Qual Life Res
DOI 10.1007/s11136-016-1406-3

The utility of patient-reported outcome measures among patients with myalgic encephalomyelitis/chronic fatigue syndrome

Kyle W. Murdock[1] · Xin Shelley Wang[2] · Qiuling Shi[2] · Charles S. Cleeland[2] · Christopher P. Fagundes[1,2] · Suzanne D. Vernon[3]

Accepted: 31 August 2016

Abstract
Purpose Debilitating fatigue is a core symptom of myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS); however, the utility of patient-reported symptom outcome measures of fatigue for ME/CFS patients is problematic due to ceiling effects and issues with reliability and validity. We sought to evaluate the performance of three patient-reported symptom measures in a sample of ME/CFS patients and matched controls.
Methods Two hundred and forty ME/CFS patients and 88 age, sex, race, and zip code matched controls participated in the study. Participants completed the Multidimensional Fatigue Inventory-20, DePaul Symptom Questionnaire, and RAND SF-36.
Results The general and physical fatigue subscales on Multidimensional Fatigue Inventory-20, as well as the role of physical health on the RAND SF-36, demonstrated questionable or unacceptable internal consistency and problematic ceiling effects. The DePaul Symptom Questionnaire demonstrated excellent internal reliability, and less than 5 % of participants were at the ceiling on each subscale. The post-exertional malaise subscale on the DePaul Symptom Questionnaire demonstrated excellent clinical utility as it was able to differentiate between ME/CFS patients and controls (OR 1.23, $p < .001$) and predicted ceiling effects on other patient-reported outcome subscales. A score of 20 on the post-exertional malaise subscale of the DePaul Symptom Questionnaire optimally differentiated between patients and controls.
Conclusions Significant ceiling effects and concerns with reliability and validity were observed among Multidimensional Fatigue Inventory-20 and RAND SF-36 subscales for ME/CFS patients. The DePaul Symptom Questionnaire addresses a number of concerns typically identified when using patient-reported outcome measures with ME/CFS patients; however, an improved multidimensional patient-reported outcome tool for measuring ME/CFS-related symptoms is warranted.

Keywords Chronic fatigue syndrome · Myalgic encephalomyelitis · Fatigue · Symptom measurement · Sensitivity

Introduction

Individuals diagnosed with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), a term which was introduced in 1980s and redefined in 1994, exhibit impaired functioning, health status, and quality of life that contributes to significant physical, emotional and economic burden due to lost productivity [1, 2]. Approximately 1 % of the general population meets criteria for ME/CFS [3]. Although diagnostic criteria for ME/CFS have been widely discussed [4, 5], the only mandatory feature of ME/CFS is a period of at least 6 months of debilitating fatigue, which can be accompanied by a variety of other symptoms such as post-exertional malaise, unrefreshing sleep, cognitive impairment, orthostatic intolerance, and a list of a number

✉ Xin Shelley Wang
xswang@mdanderson.org

[1] Department of Psychology, Rice University, Houston, TX, USA

[2] Department of Symptom Research, The University of Texas MD Anderson Cancer Center, 1515 Holcombe Blvd., Unit 1450, Houston, TX 77030, USA

[3] Bateman Horne Center of Excellence, Salt Lake City, UT, USA

Published online: 06 September 2016

of other symptoms [5–11]. Despite much debate regarding the diagnostic criteria for ME/CFS, the quality of patient symptom reports has not been well studied. Such research is urgently needed for improving symptom assessment accuracy, which is important for symptom management and intervention trials for ME/CFS.

Many of the patient-reported outcome measures that are typically utilized in clinical studies of ME/CFS have demonstrated problematic utility. Identified problems among patient-reported outcome measures utilized to study ME/CFS include low test–retest reliability, questionable internal consistency, unknown construct validity, and frequent ceiling effects [12]. Moreover, patient-reported outcome measures of ME/CFS-related symptoms are unable to differentiate among those who do and do not have ME/CFS [13]. Collectively, poor performance of the patient-reported outcome measures that have been utilized, which were not originally constructed to measure symptoms of ME/CFS, has likely contributed to the disparate findings identified in factor analytic studies of ME/CFS symptoms [14–20]. Indeed, the poor performance of patient-reported outcome measures utilized among ME/CFS studies has led some to conclude that patient-reported fatigue as a symptom in ME/CFS may not be measurable [12, 21].

Ceiling effects on patient-reported outcome measures are particularly problematic as they have been identified across many patient populations [22–25]. Ceiling effects, defined as a proportion of patients reporting maximum scores on a given measure more than 40 % of the time [23, 25], are particularly concerning for evaluating response to intervention as the range of the patient-reported outcome measure may be too restrictive to capture clinically significant change. That is, a patient scoring at the maximum of a given measure at baseline may continue to score at the maximum after receiving treatment even if clinically significant improvement has occurred. As a result, intervention trials may underestimate ME/CFS patient response to treatment when using patient-reported outcome measures as a result of ceiling effects, which may contribute to findings indicating many ME/CFS patients do not respond to treatment [26, 27]. An accurate understanding of patient response to treatment is needed in order to move research and patient care forward [28].

The vast majority of patient-reported outcome measures targeting fatigue were not designed to measure post-exertional malaise, a hallmark of ME/CFS [20]. The DePaul Symptom Questionnaire [29] was designed to evaluate symptoms of ME/CFS, including post-exertional malaise and has demonstrated excellent convergent and discriminant validity [19], as well as test–retest reliability [10]. However, the performance of the DePaul Symptom Questionnaire in comparison with other patient-reported outcome measures that have been previously used in ME/CFS research is largely unknown. For instance, ceiling effects and the ability to differentiate ME/CFS patients from other populations have yet to be reported for the DePaul Symptom Questionnaire. Further evaluation of patient-reported outcome measures is necessary for improving diagnostic accuracy through identification of measurement strategies that are able to discriminate between ME/CFS patients and other populations. Patient care can also be improved through identification of measures that are able to capture clinically significant change among all patients being treated for ME/CFS due to a lack of ceiling effects. Indeed, methodological concerns such as measurement of a variety of different symptoms were noted in a review of the effectiveness of interventions for treating ME/CFS symptoms [30]. Therefore, improved measurement quality and consistency in measurement of symptoms across studies are clearly needed.

The current study aimed to examine the performance of patient-reported outcome measures (Multidimensional Fatigue Inventory-20, RAND SF-36 (Version 2) and DePaul Symptom Questionnaire) for measuring fatigue and other critical symptoms of ME/CFS in a sample of ME/CFS patients and age, sex, race, and zip code matched controls. In particular, ceiling effects, internal consistency, sensitivity, and specificity were evaluated. Consistent with the available literature, we hypothesized that significant ceiling effects would be observed among the ME/CFS sample. Furthermore, we hypothesized that DePaul Symptom Questionnaire subscales would significantly differentiate ME/CFS patients from matched controls due to the measure being designed to measure symptoms specific to ME/CFS (i.e., post-exertional malaise).

Methods

Participants and procedure

Data were obtained from the Solve ME/CFS Initiative (formerly known as the CFIDS Association of America) following Institutional Review Board approval. The original study was designed to evaluate the role that methodological approaches had in the previously identified, and now largely rejected, association between xenotropic murine leukemia virus and ME/CFS [31]. The data were collected between June and August 2010. ME/CFS participants were eligible for inclusion in the study if they had been previously diagnosed with ME/CFS using the Fukuda [3] or Canadian criteria [11]. Participants with ME/CFS also had to have post-exertional malaise lasting >24 h and significant cognitive impairment in short-term memory and concentration. The final sample included 240 ME/CFS patients who were recruited from four clinical sites as

previously report by Irlbeck et al. [31]. A total of 88 healthy individuals were recruited for the control group, who were matched to ME/CFS patients by zip code (excluded if they lived within the same household), age, sex, and race.

All questionnaires described below were self-reported by participants. Participants were required to be capable of giving written informed consent and be between 18 and 65 years of age in order to be eligible for the study. The mean time taken to complete the measures was 15 min in total.

Measures

Multidimensional Fatigue Inventory-20

The Multidimensional Fatigue Inventory-20 comprises five subscales: general fatigue, physical fatigue, mental fatigue, reduced activity, and reduced motivation [32]. Each subscale includes four items with five-point scales ranging from 1 (Yes, that is true) to 5 (No, that is not true). Scores on each subscale range from 4 to 20, with higher scores indicating greater fatigue. The Multidimensional Fatigue Inventory-20 is a reliable and valid measure of fatigue in many patient populations [23]; however, the Multidimensional Fatigue Inventory-20 has demonstrated problems with sensitivity and specificity when utilized with ME/CFS patients [13].

RAND SF-36 (version 2)

The RAND SF-36 contains eight multi-item subscales: physical functioning, role limitations due to physical health, role limitations due to emotional problems, energy/fatigue, pain, general well-being, social functioning, and general health [33]. The number of response choices per item ranges from two to six. Although a widely utilized measure evidencing strong psychometric characteristics [33], the RAND SF-36 has demonstrated problems measuring substantial reductions in functioning among ME/CFS patients [13]. Scores were reverse coded such that higher scores indicate more severe impairment in functioning.

DePaul Symptom Questionnaire

The DePaul Symptom Questionnaire is a self-report measure of symptomatology, demographics, and medical, occupational, and social history [29]. In regard to symptoms, participants were asked to rate the frequency and severity of each of 54 symptoms over the past 6 months on a five-point scale. The DePaul Symptom Questionnaire has evidenced good test–retest reliability among both patient and control groups [10]. Frequency and symptoms severity scores are combined to form an overall indicator for each symptom. Four factors (i.e., autonomic/neuroendocrine/immune, cognitive, post-exertional malaise, and sleep) have been identified as core domains of ME/CFS using 40 of the symptoms measured on the DePaul Symptom Questionnaire [20]. The four factors identified by Jason et al. [10] were utilized in the analyses described below. Higher scores on the DePaul Symptom Questionnaire indicate more severe impairment in functioning.

Statistical analysis

We utilized SPSS statistical software [34] for all data analyses. Descriptive statistics were generated to characterize the study samples in terms of sociodemographic parameters. Moreover, Cronbach's alpha coefficients [35] were calculated for each subscale on study measures. Means and standard deviations, as well as one-way ANOVAs comparing the ME/CFS and control samples, were calculated for all study symptom subscales. Variables indicating whether or not participants were at the ceiling of each subscale on study measures were also calculated. Logistic regression analyses were generated to examine the sensitivity and specificity of each subscale in predicting whether or not participants were in the ME/CFS or control sample. The standard cut-point of 0.5 was utilized in all logistic regression analyses [36]. Furthermore, logistic regression analyses were utilized to determine whether or not DePaul Symptom Questionnaire subscales, as well as the role limitations due to physical functioning subscale on the RAND SF-36, were able to predict if participants would be at the ceiling on other Multidimensional Fatigue Inventory-20 and RAND SF-36 subscales. A receiver operating characteristics curve was calculated in ancillary analyses to determine an optimal cut-point for differentiating between ME/CFS and control patients using the post-exertional malaise subscale of the DePaul Symptom Questionnaire.

Results

Sample characteristics and internal consistency

As shown in Table 1, ME/CFS patients were less likely to be working [$F(1,315) = 161.39$, $p < .001$] and more likely to be disabled [$F(1,315) = 173.37$, $p < .001$] than individuals in the control group. No other group differences were identified on demographic variables.

Internal consistency was questionable for the reduced motivation subscale ($\alpha = .69$) on the Multidimensional Fatigue Inventory-20 for the ME/CFS sample. For the

Table 1 Demographic characteristics of the study sample

Variable	ME/CFS patients ($n = 240$) Mean (SD) or number (%)	Controls ($n = 88$) Mean (SD) or number (%)
Age, years, mean (SD)	49.72 (12.89)	49.60 (13.30)
Sex		
Female	174 (73)	70 (80)
Male	65 (27)	18 (20)
Race (%)		
Black or AA	0 (0)	1 (1)
White	235 (98)	85 (97)
All others	5 (2)	2 (2)
Educational status (%)		
High school graduate/GED	27 (11)	11 (13)
Some college	49 (20)	22 (25)
College graduate	101 (42)	35 (40)
Graduate degree	37 (15)	14 (16)
Professional degree	23 (10)	5 (6)
Missing/prefer not to answer	3 (2)	0 (0)
Marital status (%)		
Married/cohabitating	137 (57)	57 (64)
Widowed	4 (2)	5 (6)
Divorced	35 (14)	11 (13)
Separated	1 (1)	1 (1)
Missing/prefer not to answer	63 (26)	14 (16)
Employment status (%)		
Working	26 (11)*	59 (67)*
Not employed	3 (1)	3 (3)
Retired	23 (10)	11 (13)
Disabled	156 (65)*	0 (0)*
Homemaker	13 (5)	7 (8)
Student	11 (5)	5 (6)
Missing/prefer not to answer	8 (3)	3 (3)

* $p < .001$

control group, internal consistency on the physical fatigue scale was in the questionable range on the Multidimensional Fatigue Inventory-20 ($\alpha = .69$). All other scales were within the acceptable to good range on the Multidimensional Fatigue Inventory-20 ($\alpha = .72$–$.81$) for ME/CFS and control participants. On the RAND SF-36, internal consistency was unacceptable for the general health subscale ($\alpha = .56$), and the energy/fatigue subscale was within the questionable range ($\alpha = .68$), for the ME/CFS sample. In the control sample, two RAND SF-36 subscales were within the unacceptable range (i.e., role limitations due to emotional problems ($\alpha = .48$) and emotional well-being ($\alpha = .46$), while the general health subscale was within the questionable range ($\alpha = .66$). Other RAND SF-36 subscales were within the acceptable to excellent range ($\alpha = .71$–$.91$) for ME/CFS and control participants. All subscales on the DePaul Symptom Questionnaire were within the good to excellent range for internal consistency ($\alpha = .89$–$.96$) for both groups.

Distribution of patient-reported outcome measures

Means and standard deviations for study measure subscales are presented in Table 2. As expected, significant mean differences were identified for each of the subscales measured when comparing the ME/CFS and control samples. Furthermore, ceiling effects (i.e., maximum scores on a given measure among more than 40 % of patients; [23, 25]) were identified for ME/CFS sample participants on the general fatigue subscale of the Multidimensional Fatigue Inventory-20, while the physical fatigue subscale nearly demonstrated ceiling effects with 38 % of ME/CFS patients at the ceiling. The role limitations due to physical

Table 2 Means, standard deviations, and ceiling effects for subscales among the ME/CFS and control samples

Measure	ME/CFS patients ($n = 240$)	Controls ($n = 88$)	Number (%) of ME/CFS patients at ceiling	Number (%) of controls at ceiling
Multidimensional Fatigue Inventory-20				
General fatigue	17.83 (2.81)	8.42 (3.23)	95 (40)	0 (0)
Physical fatigue	17.44 (3.10)	6.67 (2.73)	90 (38)	0 (0)
Mental fatigue	14.16 (3.59)	8.02 (3.62)	17 (7)	0 (0)
Reduced activity	15.97 (3.57)	6.64 (3.06)	48 (20)	1 (1)
Reduced motivation	12.25 (3.82)	6.65 (3.13)	3 (1)	1 (1)
RAND SF-36				
Physical functioning	63.41 (23.59)	5.23 (9.62)	14 (6)	0 (0)
Role limitations due to physical health	94.90 (16.73)	5.23 (17.20)	214 (89)	1 (1)
Role limitations due to emotional problems	29.55 (41.76)	6.98 (17.80)	51 (21)	1 (1)
Energy/fatigue	82.99 (16.94)	30.30 (15.60)	54 (23)	0 (0)
Emotional well-being	30.83 (19.47)	19.16 (11.61)	0 (0)	0 (0)
Social functioning	71.35 (25.00)	6.47 (12.33)	54 (23)	0 (0)
Pain	57.40 (28.50)	11.63 (14.19)	33 (14)	0 (0)
General health	75.18 (17.98)	17.42 (14.95)	22 (9)	0 (0)
DePaul Symptom Questionnaire				
Autonomic/neuroendocrine/immune	60.87 (31.35)	11.83 (10.66)	0 (0)	0 (0)
Cognitive	37.71 (15.99)	6.23 (7.14)	3 (1)	0 (0)
Post-exertional malaise	45.06 (12.55)	6.05 (7.26)	7 (3)	0 (0)
Sleep	12.53 (6.11)	5.39 (4.81)	11 (5)	0 (0)

All ANOVA comparisons of group means were significant at $p < .001$

health subscale on the RAND SF-36 were particularly concerning given that 89 % of the ME/CFS sample reported the highest score possible. Less than 5 % of the sample reported symptoms at the ceiling of on any of the subscales of the DePaul Symptom Questionnaire.

Differentiating ME/CFS patients and control participants

To determine if measure subscales were able to effectively differentiate between ME/CFS and control participants, a series of logistic regression analyses were conducted. Separate analyses were run for each measure (see Table 3). For the Multidimensional Fatigue Inventory-20, the physical fatigue subscale significantly predicted group membership. In regard to the RAND SF-36, the role limitations due to physical health and general health subscales were significant predictors of whether or not participants were in the ME/CFS and controls samples. For the DePaul Symptom Questionnaire, the cognitive, post-exertional malaise, and sleep subscales significantly differentiated between the ME/CFS and control participants.

Further analysis of ceiling effects

As DePaul Symptom Questionnaire subscales did not exhibit significant ceiling effects and were able to differentiate between ME/CFS and control samples, we ran a series of regression analyses to evaluate whether DePaul Symptom Questionnaire subscales could differentiate among those who were at the ceiling on Multidimensional Fatigue Inventory-20 and RAND SF-36 subscales from those who were not (see Table 4). The autonomic/neuroendocrine/immune subscale of the DePaul Symptom Questionnaire significantly predicted those who were at the ceiling for the Multidimensional Fatigue Inventory-20 reduced activity subscale, as well as the physical functioning and pain subscales of the RAND SF-36, from those who did not. The cognitive scale of the DePaul Symptom Questionnaire differentiated among those who were and were not at the ceiling for the mental fatigue subscale on the Multidimensional Fatigue Inventory-20, as well as the physical functioning, and role limitations due to emotional health subscales on the RAND SF-36. The post-exertional malaise subscale predicted whether or not participants were

Table 3 Logistic regressions for measure subscales predicting ME/CFS and control sample

Measure	Odds ratio	95 % CI	*p*
Multidimensional Fatigue Inventory-20			
General fatigue	1.15	.93–1.43	0.2
Physical fatigue	1.65	1.29–2.11	<.001
Mental fatigue	0.99	.83–1.12	0.94
Reduced activity	1.06	.82–1.39	0.66
Reduced motivation	0.94	.76–1.16	0.55
RAND SF-36			
Physical functioning	1.02	.96–1.09	0.49
Role limitations due to physical health	1.07	1.03–1.11	0.001
Role limitations due to emotional problems	0.99	.96–1.02	0.6
Energy/fatigue	0.99	.93–1.06	0.77
Emotional well-being	0.98	.91–1.05	0.5
Social functioning	1.01	.94–1.09	0.78
Pain	0.97	.92–1.02	0.21
General health	1.1	1.03–1.18	0.006
DePaul Symptom Questionnaire			
Autonomic/neuroendocrine/immune	0.98	.93–1.04	0.5
Cognitive	1.16	1.06–1.27	0.002
Post-exertional malaise	1.23	1.12–1.34	<.001
Sleep	0.84	.73–.97	0.02

Table 4 Logistic regressions for DePaul Symptom Questionnaire subscales predicting ceiling effects among Multidimensional Fatigue Inventory-20 and RAND SF-36 subscales

Measure	Autonomic/ neuroendocrine/ immune		Cognitive		Post-exertional malaise		Sleep	
	Odds ratio	95 % CI	Odds ratio	95 % CI	Odds ratio	95 % CI	Odds ratio	95 % CI
Multidimensional Fatigue Inventory-20								
General fatigue	1.00	.99–1.02	1.00	.98–1.03	1.01***	1.06–1.13	1.01	.96–1.07
Physical fatigue	.99	.98–1.00	.98	.96–1.01	1.12***	1.08–1.16	1.02	.96–1.08
Mental fatigue	1.00	.98–1.03	1.12***	1.05–1.20	1.01	.94–1.08	.95	.84–1.07
Reduced activity	1.02*	1.00–1.03	.98	.95–1.01	1.09***	1.05–1.14	.94	.88–1.01
RAND SF-36								
Physical functioning	1.06**	1.02–1.11	.93*	.87–.99	1.16*	1.02–1.31	.99	.86–1.13
Role limitations due to physical health	.98	.96–1.00	1.04	1.01–1.08	1.14***	1.09–1.19	.98	.91–1.05
Role limitations due to emotional health	.99	.98–1.01	1.04**	1.01–1.07	1.03	1.00–1.07	1.07	1.00–1.14
Energy/fatigue	1.00	.99–1.02	1.00	.97–1.03	1.20***	1.13–1.28	.91*	.84–.99
Social functioning	1.00	.99–1.02	1.00	.97–1.03	1.12***	1.12–1.27	.95	.88–1.02
Pain	1.03**	1.01–1.05	.97	.93–1.00	1.09**	1.03–1.15	.99	.91–1.07
General health	1.00	.98–1.02	1.03	.99–1.07	1.05	.99–1.10	1.04	.95–1.13

* $p < .05$; ** $p < .01$; *** $p < .001$

at the ceiling for the majority of subscales measured. Indeed, the post-exertional malaise scale significantly differentiated those who were at the ceiling from those who were not for the general fatigue, physical fatigue, and reduced activity subscales of the Multidimensional Fatigue Inventory-20, as well as the physical functioning, role limitations due to physical health, energy/fatigue, social functioning, and pain subscales of the RAND SF-36. The sleep subscale on the DePaul Symptom Questionnaire significantly predicted whether or not participants were at

the ceiling on the energy/fatigue subscale on the RAND SF-36.

The utility of the role limitations due to physical health subscale on the RAND SF-36 in predicting ceiling effects on the Multidimensional Fatigue Inventory-20 was also evaluated. The role limitations due to physical health subscale of the RAND SF-36 were unable to differentiate between patients at and below the ceiling on the general fatigue (OR 1.00, $p = .79$), physical fatigue (OR 1.00, $p = .73$), mental fatigue (OR 1.01, $p = .75$), reduced activity (OR 1.00, $p = .90$), and reduced motivation (OR .95, $p = .11$) subscales on the Multidimensional Fatigue Inventory-20.

Sensitivity and specificity of DePaul Symptom Questionnaire

Given the superior performance of the post-exertional malaise subscale on the DePaul Symptom Questionnaire, we conducted a receiver operating characteristics curve analysis to determine an ideal cutoff score to differentiate between ME/CFS patients and controls. The area under the receiver operating characteristics curve value was excellent (.98). Optimal sensitivity (i.e., positive predictive value; .95) and specificity (i.e., negative predictive value; .94) were identified when using a cutoff score of 20. Accordingly, a score of 20 on the post-exertional malaise subscale of the DePaul Symptom Questionnaire could be utilized to differentiate those with ME/CFS from those in the general population.

Discussion

Use of the Multidimensional Fatigue Inventory-20 and RAND SF-36 with ME/CFS patients has been challenged methodologically, and neither measures were not constructed for this disease [12]. This study is the first to identify the critical patient-reported outcome measure scales for differentiating ME/CFS patients and a control sample including physical fatigue subscale from the Multidimensional Fatigue Inventory-20, the role limitations due to physical health subscale of the RAND SF-36, and the post-exertional malaise subscale from the DePaul Symptom Questionnaire. Such findings should be considered for future symptom assessment of ME/CFS.

In addition, our study demonstrated ceiling effects among patient samples and questionable reliability and validity of the Multidimensional Fatigue Inventory-20 and RAND SF-36, indicating questionable utility for being utilized with ME/CFS patients [22–25]. Indeed, for ME/CFS patients, questionable and poor internal consistency were identified on subscales of the Multidimensional Fatigue Inventory-20 (i.e., reduced motivation) and RAND SF-36 (i.e., energy/fatigue and general health), respectively. Furthermore, ceiling effects were observed for the general fatigue subscale of the Multidimensional Fatigue Inventory-20 and the role limitations due to physical health subscale on the RAND SF-36. Such findings warrant research in development of a new measure of fatigue and other symptoms among those with ME/CFS. Focusing on fatigue may not help to distinguish ME/CFS from other conditions that also present with fatigue. However, based on research on cancer-related fatigue [37], we highlight that a reliable and valid symptom measure of fatigue should be a necessary tool in patient care for diagnosed cases of ME/CFS during an intervention.

While the DePaul Symptom Questionnaire has demonstrated good validity in prior work [10], the present study extends the literature by demonstrating reliability and clinical utility of DePaul Symptom Questionnaire subscales with ME/CFS patients. Each subscale on the DePaul Symptom Questionnaire demonstrated good to excellent internal consistency. Ceiling effects were relatively rare (<5 %) on the DePaul Symptom Questionnaire. Further, the cognitive, post-exertional malaise, and sleep subscales of the DePaul Symptom Questionnaire significantly predicted whether or not participants were ME/CFS patients or matched controls, highlighting the utility of the DePaul Symptom Questionnaire as a patient-reported outcome measure tool for ME/CFS patient in clinical research care and research.

Given that post-exertional malaise is the hallmark symptom of CFS, there is no surprise that the post-exertional malaise subscale of the DePaul Symptom Questionnaire was particularly effective at predicting whether or not participants were at the ceiling on Multidimensional Fatigue Inventory-20 and RAND SF-36 subscales. Accordingly, the post-exertional malaise subscale was useful for capturing meaningful differences among ME/CFS patients. As a result, present study findings suggest that the available literature has not yet fully captured the true distribution of symptoms among ME/CFS patients due to limitations of traditionally utilized measures, especially fatigue measures [12]. Perhaps such differences were due to the DePaul Symptom Questionnaire being designed to measure both symptom severity and frequency of post-exertional malaise [29], a core symptom of ME/CFS. Moreover, such findings suggest that traditionally utilized patient-reported outcome measures, such as the Multidimensional Fatigue Inventory-20 and RAND SF-36, are unable to identify clinically significant differences among patients being diagnosed or treated with ME/CFS. Such measurement concerns reduce the ability to accurately assess and revise interventions in a manner that supports best practice in patient care [38], consistent with a prior

literature review [30]. Accordingly, further research is warranted for establishing a comprehensive patient-reported outcome measure symptom tool in ME/CFS patients targeting fatigue severity, fatigability on functioning, as well as post-exertional malaise. Moreover, further research of the DePaul Symptom Questionnaire among other disease groups known to have fatigue (e.g., multiple sclerosis, fibromyalgia) is warranted to address concerns about measurement of fatigue when patients are experiencing post-exertional malaise.

The present study is limited as patient groups that have also demonstrated ceiling effects on patient-reported outcome measures of fatigue (e.g., multiple sclerosis, fibromyalgia) [22] were not included. Additionally, treatment progress as it relates to patient-reported outcome measures was not evaluated in this cross-sectional study. Moreover, the matched control sample was smaller than the ME/CFS patient sample; however, the ability to detect significant differences between each group was unaffected despite sample size differences. The study is also limited given that the timing of the data collection was prior to the establishment of the ME International Consensus Criteria (MEICC) [11]. The MEICC is an evolution of the Fukuda [7] and Canadian [11] criteria that removed the requirement for symptoms needing to be present for at least 6 months and is purported to improve specificity of diagnoses via improved interpretation of symptoms. However, comparisons of the Fukuda, Canadian, and MEICC have not been provided. It is clear that the present study sample experienced symptoms consistent with ME/CFS in comparison with the control sample, and as such, findings provide an improved understanding of ME/CFS symptom measurement. Indeed, authors of the MEICC state that it is important to identify a quantitative score for diagnostic instruments that are measurable and most relevant to the illness [11], and our findings provide a step towards meeting this goal.

Conclusions

The present study highlighted critical scales for measuring symptoms of ME/CFS. Questionable/poor reliability and problematic ceiling effects were identified on Multidimensional Fatigue Inventory-20 and RAND SF-36, which have often been utilized with ME/CFS patients. The DePaul Symptom Questionnaire, a patient-reported outcome measure designed to measure ME/CFS symptoms, demonstrated excellent internal consistency, sensitivity, specificity, and a lack of ceiling effects. The post-exertional malaise subscale of the DePaul Symptom Questionnaire was able to identify clinically meaningful differences among those who were at the ceiling on other patient-reported outcome measure subscales. Therefore, use of the post-exertional malaise subscale of the DePaul Symptom Questionnaire represents best practice for tracking patient response to intervention given the current state of the literature. A cutoff score of 20 on the post-exertional malaise subscale of the DePaul Symptom Questionnaire is optimal for differentiating ME/CFS patients from those in the general population. Multidimensional symptom assessment tools have been established to address patient's perspective for a complicated disease such as cancer [39], and our study supports the need for a patient-reported outcome measure tool to be established that includes components measuring fatigue, fatigability, and post-exertional malaise among ME/CFS patients.

Funding Preparation of the manuscript was supported by a Grant from the National Heart, Lung, and Blood Institute (1R01HL127260-01; 1F32HL131353), and an MD Anderson Cancer Center Support Grant from the National Cancer Institute (P30 CA016672). The funding agency played no role in study design, in the collection, analysis and interpretation of data, in the writing of the report, or in the decision to submit the article for publication.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Reynolds, K. J., et al. (2004). The economic impact of chronic fatigue syndrome. *Cost Effectiveness and Resource Allocation, 2*, 4.
2. US Food and Drug Administration. (2013). *The voice of the patient: Chronic fatigue syndrome and myalgic encephalomyelitis.* http://www.fda.gov/downloads/ForIndustry/UserFees/PrescriptionDrugUserFee/UCM368806.pdf.
3. Fukuda, K., et al. (1994). The chronic fatigue syndrome—A comprehensive approach to its definition and study. *Annals of Internal Medicine, 121*(12), 953–959.
4. Haney, E., et al. (2015). Diagnostic methods for myalgic encephalomyelitis/chronic fatigue syndrome: A systematic review for a National Institutes of Health Pathways to Prevention Workshop. *Annals of Internal Medicine, 162*(12), 834–840.
5. Institute of Medicine of the National Academies (2015). *Beyond myalgic encephalomyelitis/chronic fatigue syndrome: Redefining an illness-report guide for clinicians*. Washington, DC: Institute of Medicine of the National Academies.
6. Holmes, G. P., et al. (1988). Chronic fatigue syndrome: A working case definition. *Annals of Internal Medicine, 108*, 387–389.
7. Fukuda, K., et al. (1994). The chronic fatigue syndrome: A comprehensive approach to its definition and study. *Annals of Internal Medicine, 121*, 953–959.
8. De Becker, P., McGregor, N., & De Meirleir, K. (2001). A definition-based analysis of symptoms in a large cohort of patients with chronic fatigue syndrome. *Journal of Internal Medicine, 250*(3), 234–240.
9. Reeves, W. C., et al. (2003). Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and

recommendations for resolution. *BMC Health Services Research, 3*(1), 1–9.
10. Jason, L. A., et al. (2015). Test–retest reliability of the DePaul Symptom Questionnaire. *Fatigue: Biomedicine, Health & Behavior, 3*(1), 16–32.
11. Carruthers, B. M., et al. (2003). Myalgic encephalomyelitis/chronic fatigue syndrome: Clinical working case definition, diagnostic and treatment protocols (Canadian case definition). *Journal of Chronic Fatigue Syndrome, 11*(1), 7–115.
12. Haywood, K. L., Staniszewska, S., & Chapman, S. (2012). Quality and acceptability of patient-reported outcome measures used in chronic fatigue syndrome/myalgic encephalomyelitis (CFS/ME): A systematic review. *Quality of Life Research, 21*(1), 35–52.
13. Jason, L. A., et al. (2011). Fatigue scales and chronic fatigue syndrome: Issues of sensitivity and specificity. *Disability Studies Quarterly, 31*(1), 1375.
14. Friedberg, F., et al. (2000). Symptom patterns in long-duration chronic fatigue syndrome. *Journal of Psychosomatic Research, 48*(1), 59–68.
15. Jason, L. A., Corradi, K. C., & Torresharding, S. (2007). Toward an empirical case definition of CFS. *Jounral of Social Service Research, 34*(2), 43–54.
16. Nisenbaum, R., et al. (1998). Factor analysis of unexplained severe fatigue and interrelated symptoms—Overlap with criteria for chronic fatigue syndrome. *American Journal of Epidemiology, 148*(1), 72–77.
17. Nisenbaum, R., et al. (2004). Factor analysis of symptoms among subjects with unexplained chronic fatigue—What can we learn about chronic fatigue syndrome? *Journal of Psychosomatic Research, 56*(2), 171–178.
18. Hickie, I. B., et al. (2001). Development of a simple screening tool for common mental disorders in general practice. *The Medical Journal of Australia, 175*, S10-7.
19. Brown, A. A., & Jason, L. A. (2014). Validating a measure of myalgic encephalomyelitis/chronic fatigue syndrome symptomatology. *Fatigue: Biomedicine, Health & Behavior, 2*(3), 132–152.
20. Jason, L. A., et al. (2015). Factor analysis of the DePaul Symptom Questionnaire: Identifying core domains. *Journal of Neurology and Neurobiology, 1*(4), 1–9.
21. Whitehead, L. (2009). The measurement of fatigue in chronic illness: A systematic review of unidimensional and multidimensional fatigue measures. *Journal of Pain and Symptom Management, 37*(1), 107–128.
22. Amtmann, D., et al. (2012). Comparison of the psychometric properties of two fatigue scales in multiple sclerosis. *Rehabilitation Psychology, 57*(2), 159–166.
23. Elbers, R. G., et al. (2012). Reliability and structural validity of the Multidimensional Fatigue Inventory (MFI) in patients with idiopathic Parkinson's disease. *Parkinsonism & Related Disorders, 18*(5), 532–536.
24. Ferentinos, P., et al. (2011). Psychometric evaluation of the Fatigue Severity Scale in patients with major depression. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care & Rehabilitation, 20*(3), 457–465.
25. Stouten, B. (2005). Identification of ambiguities in the 1994 chronic fatigue syndrome research case definition and recommendations for resolution. *BMC Health Services Research, 5*, 37.
26. Knoop, H., et al. (2007). Is a full recovery possible after cognitive behavioural therapy for chronic fatigue syndrome? *Psychotherapy and Psychosomatics, 76*(3), 171–176.
27. White, P. D., et al. (2013). Recovery from chronic fatigue syndrome after treatments given in the PACE trial. *Psychological Medicine, 43*(10), 2227–2235.
28. Deshpande, P. R., et al. (2011). Patient-reported outcomes: A new era in clinical research. *Perspectives in Clinical Research, 2*(4), 137–144.
29. Jason, L. A., et al. (2010). The development of a revised Canadian myalgic encephalomyelitis chronic fatigue syndrome case definition. *American Journal of Biochemistry and Biotechnology, 6*(2), 120–135.
30. Whiting, P., et al. (2001). Interventions for the treatment and management of chronic fatigue syndrome—A systematic review. *Journal of the American Medical Association, 286*(11), 1360–1368.
31. Irlbeck, D. M., Vernon, S., McCleary, K., Bateman, L., Klimas, N. G., Lapp, C. W., et al. (2014). No association found between the detection of either xenotropic murine leukemia virus-related virus or polytropic murine leukiemia virus and chronic fatigue syndrome in a blinded, multi-site, prospective study by the establishment and use of the SolveCFS BioBank. *BMC Research Notes, 4*(7), 461.
32. Smets, E. M. A., et al. (1995). The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *Journal of Psychosomatic Research, 39*(3), 315–325.
33. Ware, J. E., & Sherbourne, C. D. (1992). The Mos 36-item short-form health survey (Sf-36): Conceptual-framework and item selection. *Medical Care, 30*(6), 473–483.
34. IBM (2012) *IBM SPSS statistics for windows, version 21.0*. Armonk, NY: IBM Corp.
35. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334.
36. Cohen, J., et al. (2003). Alternative regression models: logistic, poisson regression, and the generalized linear model. In *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd edn, pp. 479–535). New Jersey: Lawrence Erlbaum Associates in Mawah.
37. Wang, X. S. (2013). Cancer-related fatigue. In A. M. Berger, J. L. Shuster, & J. H. von Roenn (Eds.), *Principles and practice of palliative care and supportive oncology* (pp. 84–95). Philadelphia, PA: Wolters Kluwer/Lippincott Williams & Wilkins.
38. Mokkink, L. B., et al. (2009). Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Quality of Life Research, 18*(3), 313–333.
39. Kirkova, J., et al. (2006). Cancer symptom assessment instruments: A systematic review. *Journal of Clinical Oncology, 24*(9), 1459–1473.