VALUE IN HEALTH 21 (2018) 1313–1321

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jval

Patient-Reported Outcomes

General Population Norms for the Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue Scale

Inka Montan, PhD[1,2,*], *Bernd Löwe, MD, PhD*[2], *David Cella, PhD*[3], *Anja Mehnert, PhD*[4], *Andreas Hinz, PhD*[4]

[1]*Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Heidelberg, Germany;* [2]*Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany;* [3]*Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA;* [4]*Department of Medical Psychology and Medical Sociology, University of Leipzig, Leipzig, Germany*

ABSTRACT

Background: The Functional Assessment of Chronic Illness Therapy (FACIT)-Fatigue Scale is an internationally used validated measure. General population–based age- and sex-specific percentile norms are, however, not published to date, although these are needed as reference for the interpretation of clinical research data. **Objectives:** To assess the FACIT-Fatigue Scale in a large representative sample of the German general population to examine psychometric characteristics and factorial structure and to provide population-based norms. **Methods:** A nationally representative face-to-face household survey was conducted in Germany using the FACIT-Fatigue Scale. Item characteristics were examined. Internal consistency was determined using the Cronbach α. Dimensionality was analyzed using confirmatory factor analysis (CFA) and bifactor analysis. Scale score differences relating to sex and age were assessed. Sex- and age-specific percentiles were computed for the entire scale range. **Results:** Of 2426 participants, 55.7% were women, and the mean age was 49.8 ± 17.4 years. The FACIT-Fatigue Scale mean was 43.5 ± 8.3. Cronbach α was high at 0.92. Although fit indices of the CFA were below desired levels (root mean squared error of approximation = 0.144, comparative fit index = 0.846, and Tucker-Lewis index = 0.815), item loadings in the CFA and bifactor analysis confirm the scale's unidimensionality. Women were more fatigued than men, and participants who were 70 years or older showed higher fatigue scores than younger respondents. Thus, sex- and age-specific population-based percentiles were provided. **Conclusions:** Reliability and validity of the German translation of the FACIT-Fatigue Scale were confirmed. This study provides general population–based sex- and age-specific FACIT-Fatigue Scale percentile norms for the first time, thereby contributing to a meaningful interpretation of clinical research data.

Keywords: FACIT-Fatigue Scale, general population, patient-reported outcomes, percentile norms.

Introduction

Fatigue is among the most prevalent and burdensome symptoms in patients and survivors of cancer [1,2]. The US National Cancer Institute recommends to measure fatigue as part of a core set of symptoms in adult cancer treatment trials [3]. Fatigue measures are classic "patient-reported outcomes" being "reports coming directly from patients about how they feel or function in relation to a health condition and its therapy, without interpretation of the patient's responses by a clinician, or anyone else" [4].

The Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue) Scale [5,6] is a legacy fatigue measure based on classical test theory. The 13-item questionnaire assesses self-reported fatigue and its impact on daily activities and function. It was originally developed as an addition to the Functional Assessment of Cancer Therapy (FACT) measurement system [6]. The 28-item FACT-General (FACT-G) plus those 13 fatigue items build the FACT-Fatigue (FACT-F). Nevertheless, the scale developers demonstrated that the fatigue subscale can stand alone as a brief, reliable, and valid measure [6]. In the last 20 years their assumption was confirmed. The FACIT-Fatigue Scale fulfills the criteria defined to describe a measure of high quality. Item development and validation were done according to state-of-the-art criteria based on semistructured interviews with patients and medical experts [6].

In line with other FACIT scales, high scores connote better health status. The recall period refers to the "past 7 days," a time frame delivering interpretable information on self-reported fatigue [7]. With an average completion time of 2 to 3 minutes, response burden is low. The scale fulfills the criteria of objectivity and ease in administration because of clear instructions, detailed administration, scoring, and interpretation guidelines available on www.facit.org.

* *Address correspondence to:* Inka Montan, Department of General Internal Medicine and Psychosomatics, Heidelberg University Hospital, Thibautstr. 4, Heidelberg 69115, Germany.

E-mail: inka.montan@med.uni-heidelberg.de.

1098-3015$36.00 – see front matter

https://doi.org/10.1016/j.jval.2018.03.013

Originally, the Patient-Reported Outcomes Measurement Information System Cooperative Group proposed two conceptual dimensions of fatigue, namely, fatigue experience and fatigue impact (meaning the impact of fatigue on daily functioning) [8]. With this in mind, Cella et al. [9] assessed the dimensionality of the FACIT-Fatigue Scale in the general population and in clinical samples. They found that the unidimensional model fitted the data well. Further analysis led to their conclusion that "fatigue experience and the impact of fatigue upon function are reported along a single dimensional continuum, but that *experience* is more likely than *impact upon function* to be endorsed at lower levels of fatigue" [9,p1441]. The items of the FACIT-Fatigue Scale are part of the Patient-Reported Outcomes Measurement Information System fatigue item bank [10–12].

Internal consistency of the FACIT-Fatigue Scale is high with the Cronbach coefficient α exceeding 0.90 in the US general population [11], in patients with cancer [9,13], and in other disease populations [13–15]. Convergent and discriminant validity of the scale was proven by the scale's authors [6] and confirmed in subsequent studies [16]. Criterion-related validity of the FACIT-Fatigue Scale has been shown on behalf of many criteria, for example, hemoglobin level and performance status [6,17]. Comparisons of patient groups, for example, patients with Parkinson disease [18], chronic kidney disease [16], or chronic obstructive pulmonary disease [19], with controls also supported the scale's criterion-related validity.

The scale was translated into more than 50 languages following state-of-the-art criteria [20,21] (visit www.facit.org for an overview of current languages).

The FACIT-Fatigue Scale has proven its reliability and validity in numerous clinical reference groups of patients with cancer [5,6,13,22] and other diseases [13,16,18,19,23]. In contrast, representative data of the scale in general populations are scarce [5,11,24]. Although FACIT.org provides information on general population testing, published sex- and age-specific general population percentile norms for the FACIT-Fatigue Scale are lacking. Percentile rank values would allow ranking patients' scores in relation to the reference data, thereby contributing to a meaningful interpretation of fatigue severity levels.

Systematic differences in fatigue severity are described in association with sociodemographic characteristics. Consistently, higher fatigue severity is described for women than for men in clinical samples [25–27] and in general population samples [11,28–31]. Results with respect to fatigue and age are less clear; some studies found younger people to be more fatigued than older ones [11,32], whereas other studies showed the opposite association [30,31]. One study found an age effect in a sample of patients with cancer, but not in the general population [24]. With the aim to provide normative data, such potential associations need to be addressed.

The objective of our study was to assess the FACIT-Fatigue Scale in a representative sample of the German general population. In detail, we aimed to 1) examine the psychometric characteristics of the scale, 2) assess its dimensionality, 3) describe associations between the FACIT-Fatigue Scale and sociodemographic characteristics, and 4) deliver percentile ranks of the FACIT-Fatigue Scale in the German general population, taking potential effects of sex and age into account.

Methods

Study Design and Participants

The validation and standardization of the FACIT-Fatigue Scale in the German general population were part of a nationally representative face-to-face household survey conducted with the assistance of a demographic consulting company (USUMA, Berlin, Germany). The country was separated into 258 sample areas. Within these areas, street, house, and household were selected randomly with the random-route technique. Members of this household fulfilling the inclusion criteria (age ≥14 years and fluent in German language) were also selected randomly by using the Kish technique. The interviewers asked the target person to take part in the study. If not found at home, four attempts were made to contact the selected person. All participants were visited by a study assistant and informed about the investigation, and self-rating questionnaires were presented. Informed consent was obtained from all participants. The study was approved by the Ethics Committee of the University of Leipzig.

Measures

FACIT-Fatigue Scale

Survey participants completed the FACIT-Fatigue Scale [5,6]. It is formatted for self-administration on one page and uses a five-point Likert-type scale (0 = not at all; 1 = a little bit; 2 = somewhat; 3 = quite a bit; and 4 = very much). All items contribute to the sum score with equal weight. The scale range is 0 to 52, with 0 being the worst possible score and 52 being the best possible score indicating no fatigue. For this purpose, each negatively worded item response is recoded.

Patient Health Questionnaire-4

The Patient Health Questionnaire-4 (PHQ-4) [33] is an ultra-brief self-report instrument that consists of a two-item subscale assessing depressive symptoms (PHQ-2) and a two-item subscale assessing anxiety symptoms (GAD-2) using a four-point Likert-type scale. The PHQ-2 and GAD-2 sum scores' range is 0 to 6 and the composite PHQ-4 score range is 0 to 12, with higher scores meaning higher symptom burden. The German translation followed state-of-the art criteria including seven steps of translation and blind back-translation by four independent translators [33]. Reliability and validity of the measure have been shown [33–35].

Data Analyses

Internal consistency of the scale was assessed using Cronbach α [36]. Item characteristics including item means and SDs, part-whole-corrected item-test correlations (r_{it}), and Cronbach α if the respective item was deleted were determined. The r_{it} describes the correlation between the respective item and the scale sum score without this item. Cronbach α of 0.70 or more is considered to be sufficiently high for comparing groups; for the clinical situation, however, when an individual's scale value is of interest, Cronbach α of 0.90 is the minimum required [36].

We tested for the unidimensional structure of the scale, which was empirically proven before [9] with confirmatory factor analysis (CFA) and bifactor analysis. Criteria of a root mean squared error of approximation (RMSEA) less than 0.10, the Tucker-Lewis index (TLI), and a comparative fit index (CFI) more than 0.90 are set to define acceptable model fit [37,38].

In addition, we tested a model that was hypothesized before but rejected in favor of the unidimensional solution in a previous study [9], a bifactor model with a general factor "overall fatigue" (defined by loadings of all 13 items) and two subdomain factors "experience" (defined by loadings of the items An1, An2, An5, HI7, and HI12) and "impact" (defined by loadings of the items An3, An4, An7, An8, An12, An14, An15, and An16). A general factor is justified if the loadings of all items on this factor are greater than 0.3 and if the loadings of the items are higher on the general factor than on the subdomain factors. A subdomain factor is justified if loadings of the respective items on this factor are greater than 0.3 [9,39,40]. χ^2 values are not considered as absolute fit indicators because of the large sample size.

Age and sex differences in fatigue mean scores were tested with a two-factor analysis of variance (ANOVA), using the factors sex (two categories) and age group (six categories as presented in Table 1). Effect sizes were calculated using Cohen *d* [41], relating the mean score differences to the pooled SD. Differences in fatigue mean scores referring to civil status, education, and employment status were each tested with one-way ANOVAs. To further evaluate construct validity, intercorrelations of the FACIT-Fatigue Scale with the PHQ-4 and its subscales for depression and anxiety were determined using Pearson correlations. To provide population-based normative data for the FACIT-Fatigue Scale, we computed sex- and age-specific percentiles for the whole score range (0–52).

Statistical analyses were conducted using SPSS version 20 (SPSS Inc., Chicago, IL) and AMOS version 5 (Chicago, IL).

Results

Sample Characteristics

The study was performed between March and May 2015. A first attempt to contact study candidates was made at 4902 addresses, of which 4844 were valid. Out of these, 1456 subjects (30.0%) refused participation, 766 subjects (15.8%) were not reached after four attempts, 27 subjects (0.6%) were out of town, and 19 subjects (0.4%) refused participation because of severe health problems. A total of 2576 people (aged 14–94 years) agreed to participate, gave informed consent, and completed the questionnaires. Participation rate was 53.2%. Our study population was restricted to adults. Therefore, 66 (2.6%) subjects younger than 18 years were excluded. A data set of 2510 people was included in this study. The sample is representative of the adult German population in terms of age and sex.

Among the 2510 persons, 2426 (44.3% men) completed the FACIT-Fatigue Scale with fewer than three missing items. The guidelines (www.facit.org) recommend that at least 80% of the items should be valid. Missing items were replaced with the rounded mean of the person's valid items. The data of these 2426 persons built the basis for the subsequent analyses. Sample characteristics are presented in Table 1.

Item Characteristics, Internal Consistency, and Factor Structure

Item characteristics are presented in Table 2. Part-whole-corrected item-test correlations (r_{it}) were high for all items being greater than 0.50, except for item An7 ("I am able to do my usual activities") being still moderate with an r_{it} of 0.41.

The Cronbach α of the FACIT-Fatigue Scale was high at 0.92, thus being sufficiently precise for individual classification. The coefficients for the Cronbach α if one item was deleted were between 0.91 and 0.93. Every single item contributed to the internal consistency in slightly improving the Cronbach α except item An7. Nevertheless, if this item was not part of the scale, the Cronbach α would increase only marginally from 0.92 to 0.93.

CFA results supported the unidimensionality of the scale by factor loadings of all 13 items on a single factor ranging from 0.38

Table 1 – Sociodemographic characteristics of the study sample.

Characteristic	Male (n = 1074)	Female (n = 1352)	Total (N = 2426)
Age (y), mean ± SD	49.7 ± 17.3	49.9 ± 17.5	49.8 ± 17.4
Age category (y), n (%)			
18–29	176 (16.4)	211 (15.6)	387 (16.0)
30–39	163 (15.2)	215 (15.9)	378 (15.6)
40–49	183 (17.0)	239 (17.7)	422 (17.4)
50–59	211 (19.6)	252 (18.6)	463 (19.1)
60–69	184 (17.1)	219 (16.2)	403 (16.6)
≥70	157 (14.6)	216 (16.0)	373 (15.4)
Civil status, n (%)			
Married, living together	533 (49.6)	565 (41.8)	1098 (45.3)
Married, living apart	21 (2.0)	32 (2.4)	53 (2.2)
Single	330 (30.7)	334 (24.7)	664 (27.4)
Divorced	125 (11.6)	209 (15.5)	334 (13.8)
Widowed	62 (5.8)	206 (15.2)	268 (11.0)
Missing	3 (0.3)	6 (0.4)	9 (0.4)
Education (number of years), n (%)			
≤9	386 (35.9)	456 (33.7)	824 (34.7)
10–11	446 (41.5)	617 (45.6)	1063 (43.8)
≥12	241 (22.4)	276 (20.4)	517 (21.3)
Missing	1 (0.1)	3 (0.2)	4 (0.2)
Religion, n (%)			
No religious affiliation	347 (32.3)	328 (24.3)	675 (27.8)
Religious affiliation	721 (67.1)	1020 (75.4)	1741 (71.8)
Missing	6 (0.6)	4 (0.3)	10 (0.4)
Employment, n (%)			
Working full-time	594 (55.3)	400 (29.6)	994 (41.0)
Working part-time	50 (4.7)	324 (24.0)	374 (15.4)
Unemployed	67 (6.2)	68 (5.0)	135 (5.6)
Housewife/man	7 (0.7)	100 (7.4)	107 (4.4)
Retired	297 (27.7)	386 (28.6)	683 (28.2)
In education*	59 (5.5)	74 (5.5)	133 (5.5)

* School, university, professional education.

Table 2 – Item characteristics of the FACIT-Fatigue Scale.

Item		Mean ± SD	r_{it}	α(del.)
HI7	I feel fatigued.	3.03 ± 0.99	0.71	0.91
HI12	I feel week all over.	3.38 ± 0.90	0.79	0.91
An1	I feel listless ("washed out").	3.35 ± 0.93	0.77	0.91
An2	I feel tired.	3.07 ± 0.99	0.75	0.91
An3	I have trouble starting things because I am tired.	3.42 ± 0.91	0.83	0.91
An4	I have trouble finishing things because I am tired.	3.49 ± 0.87	0.79	0.91
An5	I have energy.	2.57 ± 1.04	0.57	0.92
An7	I am able to do my usual activities.	3.14 ± 1.10	0.41	0.93
An8	I need to sleep during the day.	3.16 ± 1.01	0.62	0.92
An12	I am too tired to eat.	3.84 ± 0.51	0.55	0.92
An14	I need help doing my usual activities.	3.76 ± 0.65	0.52	0.92
An15	I am frustrated by being too tired to do the things I want to do.	3.64 ± 0.78	0.74	0.91
An16	I have to limit my social activity because I am tired.	3.67 ± 0.72	0.71	0.91

Note. Means are reported after transformation according to scoring guidelines. All items other than An5 and An7 are reversed such that 0 = 4, 1 = 3, 2 = 2, 3 = 1, and 4 = 0. r_{it} is part-whole-corrected item-test correlation and α(del.) is α if item deleted.
FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue.

to 0.89 (Fig. 1A). Nevertheless, fit statistics with an RMSEA of 0.144 (criterion <0.10), a CFI of 0.846 (criterion >0.90), and a TLI of 0.815 (criterion >0.90) were below desired levels.

Bifactor analysis as well supported a general factor solution with loadings of all 13 items ranging between 0.37 and 0.94 (Fig. 1B) and loadings of each of the items being higher on the general factor than on the respective subdomain factor. The analysis, however, did not support the existence of the subdomain factors "experience" and "impact" because not all the loadings of the items on the respective factor were higher than 0.3 (Fig. 1B). Fit statistics of the bifactor model with a CFI of 0.928 (criterion >0.90), a TLI of 0.892 (criterion >0.90), and an RMSEA of 0.110 (criterion <0.10) were better than for the unidimensional model but still only in part reached the outlined criteria for acceptable fit.

Fatigue Scores in Relation to Sociodemographic Characteristics, Construct Validity

The FACIT-Fatigue Scale mean score of the total sample was 43.5 ± 8.3, and corresponding scores for males and females, respectively, were 44.6 ± 7.4 and 42.7 ± 8.9, indicating more fatigue among women. Figure 2 illustrates fatigue mean scores depending on sex and age. In all age groups, females had lower mean scores than males. In the age range between 18 and 69 years, there were only marginal differences in fatigue mean scores between the defined age groups (see Fig. 2). The oldest age group (≥70 years), however, showed higher fatigue severity compared with the others. The ANOVA results were as follows: age group: $F = 23.18$, $P < 0.001$; sex: $F = 4.86$, $P < 0.001$; Age group × Sex: $F = 0.30$, $P = 0.878$ (i.e., no interaction). There was a small effect of fatigue differences due to sex with Cohen d of 0.23.

Other sociodemographic characteristics also explained differences in the fatigue level. There were highly significant differences in fatigue severity with respect to socioeconomic status (Table 3). Married people living together with one's partner showed least fatigue, whereas divorced people showed the highest fatigue levels. For women, the employment status also accounted for highly significant differences, with those working full-time or being a housewife showing least fatigue severity and those being unemployed showing the highest fatigue burden. Years of education did not account for any differences in the fatigue level.

The FACIT-Fatigue Scale score was correlated to the PHQ-4 with an r of −0.71. It was correlated to the PHQ-2 depression scale with an r of −0.69, and to the GAD-2 anxiety scale with an r of −0.63, indicating higher fatigue severity to be associated with higher levels of depression and anxiety.

Normative Data

Table 4 presents percentiles for the FACIT-Fatigue Scale scores on the basis of a representative sample of the German general population. Because males and females systematically differed in their fatigue levels and because the age group of 70 years and older systematically differed from the age groups between 18 and 69 years, percentiles for these two age groups are provided separately for males and females. Reference values based on the total sample of the German general population are presented in the last column of Table 4.

An example may illustrate the application of Table 4. For a 40-year-old man with a FACIT-Fatigue Scale score of 36, Table 4 indicates a percentile rank of 12.9% in his respective sex- and age-specific group. The percentile rank for a woman of the same age with the same FACIT-Fatigue Scale score would be 18.3%.

Discussion

In this study, the internationally used FACIT-Fatigue Scale was assessed in a large representative sample of the German general population. Reliability, validity, and the factorial structure of the scale were examined, and population-based sex- and age-specific percentile norms were provided.

Sample Characteristics

To our knowledge, to date there are few studies available of the FACIT-Fatigue Scale assessed in representative general population samples. Of the US population, one study included 1010 participants assessed via phone interviews in the 1990s [5]. Another study reported 1072 respondents approached via Internet survey panel, also in the 1990s [24,42]. A third study assessed 666 participants also via Internet survey in the years 2006 and 2007 [11,43]. Now, with the German general population, we extend these results to a nationally representative sample of 2426 participants, assessed via face-to-face household survey in the year 2015.

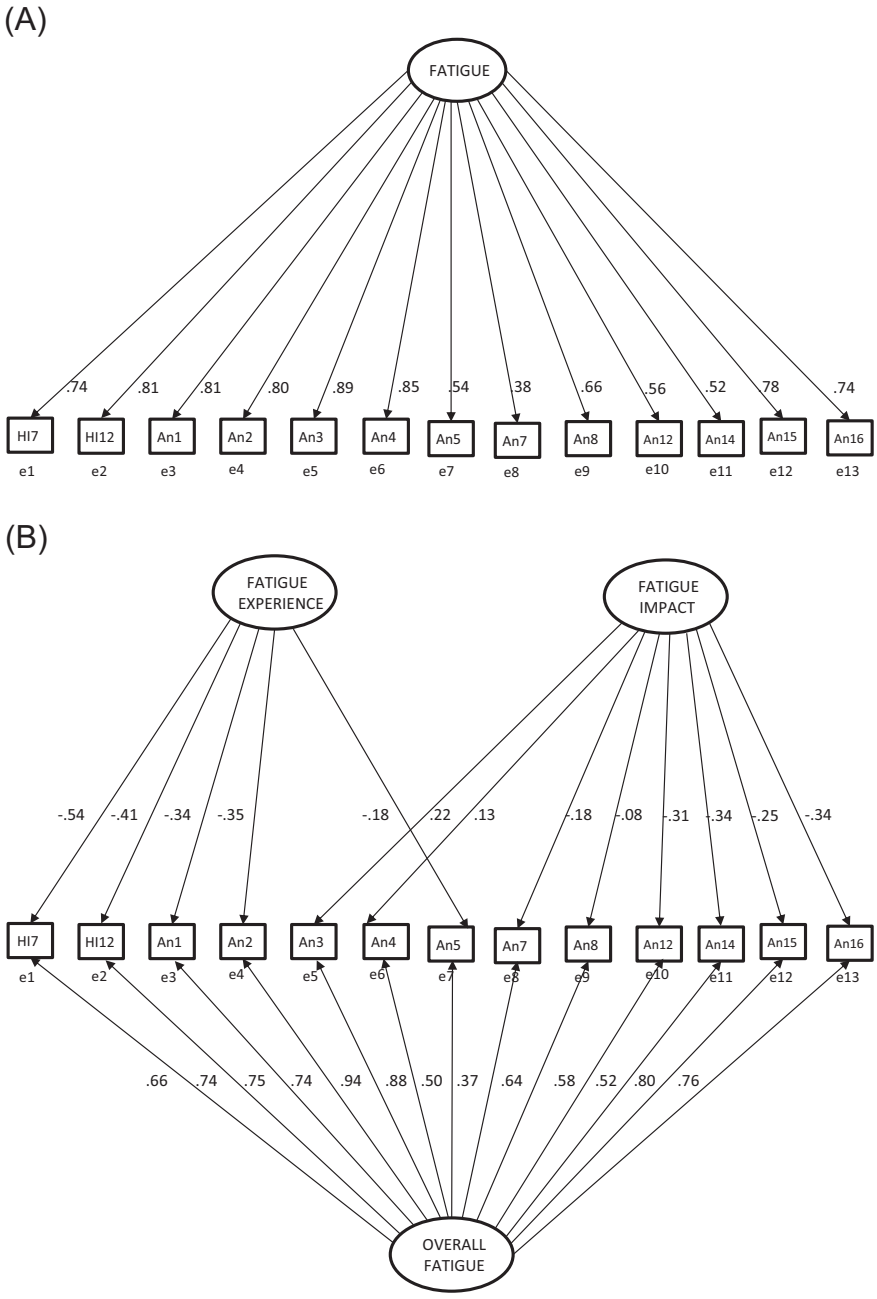

Fig. 1 – (A) CFA result (one-factor model) and (B) bifactor analysis result of the FACIT-Fatigue Scale. CFA, confirmatory factor analysis; FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue.

The FACIT-Fatigue Scale mean score of the German general population sample (43.5 ± 8.3) was almost identical to that of the US general population sample (43.6 ± 9.4) as published in Cella et al. [5] and similar to the mean score of another US general population sample as reported by Butt et al. [24] (46.6 ± 7.2). Compared with these results, fatigue severity was higher in the study by Junghaenel et al. [11], with the FACIT-Fatigue Scale mean score being approximately 38.5 (we inferred this mean score from the reported data in the cited article). One reason for the relatively high levels of fatigue in the study by Junghaenel et al. may be the high proportion of people with health problems in this Internet survey. Among the study participants, 34.5% suffered from hypertension, 20.4% from depression, and 15.8% from anxiety.

Our study showed that, as to be expected, the general population proved to be markedly less fatigued than clinical samples, for example, patients suffering from stroke (FACIT-Fatigue 38.1 ± 9.6) [13], cancer (FACIT-Fatigue 34.6 ± 10.1) [22], or iron deficiency anemia (FACIT-Fatigue 24.1 ± 11.8) [44], thereby providing evidence for the construct validity of the FACIT-Fatigue Scale.

Internal Consistency and Factorial Validity

Internal consistency of the FACIT-Fatigue Scale with an α of 0.92 found in our study was similar to results in the US general population ($\alpha = 0.96$) [11], and to results in clinical groups, for example, patients receiving maintenance dialysis ($\alpha = 0.92$) [45] or

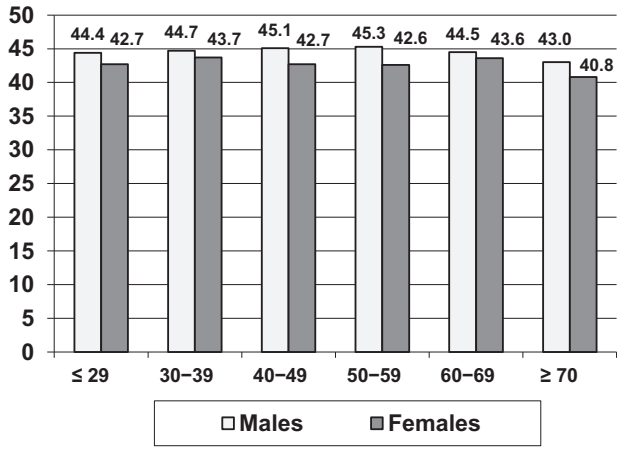

Fig. 2 – FACIT-Fatigue Scale mean scores by sex and age group. FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue.

patients suffering from cancer, stroke, or HIV ($\alpha \geq 0.91$ for each group) [13].

Analogous to the results reported by Cella et al. [9], CFA as well as bifactor analysis in our study supported the one-factor structure of the FACIT-Fatigue Scale. Both analyses showed loadings of all 13 items greater than 0.3 on a common factor, whereas loadings on the hypothesized subdomains "experience" and "impact" did not reach this criterion in the bifactor analysis.

Fit indices of the CFA did not meet the criteria; for the bifactor model they were slightly better but still suboptimal, with only the CFI (0.092) reaching the criterion. Compared with our results, Cella et al. [9] reported better fit indices for the CFA and bifactor analysis of the FACIT-Fatigue Scale in the US general population, with only the RMSEA in the CFA being below the set criterion (RMSEA = 0.176).

The fact that in our study as well as in the study by Cella et al. [9] the bifactor model has slightly better fit indices than the one-factorial CFA is not surprising to us because every factor added in an analysis improves explained variance. In our view, the improvement of fit indices by adding factors alone does not justify to favor this model if it shows insufficient item factor loadings.

The unsatisfying fit indices of the CFA results in our study remain an open question. Poor fit indices for factorial analyses of a 72-item fatigue item bank that nevertheless supports the unidimensional nature of the scale by the item loadings on the common factor were also reported in another study [39].

Taken all together, several methods used in our study confirm the unidimensionality of the FACIT-Fatigue Scale, namely, 1) a high internal consistency of $\alpha = 0.92$, 2) all item-total correlations being greater than 0.3, 3) loadings of all 13 items on one factor being greater than 0.3 in the CFA, and 4) loadings of all 13 items on the general factor being greater than 0.3 and the loading of each item on the general factor being higher than on the subdomain factor, whereas loadings on the subdomain factors in part did not reach the criterion (>0.03) to support the subdomains "experience" and "impact" in the bifactor analysis.

Fatigue Scores in Relation to Sex and Age, Construct Validity

In accordance to other studies assessing fatigue in general population samples [11,28–31], we also found women to be slightly (but significantly) more fatigued than men. Because this sex difference is well documented in other surveys, this further supports the construct validity of the FACIT-Fatigue Scale.

With respect to age, the picture is less clear. The age range of our sample was 18 to 94 years. We found the oldest group of 70 years or older to be significantly more fatigued compared with the younger age groups. Skapinakis et al. [30] (age range 16–64 years) and Schwarz et al. [31] (age range 14–92 years) also found higher fatigue levels in older people. They described an almost linear positive association between age and fatigue severity. In contrast to these results, Junghaenel et al. [11] (age range <30 to ≥75 years) found in their Internet survey that younger participants showed higher fatigue severity. Ricci et al. [32] (age range 18–65 years) also found younger people to be more fatigued than older ones. This result could also be explained as a reflection of sampling bias in these studies, enrolling a higher proportion of

Table 3 – FACIT-Fatigue Scale mean scores and demographic variables.

Variable	Male (n = 1074), mean ± SD	Female (n = 1352), mean ± SD	Total (N = 2426), mean ± SD
Civil status	$F = 4.9$; $P < 0.001$	$F = 5.7$; $P < 0.001$	$F = 12.1$; $P < 0.001$
Married, living together	45.6 ± 6.6	44.2 ± 7.7	44.8 ± 7.2
Married, living apart	43.7 ± 7.9	42.4 ± 7.3	42.8 ± 7.5
Single	43.7 ± 8.0	41.8 ± 9.8	42.8 ± 9.0
Divorced	42.8 ± 8.7	41.1 ± 10.4	41.8 ± 9.8
Widowed	44.1 ± 6.3	41.6 ± 8.7	42.2 ± 8.2
Education (number of years)	$F = 6.9$; $P = 0.001$	$F = 0.1$; $P = 0.906$	$F = 2.4$; $P = 0.089$
≤9	43.6 ± 7.6	42.3 ± 9.2	42.9 ± 8.6
10–11	45.8 ± 6.7	43.0 ± 8.8	44.2 ± 8.1
≥12	43.9 ± 7.9	42.6 ± 8.4	43.2 ± 8.2
Employment	$F = 4.0$; $P = 0.001$	$F = 5.8$; $P < 0.001$	$F = 9.4$; $P = 0.001$
Working full-time	45.4 ± 7.2	44.0 ± 8.3	44.9 ± 7.7
Working part-time	41.9 ± 7.5	43.2 ± 8.4	43.0 ± 8.3
Unemployed	42.6 ± 8.8	38.7 ± 12.6	40.6 ± 11.0
Housewife/man	46.1 ± 5.1	43.9 ± 7.8	44.2 ± 7.7
Retired	43.7 ± 7.2	41.7 ± 8.9	42.6 ± 8.2
In education*	44.4 ± 8.0	40.2 ± 10.2	42.0 ± 9.5

Note. *F* values refer to the ANOVA test statistics.

ANOVA, analysis of variance; FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue.

* School, university, professional education.

Table 4 – FACIT-Fatigue Scale normative data from the general population: percentile scores.

FACIT-Fatigue Scale score	Male		Female		Total
	<70 y (n = 917)	≥70 y (n = 157)	<70 y (n = 1136)	≥70 y (n = 216)	(N = 2426)
0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0
6	0.0	0.0	0.2	0.5	0.1
8	0.0	0.0	0.4	0.9	0.3
10	0.0	0.0	0.4	0.9	0.3
12	0.0	0.0	0.8	0.9	0.5
14	0.2	0.6	1.4	1.4	0.9
16	0.3	1.3	2.2	2.8	1.5
18	0.7	1.3	3.0	3.2	2.0
20	1.1	1.3	3.5	4.6	2.6
22	1.7	2.5	4.1	6.0	3.3
24	2.9	3.2	5.1	7.4	4.4
26	4.1	3.8	6.1	9.7	5.5
28	5.2	5.7	7.7	10.2	6.8
30	7.0	7.0	9.2	12.5	8.5
32	8.4	9.6	11.8	14.8	10.6
34	9.8	12.1	15.1	19.9	13.4
36	12.9	15.3	18.3	25.5	16.7
38	15.8	19.7	23.2	31.9	20.9
40	20.0	24.8	28.6	36.1	25.8
42	26.1	31.8	35.5	44.0	32.4
44	36.1	49.0	45.4	57.4	43.2
46	45.3	62.4	55.0	69.0	53.1
48	58.6	79.0	65.8	81.9	65.3
50	78.5	93.6	83.1	95.4	83.1
52	100.0	100.0	100.0	100.0	100.0

FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue.

people with limiting health conditions than one sees in a representative sample drawn from the general population. Also, only working adults aged 18 to 65 years were included in the study by Ricci et al. [32], thereby not covering the 70 years and older age group, which accounted for the age effect in our study. Butt et al. [24] (age range 18 to ≥80 years) did not find evidence for increased fatigue with age in the general population, but they did find this age effect in patients with cancer.

To summarize, there are significant sex and age effects on the FACIT-Fatigue Scale score that have to be taken into account when interpreting those scores.

In our study, fatigue was highly associated to depression and anxiety, with slightly higher intercorrelations to depression than to anxiety. The same constellation was found in another study that used different instruments to assess these constructs [31], thereby providing further evidence for the construct validity of the measure.

Sex- and Age-Specific Percentile Norms

The FACIT-Fatigue Scale is used worldwide to assess fatigue severity in diverse clinical groups [5,6,13,16,18,19,22,23,44–46]. In light of this background it is surprising that, according to our knowledge, sex- and age-specific population-based percentile norms of the FACIT-Fatigue Scale are not available in the literature yet. This, however, is helpful and is needed to interpret and understand the respective clinical data [47]. There are studies available reporting FACIT-Fatigue data described as representative for the US general population [5,11,24,42]. These, however, do not provide sex- and age-specific percentile norms. There is one study providing sex-specific percentile norms for FACT-General [47] but not for the FACIT-Fatigue Scale. To our knowledge, our study is the first in providing sex- and age-specific population-based percentile norms for the FACIT-Fatigue Scale.

Study Limitations

The response rate of this study (53.2%) was sufficient but not optimal. Although the sample was fairly representative of people living in private households in terms of age and sex, people with severe degrees of fatigue might be under-represented because of their symptom burden—a potential limitation common to general population studies.

With the intention to adopt an instrument that was developed in another language and—inherent with that—in another culture, you have to decide whether you want to keep the translated instrument with its psychometric strengths and weaknesses in its original structure and content and thereby ensure comparability or you want to change single items or wordings to improve, for example, factor loadings or fit indices at the expense of comparability.

Our opinion is to keep the structure and content of the original instrument as far as psychometric conditions are not violated. In our study, for example, item An7 showed a markedly lower factor loading on the common factor in the CFA (0.38 compared with the other item factor loadings ranging between 0.52 and 0.89), a result described previously [46]. Nevertheless, because psychometric conditions are not violated, it is justified

and even necessary to keep this item for the sake of international comparability. As a possible next research step, language-related differential item functioning analyses could be conducted to ensure that the instrument "works" after its translation into the other language [48–50].

A limitation in the application of our study results is that for the international use of the provided data it is a precondition that characteristics of the general population sample of the respective country closely resemble those of the German general population sample the normative data are based on.

Conclusions

In our study, the FACIT-Fatigue Scale was assessed in a large representative sample of the German general population. Psychometric characteristics, factorial analyses, and associations to sociodemographic data and related constructs confirmed the reliability and validity of the unidimensional scale. Sex- and age-specific percentile norms for the entire range of FACIT-Fatigue Scale scores were provided for the first time, thereby contributing to a meaningful interpretation of the respective clinical research data.

Source of financial support: The study was supported by intramural means of the University Medical Center Hamburg-Eppendorf, Germany, and the University of Leipzig, Germany. External funding was not received.

REFERENCES

[1] Ramsenthaler C, Kane P, Gao W, et al. Prevalence of symptoms in patients with multiple myeloma: a systematic review and meta-analysis. Eur J Haematol 2016;97:416–29.
[2] Wang XS, Zhao F, Fisch MJ, et al. Prevalence and characteristics of moderate to severe fatigue: a multicenter study in cancer patients and survivors. Cancer 2014;120:425–32.
[3] Reeve BB, Mitchell SA, Dueck AC, et al. Recommended patient-reported core set of symptoms to measure in adult cancer treatment trials. J Natl Cancer Inst 2014;106:dju129.
[4] Patrick D, Gyatt G, Acquadro C. Patient-reported outcomes. In: Higgins J, Green S, eds., Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. The Cochrane Collaboration, 2011. Available from: www.handbook.cochrane.org. [Accessed September 21, 2017].
[5] Cella D, Lai JS, Chang CH, et al. Fatigue in cancer patients compared with fatigue in the general United States population. Cancer 2002;94:528–38.
[6] Yellen SB, Cella DF, Webster K, et al. Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. J Pain Symptom Manage 1997;13:63–74.
[7] Lai JS, Cook K, Stone A, et al. Classical test theory and item response theory/Rasch model to assess differences between patient-reported fatigue using 7-day and 4-week recall periods. J Clin Epidemiol 2009;62:991–7.
[8] Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care 2007;45 (Suppl. 1):S3–11.
[9] Cella D, Lai JS, Stone A. Self-reported fatigue: one dimension or more? Lessons from the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) questionnaire. Support Care Cancer 2011;19:1441–50.
[10] Christodoulou C, Junghaenel DU, DeWalt DA, et al. Cognitive interviewing in the evaluation of fatigue items: results from the Patient-Reported Outcomes Measurement Information System (PROMIS). Qual Life Res 2008;17:1239–46.
[11] Junghaenel DU, Christodoulou C, Lai JS, Stone AA. Demographic correlates of fatigue in the US general population: results from the Patient-Reported Outcomes Measurement Information System (PROMIS) initiative. J Psychosom Res 2011;71:117–23.
[12] Lai JS, Cella D, Choi S, et al. How item banks and their application can influence measurement practice in rehabilitation medicine: a PROMIS fatigue item bank example. Arch Phys Med Rehabil 2011;92(Suppl): S20–7.
[13] Butt Z, Lai JS, Rao D, et al. Measurement of fatigue in cancer, stroke, and HIV using the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F) scale. J Psychosom Res 2013;74:64–8.
[14] Hagell P, Hoglund A, Reimer J, et al. Measuring fatigue in Parkinson's disease: a psychometric study of two brief generic fatigue questionnaires. J Pain Symptom Manage 2006;32:420–32.
[15] Lai JS, Beaumont JL, Ogale S, et al. Validation of the Functional Assessment of Chronic Illness Therapy-Fatigue scale in patients with moderately to severely active systemic lupus erythematosus, participating in a clinical trial. J Rheumatol 2011;38:672–9.
[16] Macdonald JH, Fearn L, Jibani M, Marcora SM. Exertional fatigue in patients with CKD. Am J Kidney Dis 2012;60:930–9.
[17] Cella D, Eton DT, Lai JS, et al. Combining anchor and distribution-based methods to derive minimal clinically important differences on the Functional Assessment of Cancer Therapy (FACT) anemia and fatigue scales. J Pain Symptom Manage 2002;24:547–61.
[18] Lindqvist D, Kaufman E, Brundin L, et al. Non-motor symptoms in patients with Parkinson's disease—correlations with inflammatory cytokines in serum. PLoS One 2012;7:e47387.
[19] Stridsman C, Mullerova H, Skar L, Lindberg A. Fatigue in COPD and the impact of respiratory symptoms and heart disease—a population-based study. COPD 2013;10:125–32.
[20] Bonomi AE, Cella DF, Hahn EA, et al. Multilingual translation of the Functional Assessment of Cancer Therapy (FACT) quality of life measurement system. Qual Life Res 1996;5:309–20.
[21] Eremenco SL, Cella D, Arnold BJ. A comprehensive method for the translation and cross-cultural validation of health status questionnaires. Eval Health Prof 2005;28:212–32.
[22] Leung YW, Brown C, Cosio AP, et al. Feasibility and diagnostic accuracy of the Patient-Reported Outcomes Measurement Information System (PROMIS) item banks for routine surveillance of sleep and fatigue problems in ambulatory cancer care. Cancer 2016;122:2906–17.
[23] Villoria A, Garcia V, Dosal A, et al. Fatigue in out-patients with inflammatory bowel disease: prevalence and predictive factors. PLoS One 2017;12:e0181435.
[24] Butt Z, Rao AV, Lai JS, et al. Age-associated differences in fatigue among patients with cancer. J Pain Symptom Manage 2010;40:217–23.
[25] Meeuwesen L, Bensing J, Brink-Muinen A. Communicating fatigue in general practice and the role of gender. Patient Educ Couns 2002;48:233–42.
[26] Nakao M, Fricchione G, Zuttermeister PC, et al. Effects of gender and marital status on somatic symptoms of patients attending a mind/body medicine clinic. Behav Med 2001;26:159–68.
[27] Ridsdale L, Evans A, Jerrett W, et al. Patients with fatigue in general practice: a prospective study. BMJ 1993;307:103–6.
[28] Dochi M, Suwazono Y, Oishi M, et al. The relation between cumulative fatigue and marital status in Japanese workers. Behav Med 2007;33:55–65.
[29] Loge JH, Ekeberg O, Kaasa S. Fatigue in the general Norwegian population: normative data and associations. J Psychosom Res 1998;45:53–65.
[30] Skapinakis P, Lewis G, Meltzer H. Clarifying the relationship between unexplained chronic fatigue and psychiatric morbidity: results from a community survey in Great Britain. Int Rev Psychiatry 2003;15: 57–64.
[31] Schwarz R, Krauss O, Hinz A. Fatigue in the general population. Onkologie 2003;26:140–4.
[32] Ricci JA, Chee E, Lorandeau AL, Berger J. Fatigue in the U.S. workforce: prevalence and implications for lost productive work time. J Occup Environ Med 2007;49:1–10.
[33] Löwe B, Wahl I, Rose M, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. J Affect Disord 2010;122:86–95.
[34] Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. Psychosomatics 2009;50:613–21.
[35] Wild B, Eckl A, Herzog W, et al. Assessing generalized anxiety disorder in elderly people using the GAD-7 and GAD-2 scales: results of a validation study. Am J Geriatr Psychiatry 2014;22:1029–38.
[36] Bland JM, Altman DG. Cronbach's alpha. BMJ 1997;314:572.
[37] Hu L, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 2009;6:1–55.
[38] Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). Med Care 2007;45(Suppl. 1):S22–31.
[39] Lai JS, Crane PK, Cella D. Factor analysis techniques for assessing sufficient unidimensionality of cancer related fatigue. Qual Life Res 2006;15:1179–90.
[40] McDonald RP. Test Theory: A Unified Treatment. Mahwah, NJ: Lawrence Erlbaum, 1999.

[41] Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum, 1988.
[42] Cella D, Zagari MJ, Vandoros C, et al. Epoetin alfa treatment results in clinically significant improvements in quality of life in anemic cancer patients when referenced to the general population. J Clin Oncol 2003;21:366–73.
[43] Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005–2008. J Clin Epidemiol 2010;63:1179–94.
[44] Acaster S, Dickerhoof R, DeBusk K, et al. Qualitative and quantitative validation of the FACIT-Fatigue scale in iron deficiency anemia. Health Qual Life Outcomes 2015;13:60.
[45] Wang SY, Zang XY, Liu JD, et al. Psychometric properties of the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue) in Chinese patients receiving maintenance dialysis. J Pain Symptom Manage 2015;49:135–43.
[46] Kwakkenbos L, Willems LM, Baron M, et al. The comparability of English, French and Dutch scores on the Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F): an assessment of differential item functioning in patients with systemic sclerosis. PLoS One 2014;9:e91979.
[47] Brucker PS, Yost K, Cashy J, et al. General population and cancer patient norms for the Functional Assessment of Cancer Therapy-General (FACT-G). Eval Health Prof 2005;28:192–211.
[48] Fischer HF, Wahl I, Nolte S, et al. Language-related differential item functioning between English and German PROMIS Depression items is negligible. Int J Methods Psychiatr Res 2017. http://dx.doi.org/10.1002/mpr.1530.
[49] Oude Voshaar MA, Ten Klooster PM, Glas CA, et al. Calibration of the PROMIS physical function item bank in Dutch patients with rheumatoid arthritis. PLoS One 2014;9:e92367.
[50] Paz SH, Spritzer KL, Morales LS, Hays RD. Evaluation of the Patient-Reported Outcomes Information System (PROMIS®) Spanish-language physical functioning items. Qual Life Res 2013;22:1819–30.