



Research paper

Development and Validation of a PROM to capture holistic outcomes in traditional, complementary and integrative medicine - The Warwick Holistic Health Questionnaire (WHHQ-18)

Nicola Brough^{a,*}, Sarah Stewart-Brown^b, Helen Parsons^c, Chinthana Perera^d

^a Institute: Honorary Research Fellow, University of Warwick, Warwick Medical School, Health Sciences Division, Coventry, CV4 7AL

^b Emeritus Professor in Public Health, University of Warwick, Warwick Medical School, Statistics and Epidemiology Department, Coventry, CV4 7AL

^c University of Warwick, Clinical Trials Unit, Coventry, CV4 7AL

^d Epidemiology Unit, Ministry of Health, No 231, De Saram Place, Colombo 10, Sri Lanka



ARTICLE INFO

Keywords:

Patient reported outcomes
Traditional
Complementary and integrative medicine
Craniosacral therapy
Holistic health
Complementary therapies

ABSTRACT

Introduction: Existing patient reported outcome measures (PROMs) do not capture all holistic outcomes observed in Traditional, Complementary and Integrative Medicine (TCIM). This study reports development and validation of a PROM to support research on craniosacral therapy (CST) and other TCIMs.

Methods: Using a conceptual framework and items developed and evaluated with clients and practitioners in a CST setting, a questionnaire was developed and tested using mixed methods approaches. Evaluation included an iterative process. Psychometric tests: structural validity (exploratory factor analysis, EFA), internal consistency (Cronbach's alpha), convergent validity (correlations with Warwick-Edinburgh Mental Wellbeing Scale, Short Form-12v2 (SF-12), Harry Edwards Healing Impact Questionnaire), repeatability and responsiveness (t-tests; intra class coefficients, ICC).

Results: The Warwick Holistic Health Questionnaire (WHHQ-18) was resolved covering mental, physical, emotional and spiritual wellbeing, self-awareness, engaging in life, responsibility for self, living in the moment and satisfaction with life.

EFA revealed four correlated sub-scales. Internal reliability was good ($\alpha = 0.852$). Convergent validity showed strong positive correlation with other wellbeing measures, but no correlation with health-related quality of life (SF-12). Repeatability testing showed good agreement ($ICC=0.822$) and no differences in scores for test-retest (paired t-test: $t = 0.355$, $p = 0.723$). Responsiveness analysis showed significant differences in scores (paired t-test $t = 6.15$, $p < 0.001$) with 46% of participants having an effect size of 0.5 or more.

Conclusion: WHHQ-18 is the first PROM developed for CST practice and captures outcomes important to TCIM more broadly. Good internal consistency, test-retest reliability and responsive at individual and group level make this new PROM an attractive resource for evaluators. Lack of convergent validity with SF-12 scales suggests WHHQ-18 be added to rather than replace HRQoL measures in clinical studies.

1. Introduction

Outcome measures used in studies of Traditional, Complementary and Integrative Medicine (TCIM) have, in the main, been adopted from other areas of healthcare, for example disease specific measures which do not detect broader holistic effects [1], and health-related quality of life measures (HRQoL) that do not cover the positive outcomes of TCIM. This mis-match between available measures and the therapies under study limits the efficacy of research in TCIM.

Although TCIM researchers have developed explicit conceptual frameworks which depict the effects of TCIM and show the gaps in available measures [1–4], patient reported outcome measures (PROMs) which cover the wellbeing and holistic outcomes of (TCIM) are lacking [5]. The aim of this study was to develop a PROM suitable for evaluation of craniosacral therapy (CST), a TCIM approach in which the need for research, evaluation and audit had been prioritised by the responsible professional association the Craniosacral Therapy Association UK (CSTA).

* Corresponding author.

E-mail addresses: Nicola.Brough@warwick.ac.uk (N. Brough), sarah.stewart-brown@warwick.ac.uk (S. Stewart-Brown), h.parsons@warwick.ac.uk (H. Parsons).

<https://doi.org/10.1016/j.eujim.2021.101375>

Received 14 March 2021; Received in revised form 15 August 2021; Accepted 16 August 2021

Available online 28 August 2021

1876-3820/© 2021 The Authors. Published by Elsevier GmbH. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

This study reports the development and validation of the Warwick Holistic Health Questionnaire (WHHQ), designed initially to capture outcomes important to clients and practitioners of craniosacral therapy (CST)¹. People present for CST, like for other TCIM, for diverse reasons including physical (often musculoskeletal) disease, mental illness (anxiety and depression) and pursuit of greater wellbeing particularly psychological and spiritual wellbeing [7]. Stress relief, support with rehabilitation and experience of an holistic approach to healthcare are other motivations for therapy [8]. It is not uncommon within the practice of CST for a client to present with a physical problem and for psycho/emotional roots of the problem to emerge during the therapeutic process. The proposed PROM thus needed to allow change to be captured in for the full range of possible outcomes not just the presenting complaint.

Clients of CST have attributed the following effects to treatment [9]: reductions in pain and disability, improved mobility, reductions in anxiety and depression, heightened sense of psycho-emotional awareness, improvements in self-concept, understanding mind-body-spirit links, improvements in interpersonal relationships, better coping strategies, enhanced engagement in self-care and capacity to manage health problems, a deepening sense of connectedness with self, others and the wider environment, and a general sense of wellbeing. Similar outcomes to those reported for CST have been reported across many TCIM modalities [2,4,10]. The proposed PROM thus needed to address all these aspects of wellbeing as well as symptoms, and to do so in the context of an holistic framework.

In summary therefore, the aim of this study was to develop a valid and reliable new PROM which covered the broad range of outcomes reported by people using CST and addressed the need for more appropriate outcome measures in TCIM more generally.

2. Methods

The WHHQ was developed and validated in accordance with FDA guidance [11] using mixed methods approaches. Consensus-based standards for the selection of health measurement Instruments known as the COSMIN checklist [12] were adopted as guidelines for reporting results, (see additional file 1). University of Warwick Biomedical Research Ethics Committee (BREC) approved the ethics for this study on 20th May 2015, REGO-2015-1499.

Both the conceptual framework and an early version of the WHHQ had been developed with advice and support of CST practitioners and clients. The Craniosacral Therapy Association (CSTA) of the UK had facilitated access to practitioners in private practice in the UK. Practitioners were recruited following informed consent as approved by Warwick University BSREC. Practitioners recruited clients who took part in focus groups and cognitive interviews and completed versions of the WHHQ. A recruitment poster and participant information leaflet were distributed amongst the CSTA membership for display within clinics. The questionnaire evaluation inclusion criteria were new or existing clients of recruiting practitioner, aged 16 or over, had a first language of English and good comprehension skills. Previous inclusion in earlier parts of the study, excluded clients from the questionnaire evaluation, clients who may be traumatised and unable to complete a questionnaire and clients who were receiving multiple treatment modalities during their sessions e.g., psychotherapy and CST, or acupuncture and CST.

¹ CST is a 'hands-on' TCIM which is thought to assist the body's natural capacity to self-repair. Developed from clinical experience within the field of osteopathy (6), it encompasses a therapeutic process between client and practitioner which emerges from the 'intention' of both to enter a relationship from which the potential for change may arise.

2.1. The conceptual framework

The conceptual framework (CF) for the WHHQ was based on the literature about TCIM [1,13,14] and CST outcomes [9]. 7 CST practitioners in two focus groups and 3 CST clients in one focus group had assessed an early version of the framework. Discussions were recorded transcribed and analysed thematically the CF was revised based on the findings [15]. The final version is available online [16]. It covers the dimensions of mental, physical, emotional and spiritual wellbeing as well as self-awareness, engaging in life, taking responsibility for self, being present and life satisfaction.

2.2. Development of Draft Warwick Holistic Health Questionnaire

A draft PROM with 73 items (see Additional file 2) had been developed to cover all the dimensions of this conceptual framework, using, where possible, data from a qualitative study of CST outcomes [7] to generate items from verbatim quotes. A five-point Likert scale was initially used for each response. Two rounds of qualitative semi-structured interviews with 6 clients together with a meeting with 16 CST practitioners had been used to evaluate these items, propose response options, recall period, and layout. 21 items were removed, 37 items were revised and the overall number was reduced to 52 (see Additional file 3). Different response options had been tested. As a result, the response options applied were 'Little or none of the time', 'Rarely', 'Sometimes', 'Often', 'All of the time'. Cognitive interviews had been undertaken with three clients to establish the face and content validity of the WHHQ-52. All interviews were recorded, transcribed and analysed thematically; findings have been reported elsewhere [15]. No changes were made to the content after the cognitive interviews, but the response options were revised to include the term 'most' resulting in 'all or most of the time' to avoid end-aversion bias.

2.3. Evaluating the WHHQ-52 and developing and evaluating the WHHQ-18

2.3.1. Overview

This was a multistage process involving data collection on the WHHQ-52 and scale reduction to 19 items using Exploratory Factor Analysis (EFA), followed by presentation to and discussion with practitioners in a structured consensus meeting. They required 6 items to be reinstated. The WHHQ-25 was evaluated on further samples allowing convergent validity, repeatability and responsiveness to be tested. EFA on the WHHQ-25 pointed to a need to reduce the scale again to 18 items. Data collected on these 18 items were extracted from the WHHQ-25 validation for further analysis.

2.3.2. Exploratory factor analysis on Warwick Holistic Health Questionnaire-52 and derivation of the Warwick Holistic Health Questionnaire-25

142 clients completed the WHHQ-52. Exploratory Factor Analysis (EFA) using the Promax technique was used to assess factor structure. Factors were extracted if their associated Eigenvalue was greater than 1. Items with communalities >0.8 or with absolute loadings <0.3 were removed and the EFA repeated with the new item set. If multiple items had the same loading value only one was removed. A total 33 items were removed. The resulting 19 item WHHQ was assessed at a structured consensus meeting with 60 CST practitioners. Concern was expressed about the removal of 11 of 33 items on the grounds of importance for face and content validity. Using a poll, participants were asked to identify the most important of these items on grounds of influence on response rates, responsiveness to change, and face and content validity. Six items ("I've had too many demands made on me"; "I've been sleeping well"; "I've felt my inner strength". "I've felt connected to my family and friends"; "I've asked for help when I've needed it"; "I've been able to express how I feel") were reinstated creating the WHHQ-25.

2.4. Evaluation of the WHHQ-25 and development and evaluation of the WHHQ-18

2.4.1. Data collection

To assess responsiveness, 146 clients were invited to complete the WHHQ-25 over two consecutive sessions of CST or cranial osteopathy. Responses were collected with either paper based or digital PROMs depending on practitioner preference. Three comparator measures were used to assess convergent validity: the Short Form-12 (SF-12) [17,18], Warwick-Edinburgh Mental-Wellbeing Scale (WEMWBS) [19] and the Harry Edwards Healing Impact Questionnaire (HEHIQ) [20].

Repeatability was assessed on a student population ($n = 109$) from Warwick Medical School using an online version of the WHHQ-25 in Qualtrics software [21]. Students completed the measure twice over a two-week interval with no intervention together with an anchor question at the second completion.

2.4.2. Statistical methods

Exploratory Factor Analysis using principal component analysis (PCA) and Promax rotation for correlated factor structures [22] was applied to first visit data to assess the factor structure of the WHHQ 25. Items with loadings less than 0.30 were suppressed [23]. The adequacy of the sample size was assessed by Bartlett's test of sphericity and the Kaiser-Meyer-Olkin test.

Internal reliability was assessed with Cronbach's alpha. Intra-class coefficients (ICC) and Standard Error of the Measurement (SEM) [22] was calculated as $SEM = SD(\text{baseline}) \times \sqrt{1 - \text{reliability of the instrument}}$.

Convergent validity was tested by magnitude and direction of the correlations with the three comparator measures to test the hypothesis of a strong positive correlation. The comparator measures covered all aspects of wellbeing: mental wellbeing WEMWBS [17], physical health (SF-12v2) [18] Physical Component Scale (PCS), mental/emotional health on SF-12v2 [17] Mental Component Scale (MCS) and spiritual wellbeing on the HEHIQ [20] outlook domain.

Repeatability was assessed using a paired t-test for change in mean scores on all students.

Responsiveness was assessed using distributional methods on first and second visit data; these methods are valuable in determining clinically importance differences [25]. Group level responsiveness was assessed using Cohen's D effect size with pooled SD as the denominator, and Standardized Response Mean (SRM) calculated by dividing mean change score by SD of the change scores. Individual level responsiveness was assessed using Cohen's D effect size calculated as the difference between pre and post assessment scores divided by pooled SD individually.

Meaningful change at group level was described using standard cut off values for Cohen's D; 'trivial' ($ES < 0.20$), 'small' ($ES \geq 0.20 < 0.50$), 'moderate' ($ES \geq 0.50 < 0.80$), or large ($ES \geq 0.80$) [26].

SRM was assessed with probability of change statistic: 0.5 and 1 representing no ability to detect and perfect ability to detect change respectively [27] and confidence intervals [27]. At individual level a cut off value of $ES > 0.5$ was taken as the threshold for meaningful change [28] and true change was assumed when more than 2.5% of the sample, had an increased/ decreased score greater than 2.77SEM [29].

Data were analysed using SPSS 24 software. Each item was scored 0 – 4 on the basis of a Likert scale [30]. Normality of the samples was assessed using visual inspection of histograms and Kolmogorov Smirnov test. Multivariate normality was assessed using Mahalobias distance method [23].

3. Results

3.1. Demographics of the samples

Three different samples were used during the validation of the

WHHQ. The demographics of participants and sample sizes are shown in Table 1.

3.2. Exploratory factor analysis on Warwick Holistic Health Questionnaire-25 and resolution of Warwick Holistic Health Questionnaire-18

Kolmogorov-Smirnov test and Shapiro-Wilk tests of normality for the total scores of WHHQ-25 were not significant ($p = 0.20$ and $p = 0.469$).

The Kaiser-Meyer-Olkin Measure of Sampling Adequacy was 0.872 which is above the recommended level minimum 0.6 [31]. Bartlett's test of sphericity was significant ($X^2 = 1628.02$; $df = 300$; $p < 0.001$) rejecting the null hypothesis that variables in the population correlation matrix are not correlated with each other [32].

The initial EFA of the WHHQ-25 illustrated a factor structure with six latent factors with an eigen value more than 1, explaining 63.35% of the variability of the observed items. Item 22 did not load on any factor with loading more than 0.33 and was therefore omitted from the second factor analysis.

Following a series of five EFAs, a 4-factor solution (18 items) was resolved with 63.0% of the total variance explained. In this structure, item Q 23 was cross loaded with factor 1 and factor 4, but as this cross loading was less than 75% ($0.337/0.534 = 63.1$), Q23 was kept in the WHHQ-18 (see Table 2). Four latent factors were identified (illustrated in Fig. 1) and named as factor 1: meaning, purpose, connectedness; factor 2: self-awareness, self-agency; factor 3: physical wellbeing; factor 4: emotional wellbeing. Of the 7 items dropped from the WHHQ-25 to resolve the WHHQ-18 (item numbers: 12, 17, 18, 19, 20, 22 and 24) three items [19,22,24] had been reinstated at the request of the practitioners at the meeting to discuss face and content validity of the original 19 item instrument resolved in the initial EFA.

3.3. Validity and reliability of Warwick Holistic Health Questionnaire-18

Reliability: Cronbach's alpha for the WHHQ-18 was 0.852.

3.4. Test-re-test reliability

The paired t-test suggested no difference ($t = .355$, $p = 0.752$) between mean scores at time 1 (mean: 64.21; SD:10.70) and time 2 (mean:

Table 1
Sample demographics.

Psychometric validation phase	Sample size	Age	GenderFemale Male	Total	
WHHQ-52 Structural validity (EFA)	N = 142	16-20 years	7	3	10
		21-39 years	26	3	29
		40 – 64 years	54	9	63
		65+ years	27	12	39
		Missing data	-	-	1
WHHQ-25 & 18 (CST n = 105; cranial-osteopathy n = 41) Structural validity (EFA)	N = 146	16-20 years	0	0	0
		21-39 years	1	2	3
		40 – 64 years	47	19	66
		65+ years	60	12	72
		Missing data	-	-	5
Internal reliability Responsiveness Convergent validity	N = 109	18-24 years	42	14	56
		25-30 years	36	17	53

Table 2
Pattern matrix of WHHQ-18.

		Meaning, purpose, connection	Self- awareness, /agency	Physical wellbeing	Emotional wellbeing
Q2	I've felt engaged in life	0.944			
Q1	My daily life has been full of things that keep me interested	0.931			
Q4	I've been satisfied by my school, work or current role in life	0.878			
Q3	I've felt my life has meaning	0.741			
Q5	I've felt joyful	0.651			
Q23	I've felt connected to my family and friends	0.534			0.337
Q7	My awareness about my health has helped me manage life		0.902		
Q9	I'm learning about myself and my body		0.887		
Q8	I've been able to stop and reflect		0.770		
Q10	I've felt in control of my health and well being		0.668		
Q11	I've felt aware of my body's needs		0.495		
Q13	I've been in pain			-0.858	
Q14	My symptoms have limited my daily activities			-0.852	
Q15	I've been physically well			0.707	
Q16	I've had lots of energy			0.487	
Q21	I've been sleeping well				0.880
Q6	I've felt calm				0.753
Q25	I've been able to express how I feel				0.666

63.97; SD: 12.65). The ICC (0.822) exceeded the recommended value of 0.7.

Convergent validity: There was a strong positive correlation between WHHQ-18 and both WEMWBS and HEHIQ as hypothesized. However, there was no significant correlation between the WHHQ-18 and SF-12v2 PCS and MCS scales as shown in Table 3.

Responsiveness: Mean age ($n = 138$) of the 146 clients with data on two consecutive visits was 50 years ($SD=15.9$), and the majority were female. Significant change was observed ($t = 6.15, p < 0.001$) between the mean baseline score of 44.71 ($SD=9.80$) and the mean post therapeutic intervention of 49.02 ($SD=9.13$) Table 4.

Cohen's D effect size was 0.45. The probability of change of SRM was

0.6950 (95th CI 0.337-0.682) which is within the range (0.5-1).

In individual level analysis, scores of 45.9% of participants showed an effect size change of 0.5 or more. The SEM was 3.7; 17.1% (increase) and 4.1% (decrease) of the participants respectively showed more than 2.5% increase or decrease of $2.77 \times SEM$ threshold.

4. Discussion

This study aimed to develop a valid and reliable new PROM which covered the broad range of outcomes of importance to health and wellbeing that have been documented in studies of outcomes of CST and TCIM approaches [1,13,33]. Both the conceptual framework on which the new PROM was built and item generation drew heavily on a prior qualitative study of outcomes undertaken with CST clients [33] enabling coverage of aspects of spiritual wellbeing and personal development not usually covered in PROMs.

CST practitioners had been engaged in early stages of the process of development of the scale and were involved again in this phase. This approach required an extra step to be added to standard approaches: the shorter measure (WHHQ-19) derived from initial EFA was taken back to practitioners for review. The latter were disappointed that psychometric testing had resulted in the removal of some items they considered important. It was agreed to reinstate 6 of these items and to test a 25-item measure. In the event half of the reinstated items performed poorly again and only 3 appeared in the final 18 item version. These items related to quality of sleep, feeling connected to family and friends and ability to express feelings. Practitioner involvement enhances content and face validity but can create tension within the methodological process and the trade-off between psychometric robustness and face validity. Practitioners are rarely concerned with the implications of psychometric testing, and choose PROMs based on the item content. If face and content validity are deemed as poor by practitioners, they are less likely to adopt the PROM. Whilst in the end the involvement of practitioners resulted only in minor changes to the scale, the process was considered important with regard to engagement with its use.

Prior to this study four PROMs had been developed specifically for use with TCIM two in the UK (The Measure Your Medical Outcomes Profile (MYMOP) [34] and the Harry Edwards Healing Impact Questionnaire (HEHIQ) [20]) and two in the US (Complementary and Integrative Medicine Outcomes Scale (CIMOS) [35] and Self-Assessment of Change (SAC) [2]). Of these the MYMOP is the most popular in the UK. MYMOP differs from the WHHQ in that change is measured based on issues identified by the clients at first assessment. So, MYMOP does not adequately capture the change in experiences, capabilities and mindsets identified as important in the development of the Conceptual Framework, which are not part of the client's presenting complaint or belief system about health improvement. Because of this conceptual difference we chose the second UK based measure (HEHIQ) as a comparator together with two very well validated and popular generic measures both of which had been recommended as suitable for evaluating TCIM the Warwick-Edinburgh-Mental-Wellbeing Scale [19] and the Short-Form-12 [18]. The HEHIQ is unusual in that it addresses spiritual wellbeing. The WEMWBS addresses mental wellbeing and the SF-12 is a highly regarded generic HRQOL measure which addresses physical, mental and emotional health as well as disability.

We observed the expected correlation with WEMWBS and HEHIQ but the lack of correlation with the SF-12 MCS and PCS was surprising. Health related quality of life captured by the SF-12 represents a different construct from the aspects of wellbeing captured by the WHHQ-18. However, clients of CST do report improvements in symptoms of disease and disability which have a negative impact on wellbeing, and the WHHQ-18 covers aspects of mental and physical health which are similar to those covered in the SF-12. In terms of the PCS, the WHHQ-18 items 'I feel in pain' and 'my symptoms limit my daily activities' would both be expected to correlate with those in the SF-12 PCS. In terms of emotional and mental wellbeing, WHHQ-18 correlated strongly with

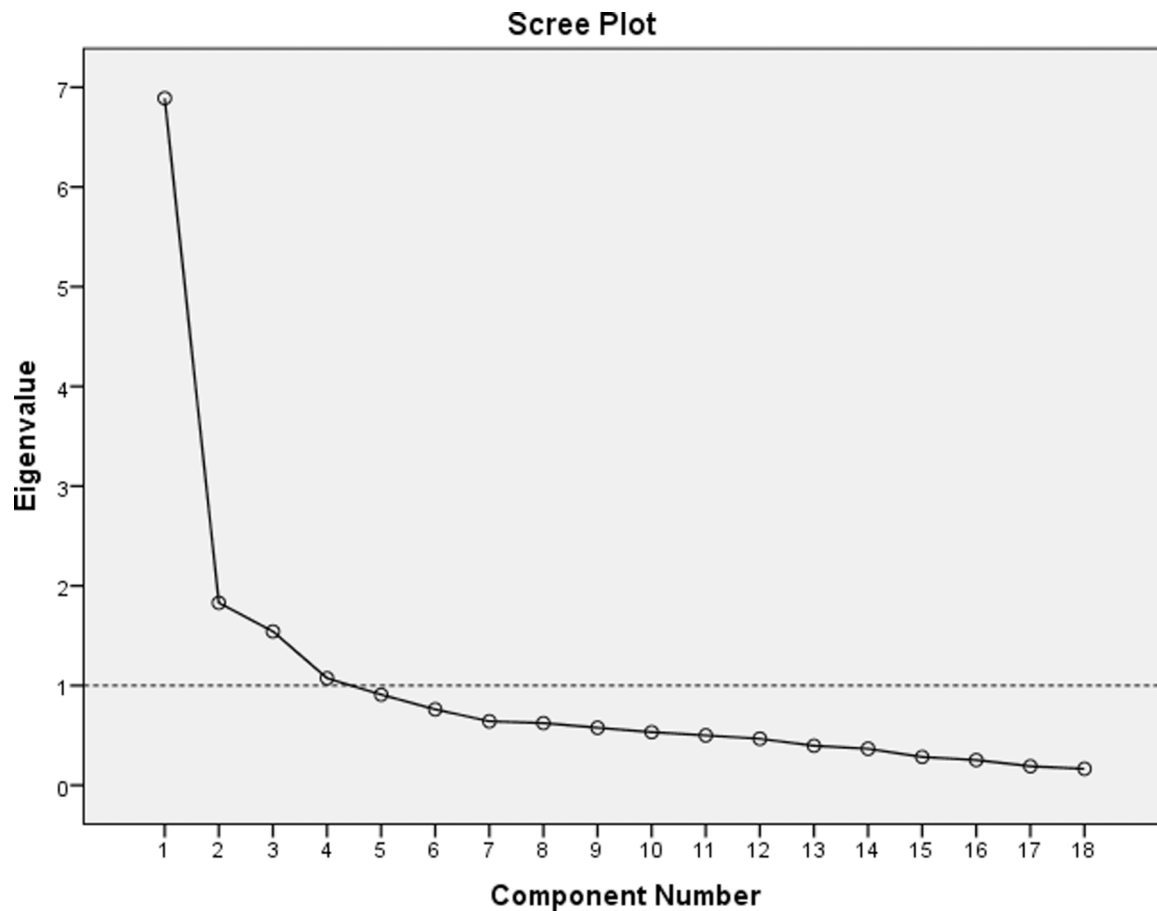


Fig. 1. Scree plot of WHHQ-18.

Table 3
Convergent validity using three comparator measures.

	Correlation with WHHQ 18 (n = 105)	2 tailed significancetest
WEMWBS (n = 96)	0.733	p < 0.01
HEHIQ (n = 105)	0.811	p < 0.01
SF-12 v2 (n = 105)		
Physical Component Summary	0.082	p = 0.408
Mental Component Summary	0.080	p = 0.418

Table 4
Distribution of mean WHHQ-18 score at baseline and after.

	N	Mean	SD	95th CI mean	Min	Max
WHHQ-18 baseline	146	44.71	9.80	43.11 – 46.32	22	69
WHHQ-18 after therapy	146	49.02	9.13	47.53 – 50.51	26	75

WEMWBS, and WEMWBS has been shown to have moderate correlation (.45 -.65) with well validated measures of mental illness in diverse populations [17,36–38]. Items in the WHHQ-18 which might be anticipated to correlate negatively with the SF-12 MH subscale include experience of joy, calm and life satisfaction, as well as possibly sleeping well and feeling in control. The lack of correlation we observed is likely to be due to both the broader range of health and wellbeing states captured by the WHHQ-18 and to ceiling effects in SF-12 such that the component scales do not register improvements in wellbeing captured by the WHHQ-18. Further investigation of this finding is warranted.

Clinicians and researchers alike need to ensure they are measuring

outcomes important to clients receiving TCIM and to capture the full range of possible change. The WHHQ-18 covers common symptoms of illness as well as a wide range of aspects of wellbeing so provides a broader perspective on outcomes and effectiveness of treatment. However, findings with regard to the SF-12 PCS and MCS suggest that TCIM researchers would be advised to include established measures of physical and mental illness alongside the WHHQ-18. These could be either generic measures like the MCS and PCS or disease specific measures.

Weaknesses of this study include the relatively small sample size and the fact that it was drawn primarily from clients of one TCIM who were keen to support research. Clients providing data for the conceptual framework and early versions of the scale were primarily female reflecting the demographic of CST clients. Collection of WHHQ-18 data on a larger sample of clients being treated in different TCIM approaches would enable confirmatory factor analysis as well investigation of face validity and responsiveness in different TCIMs. It would also enable investigation of the lack of correlation between the SF-12 physical and mental subscales and the WHHQ-18. Undertaking a comparison with MYMOP [39] and establishing the relative performance of both scales would be valuable.

The strengths of this work lie in the involvement of CST practitioners and clients throughout the process of development. The WHHQ-18 is the first PROM developed and validated to capture outcomes for CST and will enable evaluation of this and other under-researched TCIMs. The conceptual framework on which it is based has instigated new areas for investigation in terms of treatment effects so the WHHQ-18 can be used as a teaching tool. New areas include the development of self-awareness and the importance of individuals taking responsibility for their own health. Such outcomes are potentially of relevance across the spectrum of health care not just in TCIM, so the WHHQ-18 may in due course be

useful in other settings particularly if the understanding implicit in the WHHQ-18 that mental, physical, social/relational and spiritual well-being are all equally important for health begins to spread to more biomedical approaches to health care, for example palliative care settings.

5. Conclusion

This study responds to the lack of suitable PROMS in the field of Traditional, Complementary and Integrative Medicine to evaluate services being delivered within primary and secondary care settings. The WHHQ-18 is psychometrically sound and has been demonstrated to have good face and content validity by both clients and practitioners of CST. Pending further investigation of construct validity in regard to HRQoL measures, consideration needs to be given to using the WHHQ-18 alongside HRQoL and disease specific measures for research purposes. In due course the WHHQ-18 may have wider application in TCIM private practice, primary care and voluntary sector settings. Further research is needed to assess how the WHHQ-18 performs in these settings.

Financial Support

Nicola Brough was supported by a University of Warwick Chancellor's Scholarship Award. The Craniosacral Therapy Association UK funded the research costs.

Declaration of Competing Interests

Dr Nicola Brough is a member of the CSTA and may attract more clients because of publishing this paper. NB, SSB and HP are named as inventors of the WHHQ which may in the future be made available to other practitioners and researchers through a paid license agreement, www.go.warwick.ac.uk/whhq.

CRedit authorship contribution statement

Nicola Brough: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Validation, Visualization, Writing – original draft, Writing – review & editing. **Sarah Stewart-Brown:** Conceptualization, Supervision, Visualization, Writing – original draft, Writing – review & editing. **Helen Parsons:** Methodology, Supervision, Visualization, Writing – review & editing. **Chinthana Perera:** Data curation, Formal analysis, Validation, Writing – review & editing.

Acknowledgements

Dr Nicola Brough undertook this study for the degree of Doctor in Philosophy in Health Sciences, University of Warwick, Warwick Medical School. Repeatability was assessed by Simon Saldhana as part of an MBChB student project, University of Warwick, Warwick Medical School.

Data Availability Statement

The dataset(s) supporting the conclusions of this article is (are) available from the corresponding author by request.

References

- [1] MJ Verhoef, LC Vanerheyden, T Dryden, D Mallory, MA Ware, Evaluating complementary and alternative medicine interventions: in search of appropriate patient-centered outcome measures, *BMC Complement. Alternat. Med.* 6 (38) (2006).
- [2] C Ritenbaugh, M Nichter, KL Kelsy, CM Sims, IR Bell, HM Castaneda, et al., Developing a patient-centered outcome measure for complementary and alternative medicine therapies I: defining content and format, *BMC Complement. Alternat. Med.* 11 (135) (2011) 17.
- [3] C Paterson, C Baarts, L Launso, MJ Verhoef, Evaluating complex health interventions: a critical analysis of the 'outcomes' concept, *BMC Complement. Alternat. Med.* 9 (18) (2009).
- [4] EA Thompson, T Quinn, C Paterson, H Cooke, D McQuigan, G. Butters, Outcome measures for holistic, complex interventions within the palliative care setting, *Complement. Ther. Clin. Pract.* 14 (2008) 25–32.
- [5] H J, Exploring the prospect of a complementary and integrative medicine database for use in the Australian primary care setting, *Adv. Integrat. Med.* 1 (1) (2014) 25–31.
- [6] WG. Sutherland, AL Wales, Teachings in the Science of Osteopathy, Sutherland Cranial Teaching Foundation, Inc, Fort Worth, Texas, 1990.
- [7] N. Brough, *Craniosacral Therapy: Exploring Clients' Experiences. A Qualitative Study*, University of Warwick, Warwickshire, UK, 2012.
- [8] A Biggs, N Brough, C Cremer, D Hatton, R James, L Kalinowska, et al., A Survey of craniosacral therapy (CST) practitioners. Profiling practice and informing organisational strategy, *Eur. J. Integrat. Med.* (2015) 684.
- [9] N Brough, A Lindenmeyer, J Thistlethwaite, G Lewith, S Stewart-Brown, Perspectives on the effect and mechanisms of craniosacral therapy: a qualitative study of users' views, *Eur. J. Integrat. Med.* 7 (2015) 172–183.
- [10] C. Paterson, Measuring changes in self-concept: a qualitative evaluation of outcome questionnaires in people having acupuncture for their chronic health problems, *BMC Complement. Alternat. Med.* 6 (2006) 7.
- [11] Patient-Reported Outcome Measures: Use in, Medical Product Development to Support Labeling Claims, 2009.
- [12] L Morkink, C Terwee, D Patrick, J Alonso, PW Stratford, DK Knol, et al., The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study, *Qual. Life Res.* 19 (4) (2010) 539–549.
- [13] MJ Polley, HE Seers, HJ Cooke, C Hoffman, C. Paterson, How to summarise and report written qualitative data from patients: a method for use in cancer support care, *Support Care Cancer* 15 (2007) 963–971.
- [14] T Cartwright, R. Torr, Making sense of illness: the experience of users of complementary medicine, *J. Health Psychol.* 10 (4) (2005) 559–572.
- [15] N. Brough, The Warwick Holistic Health Questionnaire - The Development and Validation of a Patient-Reported Outcome Measure for Craniosacral Therapy: A mixed methods study, University of Warwick, Warwick, 2017.
- [16] N. Brough, Warwick Holistic Health Questionnaire United Kingdom, 2019 [updated October 2020. Available from: go.warwick.ac.uk/whhq.
- [17] JE Ware, M Kosinski, SD. Keller, SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales, 2nd ed., The Health Institute, Boston, MA, 1995.
- [18] JE Ware, C.D. Sherbourne, The MOS 36-item short-form health survey (SF-36), *Med. Care* 30 (1992) 473–483.
- [19] R Tennant, L Hiller, R Fishwick, S Platt, S Joseph, S Weich, et al., The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): development and UK validation, *Health Qual. Life Outc.* 5 (1) (2007) 24.
- [20] FL Bishop, F Barlow, J Walker, C McDermott, GT. Lewith, The development and validation of an outcome measure for spiritual healing: a mixed method study, *Psychosom. Psychosom.* 253 (2010) 1–13.
- [21] Q XM, . Qualtrix Online Survey Software UK: SAP, 2018.
- [22] Brown TA. *Confirmatory Factor Analysis for Applied Research*. KennyDA, New York: The Guilford press; 2006.
- [23] R Ho, *Handbook of Univariate and Multivariate Data Analysis and Interpretation with SPSS*, Tailor & Francis Group, USA, 2006.
- [24] R Jayadevappa, SB Malkowicz, M Wittink, AJ Wein, S. Chhatre, Comparison of distribution- and anchor-based approaches to infer changes in health-related quality of life of prostate cancer survivors, *Health Serv. Res.* 47 (5) (2012) 1902–1925.
- [25] RD Hays, D. Hadorn, Responsiveness to change: an aspect of validity, not a separate dimension, *Qual. Life Res.* 1 (1) (1992) 73–75.
- [26] B. Middle, EV. Sonderen, Statistical significant change versus relevant or important change in (quasi) experimental design: some conceptual and methodological problems in estimating magnitude of intervention-related change in health services research, *Int. J. Integr. Care* 2 (2002).
- [27] GY. Zou, Quantifying responsiveness of quality of life measures without an external criterion, *Qual. Life Res.* 14 (2005) 1545–1552.
- [28] GR Norman, JA Sloan, KW. Wyrwich, Interpretation of changes in health-related quality of life the remarkable universality of half a standard deviation, *Med. Care* 41 (5) (2003) 582–592.
- [29] DE Beaton, C Bombardier, JNK JN, JG Wright, A taxonomy for responsiveness, *J. Clin. Epidemiol.* 54 (2001) 1204–1217.
- [30] RA. Likert, A technique for the development of attitude scales, *Educ. Psychol. Meas.* 12 (1952) 313–315.
- [31] RG Netemeyer, WO Bearden, S. Sharma, *Scaling Procedures: Issues and Applications*, Sage Publications, 2003.
- [32] JF Hair, WC Black, BJ Babin, RE. Anderson, *Multivariate Data Analysis*, Pearson Education Limited, 2014.
- [33] N Brough, A Lindenmeyer, J Thistlethwaite, G Lewith, Perspectives on the effects and mechanisms of craniosacral therapy: a qualitative study of users' views, in: S Stewart-Brown (Ed.), *Perspectives on the effects and mechanisms of craniosacral therapy: a qualitative study of users' views*, *Eur. J. Integrat. Med.* 7 (2) (2015) 172–183.
- [34] C. Paterson, Measuring outcomes in primary care: a patient generated measure, MYMOP, compared with the SF-36 health survey, *BMJ* 312 (7037) (1996) 1016–1020.

- [35] D Eton, K Koffler, D Cella, A Eisenstein, J Astin, K Pelletier, et al., Developing a self-report outcome measure for complementary and alternative medicine, *Explore* 1 (3) (2005) 177–185.
- [36] A Clarke, T Friede, R Putz, J Ashdown, S Martin, A Blake, et al., Warwick-Edinburgh Mental Well-being Scale (WEMWBS): validated for teenage school students in England and Scotland. A mixed methods assessment, *BMC Public Health* (2011) 11.
- [37] S Stewart-Brown, A Tennant, R Tennant, S Platt, J Parkinson, S. Weich, Internal construct validity of the Warwick-Edinburgh Mental Well-being Scale (WEMWBS): a Rasch analysis using data from the Scottish Health Education Population Survey, *Health Qual. Life Outc.* 7 (2009) 15.
- [38] F Taggart, T Friede, S Weich, A Clarke, M Johnson, S Stewart-Brown, Cultural evaluation of the Warwick-Edinburgh Mental Well-Being Scale (WEMWBS) a mixed methods study, *Health Qual. Life Outc.* 11 (2013) 27.
- [39] C. Paterson, Measuring outcome in primary care: a patient-generated measure, MYMOP, compared to the SF-36 health survey, *BMJ* 312 (1996) 1016–1020.