note explaining the objectives, detailed description on CFQ and information regarding the research. In the first round, the following were assessed based on a five-point Likert scale; relevance in assessing fatigue among adult patients with dengue in Sri Lanka and the appropriateness of the words used in the local context. They were further asked to indicate their suggestions on how an item should be modified if they assign a score less than or equal to three for an item. Further, expert opinion was obtained on the scoring method for the tool and the cut-off guideline. The mean scores ranged from 3 to 5 for an item. If an item received an average mark of less than three by more than 50 % of the expert panel, it was considered as the cut off to remove an item from the questionnaire. However, no items scored less than three and therefore no item was removed from the tool. The expert panel did not suggest any additional items to be added and as a result, the item structure remained unchanged.

Item number one, four, six and eleven had received equal to or less than an average mark of four, therefore these items were highlighted in the second iteration. The principal investigator (PI) discussed with the experts regarding the suggested amendments, and modifications were done to the above items. The older version and the modified version were presented with the average marks in the concept note of the second iteration. The expert panel was revisited and opinion was obtained regarding the cultural acceptability of those items again. After the second iteration, all the item scores were summarised. An average mark of more than four was received for all the items in the tool and the tool was finalised.

The culturally adapted and translated CFQ-S was pre-tested among ten patients, aged 18 – 60 years, admitted to Colombo South Teaching Hospital (CSTH), who were diagnosed with dengue infection. They were interviewed one-month post-infection. Validity refers to how accurately a study instrument measures the intended variable (Abramson & Abramson, 1999; Friss & Sellers, 2014). 'Fatigue' is a subjective and abstract phenomenon,

which does not have a concrete gold standard (Chalder *et al.*, 1993; Dittner *et al.*, 2004).

During the development stage of the original tool, Chalder *et al.* (1993) had considered the fatigue section of the revised clinical interview schedule (RCIS) as a gold standard and reported receiver operating curve (ROC) statistics. They have taken the optimal cut off as $\frac{3}{4}$ with a sensitivity of 75.5 % and a specificity of 74.5 %, which the results should be interpreted with this limitation. In the current study, the convergent validity or the criterion validity of the Chalder fatigue scale (S) was not considered due to the unavailability of the Sinhala validated version of CIS-R among adults in the local setting.

Therefore, a triangulation approach was used; with the use of several complementary validation methods which would provide the most accurate approximate assessment. Hence, judgmental validity and construct validity were assessed in the study (Abramson & Abramson, 1999).

Face validity of the CFQs was assessed, concerning subjective qualities, such as whether it assesses the level of fatigue among dengue patients. The responses by the study participants in the pre-test were referred in assessing the face validity.

There is evidence from the literature that the CFQ-English version is a valid scale to assess the severity of fatigue. During the translation process, measures were taken to ensure the semantic and conceptual equivalence of the translated version, so that the content validity will be agreeable in the translated version.

Further, the appraisal by the multi-disciplinary team of experts ($n = 8$) approved the content and consensual validity of the CFQ. Each item in the scale was assessed for the following; relevance in assessing PIF among adult dengue patients, appropriateness of the wording used and acceptability in the local context in assessing PIF among adult dengue patients.

Procedure in appraising construct validity

A descriptive cross-sectional validation study was conducted in a tertiary care hospital in Colombo, Sri Lanka, from April to June 2018. Adults, who were resident in Colombo District for the past six months, of 18-60 years of age and clinically diagnosed and confirmed as dengue infected by a consultant physician and/or by the presence of NS I antigen and/or dengue specific IgM in their serum, comprised the study population. Those who were diagnosed with a mental illness, pregnant mothers,

those who could not comprehend well and patients who were unable to respond to an interviewer-administered questionnaire in Sinhala were excluded. For the current study, considering a subject to variable (STV) ratio of 10:1, the sample size was calculated as 110 with an added 20 % to account for loss to follow up the total sample size accounted for 138.

The non-probability consecutive sampling method was used in recruiting the patients from the medical units in CSTH and they were followed up to one month post-infection at the hospital. The socio-demographic details were collected when recruiting patients on the date of discharge from the hospital. Culturally adapted translated and validated CFQ was used for assessment of fatigue at one month post-infection via an interviewer-administered questionnaire. Ethical approval was taken from the Ethics Review Committee of the University of Kelaniya and CSTH. Permission was taken from the Director/CSTH to conduct the study. Data were collected by the nursing graduates after a training on objectives of the study and data collection techniques.

Preparatory data analysis

Scoring was done according to the instructions given with the Chalder fatigue questionnaire. According to the scoring system, the higher scores indicated a greater level of fatigue (Chalder *et al.*, 1993). Before carrying out data analysis, the dataset was evaluated for appropriateness and compliance with the assumptions required by the analytical techniques in CFA.

The CFQ is based on continuous scores and had two sub-scales of fatigue, namely, physical fatigue and mental fatigue as described by the author (Chalder *et al.*, 1993). The scores were recorded on a four-point Likert scale. The values for each item varied from zero to three and the aggregate scores varied from zero to 33. The univariate standardised skewness values in all eleven items ranged from -0.765 to 4.045 and the univariate kurtosis values ranged from -2.010 to 4.690. In the sample, three and two items out of 11 items showed high skewness and kurtosis values, respectively.

The Kolmogorov-Smirnov test and the Shapiro-Wilk test are the other tests used to determine the normality of the dataset. In the sample, both tests were significant ($p < 0.05$). The results of all three techniques showed that the data were not normally distributed. Therefore, the Robust Maximum Likelihood (RML) estimation was used in conducting confirmatory factor analysis with LISREL software.

Adequacy of the sample size was assessed by Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, which reported a value of 0.847. Bartlett's test of sphericity is used to test the null hypothesis that variables in the sample correlation matrix are not correlated and showed a chi-square value of 517.021 (df = 55, p value < 0.001). The correlation matrix was observed in the data to evaluate inter-item correlations. Most of the correlations (68%) were more than 0.3, which is indicative of a satisfactory inter-item correlation for factor analysis. Almost all item combinations showed variance inflation factor (VIF) values less than 3, suggestive of absence of multicollinearity.

Data processing and statistical analysis

Univariate analysis was conducted by SPSS-21 version to analyse the descriptive data of the sample. Descriptive statistics are presented as frequency distributions.

Construct validity of the CFQ was assessed using two methods; assessment of the hypothesised scale structure and performing confirmatory factor analysis. To assess the hypothesised scale structure, a multi-trait scaling analysis was conducted using the SPSS-21 software using two methods. This procedure is centred on an analysis of the item scale correlations. Therefore, correlations of each item with the sub-scales were assessed. Depending on the item-scale correlations, item-convergent and item-discriminant validity were evaluated. In the first method, confirmation of the item convergent validity is defined as a correlation of 0.40 or greater between an item and its own sub-scale. Confirmation of the item discriminant validity is established by comparing the degree of the correlation with an item with its particular sub-scale in comparison with other sub-scales. Item discriminant validity was supported by two criteria. First, the highest correlation in a row is the correlation between the item with its own sub-scale, discriminating the other sub-scale. Second, by checking whether each item is correlating significantly with its own sub-scale; 'whether the correlation between an item and its hypothesised scale is more than two standard errors higher, than its correlation with the other scales'. The cut off value for detecting the significance level was calculated by multiplying the standard error of each item by 1.96 and subtracting the resulting value from the correlation score of its own sub-scale. Each item was considered as a scaling success if the particular cut off value is higher than the correlation value of the other sub-scale (Hays *et al.*, 1998).

In the second method, the Average Variance Extracted (AVE) and the Composite Reliability (CR), was assessed to confirm convergent validity. The recommended minimum of AVE to have a satisfactory convergent validity was 0.5, and the CR value should be > 0.7. To assess the discriminant validity, the AVE values were compared with the squared inter-construct correlation and to fulfill satisfactory discriminant validity, the AVE of each domain should be higher than the squared inter-construct correlation (Renko *et al.*, 2001).

There is evidence that the CFQ is composed of two sub-scales; physical fatigue and mental fatigue. (Chalder *et al.*, Cho *et al.*, 2007; Cella & Chalder, 2010; 1993; Won & Fielding, 2010). A confirmatory factor analysis was conducted to assess whether the hypothesised scale structure can be reproduced in the study sample via the Linear Structural Relations (LISREL) 8.8 software. In assessing the overall goodness of fit in model evaluation, several model fit indices were looked into, since each index would provide different information regarding the assessment model. They are described under three categories of model fit indices; absolute fit indices, relative fit indices, parsimony fit indices. It is recommended that at least one index from each category should be within the expected level to decide on the acceptability of the model. The following fit indices were assessed and the desired level for model fit is presented within parenthesis; Absolute fit indices (Satorra Bentler scaled chi-square test (p > 0.05), root mean square of approximation (RMSEA- < 0.08), goodness of fit index (GFI- > 0.90) and adjusted goodness of fit index (AGFI- > 0.90). The relative fit indices were; comparative fit index (CFI - > 0.95) and non-normed fit index (NNFI - > 0.95). parsimony fit indices were; (Parsimony Goodness of Fit Index (PGFI - >0.5) and parsimonious normed fit index (PNFI - >0.5) (Brown, 2006).

Assessment of CFA

CFA was assessed in two phases;

- i. In the first phase, two-factor model was assessed: The first seven items were loaded on to a subscale named 'physical fatigue (PF)' and the last four items were loaded on to a subscale named "mental fatigue (MF)", which has been evaluated by the original author.
- ii. In the second phase, the modifications suggested by the LISREL software to improve the model fit were

considered. Several modifications were done. Two error covariance was added between the two subscales and a path was drawn from mental fatigue item 3 (MF3) to physical fatigue (PF) sub-scale.

Assessment of reliability

The reliability of the CFQ was assessed, since it is an essential technique in predicting both random and systematic error in any measurement tool (Streiner *et al.*, 2015). The internal consistency and test-retest methods were used to measure reliability by SPSS version 21. Test re-test reliability of the tool measuring post-infectious fatigue was evaluated by administering the same tool to a sub-sample of 20 patients selected randomly, with an interval of seven days. Internal consistency was evaluated by computing the Cronbach's alpha of the post-infectious fatigue assessment tool. According to Nunnally's criterion, internal consistency estimates of a magnitude of 0.7 or greater was considered acceptable (Abramson & Abramson, 1999). For test-retest reliability, a correlation coefficient (Spearman's r) of 0.70 or greater was considered acceptable (Litwin, 1995).

RESULTS AND DISCUSSION

Selection, cultural adaptation and translation of the Chalder fatigue questionnaire

Operationalisation and selection of an appropriate tool to measure post-infectious fatigue was a major challenge at the planning stage, which was achieved with a systematic process as described in the methods section. Abramson and Abramson (1999) describe that by operationalising how we measure the outcome of interest should be expressed with objectively apparent details, should be easy to understand and unambiguous.

The possibility and the implications of developing a new tool vs. validating an existing tool were assessed extensively. At the planning stage, there were no Sri Lankan studies that had assessed PIF following dengue infection. There were only two local studies that had assessed fatigue among twins (general population) and Navy officers (Ball *et al.*, 2011; Hanwella *et al.*, 2014). A survey among Sri Lankan physicians revealed that 77 % were reported as post-viral fatigue (Kularatne, 2005). There was textbook evidence of the presence of fatigue following dengue infection. Yet those descriptions were not comprehensive (Kumar & Clark, 2005; Walker *et al.*, 2014). Apart from these findings little was known about the topic. On the other hand, 'fatigue' as a generic term

has been debated and assessed at many stages by different authorities over fifty years. Therefore, considering all these factors, it was finally decided that using an already established generic tool to assess post-infectious fatigue would have more scientific advantages than developing a new tool.

Modified Delphi technique was used in deciding the most appropriate tool, in the cultural adaptation process and in assessing face, content and consensual validity of the CFQ. Following cultural adaptation, the item number did not alter after two iterations. Item number one, four, six and eleven were modified. The CFQ was translated into Sinhala language giving due consideration to safeguard semantic equivalence and theoretical equivalence.

This objective was accomplished by involving translators with technical expertise as well as language expertise. During forward translation, the translators were provided with a guide to the translation process. Further, a team comprising the translators, supervisors and the PI discussed the suitability of each translated item. In contrast, to evade bias, the back translators were not given any explanation regarding the tool and kept blind to the original English version.

Pre-testing of the tool was conducted among a similar age group (18–60 years) and considering the similar eligibility criteria used in the study, they were not included as study participants.

Validity of the Sinhala version of the CFQ

Face, content and consensual validity were confirmed by another panel of experts in the fields of Community Medicine, Psychiatry, Neurology, Immunology, Psychology and Clinical Medicine, through a Modified Delphi process. Exploratory factor analysis was not considered in the study based on several factors; the items of the Chalder Fatigue Scale did not change after the cultural adaptation process and after evaluating judgmental validity, only the wordings of items one, four, six and eleven were changed to improve how the meaning of the items was delivered to the participants and there was prior evidence from the literature on its factor structure. In conducting the validation study to appraise construct validity, a total of 140 patients were recruited and 20 patients were lost to follow up, ending with a final sample of 120 (response rate – 85.7 %). The mean age was 29.6 years with a standard deviation of 10.1 years. The socio-demographic details are presented in Table 1.

Table 1: Characteristics of the study sample

Characteristic	Frequency (n=120)	Percentage (%)
Age of the patient		
18 - 25 years	54	45.0
26 - 35 years	35	29.2
36 - 45 years	19	15.8
46 - 55 years	12	10.0
Gender		
Male	73	60.8
Female	47	39.2
Status of dengue infection		
Dengue fever	52	42.3
Dengue haemorrhagic fever	68	57.7

Multi-trait scaling analysis

Item convergent and discriminant validity were tested for a two factor model using two methods.

Method I

The first seven items were included in the 'physical fatigue' domain and the last four items were included in the "mental fatigue" domain. Item convergent validity was established in the CFQ for the two-factor structure as each item correlates with its sub-scale with a correlation of > 0.4 . The item to physical fatigue domain correlations varied from 0.778 to 0.473. The items of mental fatigue domain correlations varied from 0.557 to 0.363. In the mental fatigue sub-scale, only one item was having a correlation of 0.363, and which approximates with 0.4. The item discriminant validity was supported since each item correlates more strongly with its own sub-scale than with the other sub-scale. Further, each item was assessed for item scaling; the correlation between an item and its own sub-scale was significantly higher (> 1.96 standard errors) than the correlation with the other sub-scales. All the 11 items showed success in item scaling. When calculating item-scale correlations, the own item was excluded from the scale total to adjust for inflation of the correlation (Hays *et al.*, 1998). The results are further described in Table 2.

Table 2: Multi trait correlation matrix for the Chalder fatigue questionnaire in two-factor model

Item	Physical fatigue	Mental fatigue	Standard Error	Cut off value (-1.96SE)	Scaling success
	sub-scale score	sub-scale score			
1. Do you feel tired? (PF1)	0.657	0.173	0.059	0.541	success
2. Do you need to rest more? (PF2)	0.763	0.361	0.055	0.655	success
3. Do you feel sleepy or drowsy? (PF3)	0.554	0.398	0.058	0.44	success
4. Do you have problems in starting activities? (PF4)	0.473	0.286	0.049	0.377	success
5. Do you lack energy? (PF5)	0.732	0.345	0.054	0.626	success
6. Do you lack muscle strength? (PF6)	0.631	0.172	0.053	0.527	success
7. Do you feel weak? (PF7)	0.778	0.231	0.054	0.672	success
8. Do you have difficulties in concentrating? (MF1)	0.374	0.520	0.049	0.424	success
9. Do you make slips of the tongue when speaking? (MF2)	0.091	0.363	0.033	0.299	success
10. Do you find it more difficult to find the correct word? (MF3)	0.143	0.441	0.031	0.38	success
11. Do you have problems in remembering things? (MF4)	0.362	0.557	0.055	0.449	success

Method II

Factor loadings were explored following varimax rotation and the factors were loaded in to two domains; first seven items in to one domain and the latter four into another domain. The average variance extracted (AVE) and the composite reliability (CR) values were calculated, and presented in Table 3.

Table 3 : Average variance extracted (AVE) and the composite reliability (CR) for the two factor model

Domain	AVE (≥ 0.5)	CR (≥ 0.7)
Physical Fatigue domain	0.6	0.9
Mental Fatigue domain	0.5	0.8

The AVE for both domains were ≥ 0.5 and the CR was ≥ 0.7 , and confirms satisfactory convergent validity (Rienko *et al.*, 2001). This model was further used to evaluate discriminant validity. The average variance extracted (AVE) in both constructs were compared with the squared inter-construct correlation. The Spearman correlation coefficient between the physical fatigue and the mental fatigue domains were 0.387. The squared inter-construct correlation was calculated as 0.1497. It was evident that the AVE in both domains were more than the squared inter-construct correlation, and confirms discriminant validity (Rienko *et al.*, 2001).

Confirmatory factor analysis

The two-factor model was tested and this model showed acceptable model fit in the fit indices representing all three categories. Next, the modifications suggested by the LISREL software were tested as described in the methods section. The results are presented in Table 4.

As described in Table 4, the Satorra-Bentler scaled chi-square test, RMSEA, GFI, AGFI, CFI, NNFI, PGFI and PNFI indices showed satisfactory levels for model fit. The two-factor structure consisting of physical and mental sub-scales was the most accepted factor structure of CFQ in literature, which was confirmed in this study as well (Chalder *et al.*, 1993; Cho *et al.*, 2007; Cella & Chalder, 2010; Won & Fielding, 2010).

There were suggestions from LISREL software regarding methods to further improve model fit. These methods of improving model fit included adding error covariance between the items and adding error covariance between an item and a factor. There was evidence from the literature on testing on suggested modifications by the software (Won & Fielding, 2010). The fit indices were presented for a modified two-factor model. Considering the model fit indices of the modified models, the values were slightly better for the two-factor model in chi-square statistic, RMSEA and the parsimony fit indices. The values were similar for GFI, CFI and NNFI. Considering the above all modifications, the two-factor model added with the suggested error covariance and the path change showed the highest satisfactory indices. though with the suggested modifications the model showed improved model fit, the changes in model fit indices did not vary much. Brown (2006) argues against adding correlated error terms to an already fitting model to improve model fit. Since these suggestions by the software solely depends on the data, it may affect the generalisability of the findings (Brown, 2006). Therefore, by considering all the results and expert opinion from the consultant psychiatrist and the statistician, the original two-factor model was selected as the best-fitted model in the current study. The standardised parameter estimates for the original two-factor model is presented in figure 1.

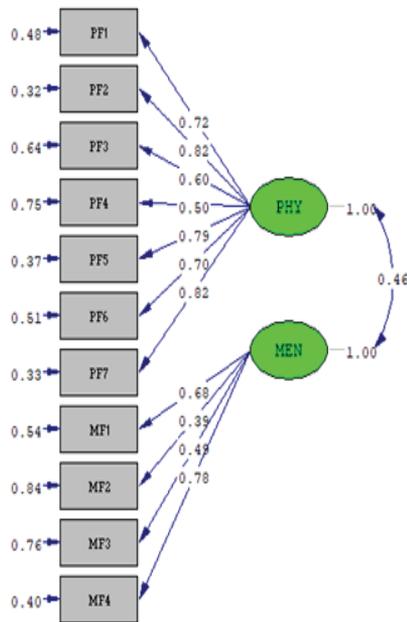
Table 4: Results of confirmatory factor analysis

Model	Absolute model fit indices				Relative fit indices		Parsimony fit indices	
	Chi-Square	RMSEA	GFI	AGFI	CFI	NNFI	PGFI	PNFI
Two factor-original	$X^2_{57.76}$ df = 43, p = 0.07	0.054	0.9	0.85	0.98	0.98	0.59	0.73
Modified two factor ^a	$X^2_{44.86}$ df = 41, p = 0.31	0.028	0.92	0.87	1.0	0.99	0.57	0.71
Modified two factor ^b	$X^2_{43.41}$ df = 40, p = 0.33	0.027	0.92	0.88	1.0	0.99	0.56	0.69

X^2 = Satorra-Bentler scaled chi square test (desired value $p > 0.05$), RMSEA = root mean square error of approximation (desired value < 0.08), GFI = Goodness of fit index (desired value > 0.9), AGFI = adjusted goodness-of-fit index (desired value > 0.9), CFI = comparative fit index (desired value > 0.95), NNFI = non-normed fit index (desired value > 0.95), PGFI = parsimony goodness of fit index (desired value > 0.5), PNFI = parsimonious normed fit index (desired value > 0.5)

^a set error covariance between mfl & pf4,

^b set error covariance between mfl & pf4 and pf2 & pf1



Chi-Square=57.86, df=43, P-value=0.06447, RMSEA=0.054

Figure 1: Standardised parameter estimates for the factor structure of the original two-factor model

Reliability

Validity and reliability are two complementing characteristics that improve the quality of data, bridging the phenomena of interest and the actual measurements. Considering the internal consistency, the Cronbach's alpha coefficient was 0.85 for the overall scale. Cronbach's alpha of 0.874 and 0.673 were reported for the physical fatigue and mental fatigue sub-scale, respectively. It is considered that the Cronbach's alpha coefficient of > 0.7 as having satisfactory internal consistency (Abramson & Abramson, 1999). It is discussed that the scales with a lesser number of items (less than ten) might get a Cronbach's alpha value up to 0.5 because the Cronbach's alpha is very sensitive to the number of items in a scale (Pallant, 2013). Hence the Cronbach's alpha of 0.673 for the mental fatigue sub-scale which contains four items was justifiable. During the development of the initial tool, the original author reported an overall Cronbach's alpha of 0.89 for the revised 11 item scale. For physical fatigue sub-scale and mental fatigue sub-scale, the Cronbach's alpha were 0.84 and 0.82, respectively (Chalder *et al.*, 1993). The test-retest reliability was assessed by calculating the intra-class correlation coefficient between the two assessments. The questionnaire was administered by the PI with an interval of one week. The correlation coefficients were more than 0.7, indicating good test-retest reliability and all the coefficients were statistically significant ($p < 0.001$).

CONCLUSION

In summary, the CFQ (S) was culturally adapted and validated ensuring scientific guidelines, as it was a relatively less investigated area locally and globally. The CFQ (S) had shown satisfactory validity and reliability among Sinhala conversant, 18 to 60 year old people at the local setting. Therefore, this study confirmed that the culturally adapted, translated and validated Sinhala version of the Chalder fatigue questionnaire is a valid tool to screen for post-infectious fatigue among adults who had dengue infection, amid those who were conversant with the Sinhala language.

Acknowledgements

Authors are grateful to the University of Sri Jayewardenepura for the financial support (ASP/01/RE/MED/2018/73) and Professor Trudie Chalder for granting permission to use the Chalder fatigue questionnaire with cultural adaptation.

Conflict of interest: The authors declare that there are no conflict of interests.

REFERENCES

- Abramson J.H. & Abramson Z.H. (1999). *Survey Methods in Community Medicine*. Churchill Livingstone; London, UK.
- Ball H.A., Siribaddana S.H., Sumathipala A., Kovas Y., Glozier N., Rijdsdijk F., McGuffin P. & Hotopf M. (2011). Genetic and environmental contributions to the overlap between psychological, fatigue and somatic symptoms: a twin study in Sri Lanka. *Twin Research and Human Genetics* **14**(1): 53–63.
DOI: <https://doi.org/10.1375/twin.14.1.53>
- Bhatt S., Gething P., Brady O., Messina J., Farlow A. & Moyes C. (2012). The global distribution and burden of dengue. *Nature* **496** (7446): 504–507.
DOI: <https://doi.org/10.1038/nature.12060>
- Brown T.A. (2006). *Confirmatory Factor Analysis for Applied Research*, Guilford Publications, Inc.72, Spring Street, New York, USA.
- Cella M. & Chalder T. (2010). Measuring fatigue in clinical and community settings. *Journal of Psychosomatic Research* **69**(1): 17–22.
DOI: <https://doi.org/10.1016/j.jpsychores.2009.10.007>
- Chalder T., Berelowitz G., Pawlikowska T., Watts L. & Wessely S.W.D. (1993). Development of fatigue scale. *Journal of Psychosomatic Research* **37**:147– 53.
- Cho H.J., Costa E., Menezes P.R., Chalder T., Bhugra D. & Wessely S. (2007). Cross-cultural validation of the Chalder Fatigue Questionnaire in Brazilian primary care. *Journal of Psychosomatic Research* **62**(3): 301–304.
DOI: <https://doi.org/10.1016/j.jpsychores.2006.10.018>

- Dittner A. J., Wessely S. C. & Brown R. G. (2004). The assessment of fatigue: A practical guide for clinicians and researchers. *Journal of Psychosomatic Research* **56**(2):157–170.
DOI: [https://doi.org/10.1016/S0022-3999\(03\)00371-4](https://doi.org/10.1016/S0022-3999(03)00371-4)
- Fong T.C., Chan J.S., Chan C.L., Ho R.T., Ziea E.T., Wong V.C., Ng B.F. & Ng S.M. (2015). Psychometric properties of the Chalder Fatigue Scale revisited: an exploratory structural equation modelling approach. *Quality of Life Research* **24**(9) : 2273-2278.
DOI: <https://doi.org/10.1007/s11136-015-0944-4>.
- Friis R.H. & Sellers T.A. (2014). *Epidemiology for Public Health Practice*, 5th Edition, Jones & Bartlett Publishers. Wall Street Burlington, USA.
- Fukuda K., Straus S.E., Hickie I., Sharpe M.C., Dobbins J.G. & Komaroff A. (1994). The chronic fatigue syndrome: a comprehensive approach to its definition and study. *Annals of Internal Medicine* **121**(12): 953–959.
DOI:<https://doi.org/10.7326/0003-4819-121-12-199412150-00009>.
- Gelder M.G., Andreasen N.C., Lopez-Ibor Jr J.J., & Geddes J.R. (eds.) (2009). *New Oxford Textbook of Psychiatry*, 2nd edition, Oxford University Press UK.
- Hanwella R., Jayasekera N.E. & De Silva V.A. (2014). Fatigue symptoms in Sri Lanka Navy personnel deployed in combat areas. *Ceylon Medical Journal* **59**(2): 39–44.
DOI: <http://dx.doi.org/10.4038/cmj.v59i2.7062>
- Hays R.D., Hayashi T., Carson S. & Ware J.E. (1998). *User's Guide for the Multitrait Analysis Program (MAP)*. The Rand Corporation, Santa Monica, USA.
- Hjollund N.H., Andersen J.H. & Bech P. (2007). Assessment of fatigue in chronic disease: a bibliographic study of fatigue measurement scales. *Health and Quality of life Outcomes* **5**(1): 12.
DOI: <https://doi.org/10.1186/1477-7525-5-12>
- Howard K.I. & Forehand G.A. (1962). A method for correcting item-total correlations for the effect of relevant item inclusion. *Educational and Psychological Measurement* **22**(4): 731–735.
DOI: <https://doi.org/10.1177/001316446202200407>
- Jing M.J., Lin W.Q., Wang Q., Wang J.J., Tang J., Jiang E.S., Lei Y.X. & Wang P.X. (2016). Reliability and construct validity of two versions of the Chalder fatigue scale among the general population in mainland China. *International Journal of Environmental Research and Public Health* **13**(1):147
DOI: <https://doi.org/10.3390/ijerph13010147>
- Krupp L.B., LaRocca N.G., Muir-Nash J. & Steinberg A.D. (1989). The fatigue severity scale: application to patients with multiple sclerosis and systemic lupus erythematosus. *Archives of Neurology* **46**(10) : 1121–1123.
DOI:<https://doi.org/10.1001/archneur.1989.00520460115022>
- Kularatne S.A. (2005). Survey on the management of dengue infection in Sri Lanka: opinions of physicians and pediatricians. *Southeast Asian Journal of Tropical Medicine and Public Health*. **36**(5): 1198.
DOI: <https://doi.org/10.1186/s13104-015-1085-0>.
- Kumar P. & Clark M. (2005). *Clinical Medicine*, 6th edition, Saunders Elsevier, London, UK.
- Litwin M.S. (1995). *How to Measure Survey Reliability and Validity*. Sage Publications, USA.
- Ministry of Health. (2019). *Epidemiology of Dengue*; National Dengue Control Unit: Ministry of Health, Colombo.
- Pallant J. (2013). *SPSS Survival Manual*, McGraw-Hill Education, UK.
- Piper B.F., Lindsey A.M., Dodd M.J., Ferketich S., Paul S.M. & Weller S. (1989). In: *Key Aspects of Comfort: Management of Pain, Fatigue, and Nausea* (eds. S.G. Funk, E.M. Tornquist, M.T. Champagne, L.A. Copp & R.A. Wiese), PP.199-240. Springer Publishing Company, New York, USA.
- Renko H.Y. & Autio E. (2001). Social capital, knowledge acquisition and knowledge exploitation in young technology-based firms. *Strategic Management Journal* **22**(6-7): 587–613.
DOI: <https://doi.org/10.1002/smj.183>
- Seet R.C., Quek A.M. & Lim E.C. (2007). Post-infectious fatigue syndrome in dengue infection. *Journal of Clinical Virology* **38**(1):1–6.
DOI:<https://doi.org/10.1016/j.jcv.2006.10.011>.
- Smets E.M., Garssen B., Bonke B.D. & De Haes J.C. (1995) The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *Journal of Psychosomatic Research* **39**(3): 315–25.
DOI:[https://doi.org/10.1016/0022-3999\(94\)00125-o](https://doi.org/10.1016/0022-3999(94)00125-o).
- Streiner D.L., Norman G.R. & Cairney J. (2015). *Health Measurement Scales: a Practical Guide to their Development and Use*. Oxford University Press, USA.
- Tsang S., Royse C.F. & Terkawi A.S. (2017). Guidelines for developing, translating, and validating a questionnaire in perioperative and pain medicine. *Saudi Journal of Anaesthesia* **11**(Suppl. 1): S8
DOI:https://doi.org/10.4103/sja.SJA_203_17
- Walker B.R., Colledge N.R., Ralston S.H. & Penman I.D. (eds.) (2014). *Davidson's Principles and Practice of Medicine*, 22nd edition. Churchill Livingstone, Elsevier Health Sciences London, UK.
- Ware Jr J.E. & Sherbourne C.D. (1992). A 36-item short-form health survey (SF-36): Results from the Medical Outcomes Study. *Med Care* **30**: 473– 83.
DOI:<https://doi.org/10.1080/17453674.2016.1181816>.
- Wong W.S. & Fielding R. (2010). Construct validity of the Chinese version of the Chalder Fatigue Scale in a Chinese community sample. *Journal of Psychosomatic Research* **68**(1): 89– 93.
DOI:<https://doi.org/10.1016/j.jpsychores.2009.05.008>.
- World Health Organization. (2017). *Process of Translation and Adaptation of Instruments*. World Health Organization, Geneva, Switzerland.