

Development and Psychometric Evaluation of the Nurses' Compassion Fatigue Inventory

Mahdieh Sabery, MSN

*Shahid Beheshti University of Medical Sciences, Tehran, Iran
Kashan University of Medical Sciences and Health Services, Kashan, Iran*

**Mansoureh Zagheri Tafreshi, PhD
Jamileh Mohtashami, PhD**

Shahid Beheshti University of Medical Sciences, Tehran, Iran

Abbas Ebadi, PhD

Baqiyatallah University of Medical Sciences, Tehran, Iran

Background and Purpose: Many efforts have been made so far to measure compassion fatigue (CF). There is no specific, precise, and comprehensive tool for CF measurement among nurses. This study aimed at developing and evaluating the psychometric properties of the Nurses' Compassion Fatigue Inventory (NCFI). **Methods:** This methodological study was made in 3 consecutive phases: the hybrid concept analysis, the item pool generation, the psychometric validation. **Results:** The initial item pool contained 98 items. The psychometric properties were examined and exploratory factor analysis revealed a 6-factor structure. NCFI contained 35 items. The Cronbach's alpha, intra-class correlation coefficient (ICC), and standard error of measurement (SEM) were 0.940, 0.935, and 7.6, respectively. **Conclusion:** The NCFI has acceptable reliability and validity. It can be easily used by nursing researchers and managers and even hospital nurses.

Keywords: compassion fatigue; secondary traumatic stress; vicarious traumatization; professional burnout; validation studies; nursing

Compassion is the core of nursing, the essence of quality health care delivery, and a significant factor behind nurses' job satisfaction (Meyer, Li, Klaristenfeld, & Gold, 2015). Because of providing empathetic care and establishing close interpersonal relationships with patients and families, some nurses may absorb their pain and agony and experience a traumatic distress (Boyle, 2011). According to Figley (2002a), compassion and empathy, which are among the main values of nursing—put nurses at risk for compassion fatigue.

The term *compassion fatigue* (CF) was first introduced in 1992 by Joinson, when assessing “nurses' burnout” in emergency departments, she noticed that they had lost their “nurturing and caring powers” (Joinson, 1992). However, she never formally explored the concept. Figley (1995) considered the concept as an appropriate term for “secondary

traumatic stress.” Nonetheless, indiscriminate use of these two terms has caused many challenges and conflicts (Coetzee & Klopper, 2010).

CF is a state of extreme fatigue and biopsychosocial disorder because of prolonged exposure to compassion stress (Melvin, 2012). Like other types of fatigue, CF also undermines one’s ability and desire to tolerate others’ agony (Lynch & Lobo, 2012). Figley (2002a, 2002b) described this phenomenon as the fine and the cost paid by care providers for their help to patients. Stamm (2010) also considered CF as the normal reward and remuneration of helping others, whereas Bride, Radey, and Figley (2007) referred to CF as an occupational hazard of clinical practice which causes different psychological injuries.

CF is associated with many negative physical, psychological, cognitive, and social consequences leading to decreased efficiency; reduced occupational satisfaction; and decreased satisfaction, safety, and care quality (Meyer et al., 2015). CF has been identified as a negative consequence and cost of caring which is capable of removing nurses’ compassionate caring forever (Sansó et al., 2015).

Because of the potential effects of CF on nurses, patients, and the organization, assessing nurses’ CF is clearly essential. Such an assessment necessitates using valid and reliable measurement tools. Careful and accurate CF assessment will pave the way for identifying its risk factors and manifestations and employing CF prevention and management strategies.

Evidence obtained from two decades of research and more than 350 peer-reviewed publications on CF indicated that the measurement of CF is challenging because of ambiguity and vicarious overlap with other surrogates such as secondary traumatic stress, vicarious traumatization, and burnout. Some scholars believe that the present tools used for measuring CF are not capable of assessing the intended construct and lack construct validity (Bride et al., 2007; Sinclair, Raffin-Bouchal, Venturato, Mijovic-Kondejewski, & Smith-MacDonald, 2017; Van Mol, Kompanje, Bakker, & Nijkamp, 2014). Van Mol et al. (2014) also postulated that there is no instrument available to measure the concept of CF as the loss of care ability. Sinclair et al. (2017) believe that CF measurement tools are based on experiences of social workers, psychologists, and counselors excluding the nurses’ experiences; so, they may not be generalized to nursing contexts. Hence, despite much research on CF in different disciplines, the shortage of valid and reliable instruments for this purpose is quite obvious and requires an operationalized definition of CF.

The aim of this study was to develop and evaluate the psychometric properties of a newly developed tool called the Nurses’ Compassion Fatigue Inventory (NCFI) in Iranian sociocultural context.

BACKGROUND AND CONCEPTUAL FRAMEWORK

There are many ambiguities in the area of the CF concept because of great number of similar concepts and lack of valid and reliable CF measurement tools. Thus, further investigations are necessary for identifying and measuring the concept (Watts & Robertson, 2015). On the other hand, there is no clear and comprehensive definition for CF. In an attempt to reduce ambiguities surrounding CF, Coetzee and Klopper (2010) analyzed the concept and defined it as “the final result of a cumulative and progressive process that is caused by prolonged, continuous, and intense contact with patients, the use of self, and exposure to stress” (p. 237).

Despite the need for valid and reliable tools to identify and measure phenomena, to the best of our knowledge, there is no specific, precise, and comprehensive tool for CF

measurement among nurses. Of course, over the last two decades, many researchers attempted to develop tools for measuring CF and its associated concepts, none of them have been developed based on nurses' experiences, and thus, their appropriateness for nurse populations is questionable (Bride, Robinson, Yegidis, & Figley, 2004).

Compassion Fatigue Self-Test (CFST) was developed by Figley (1995) based on clinical experiences and to measure both CF and burnout. Figley and Stamm (1996) further developed the CFST and added to it the Compassion Satisfaction subscale. Then, Stamm (2005) worked on the Satisfaction and Compassion Fatigue Test (SCFT) and developed its newer version, named Professional Quality of Life Scale (Pro-QOL), which comprises three subscales, namely, Compassion Satisfaction, Burnout, and Compassion Fatigue and Secondary Traumatic Stress (Bride et al., 2007; Stamm, 2010; Watts & Robertson, 2015). Another tool for measuring CF is the Secondary Traumatic Stress Test (STST) which contains the three subscales (Intrusion, Avoidance, and Arousal) and the primary target population was social workers. Other CF measurement tools include the 30-item Compassion Fatigue Scale, the results on its validity and reliability have not been yet published, and the 13-item Short Compassion Fatigue Scale, which has been used infrequently (Bride et al., 2007).

There are some controversies about the mentioned instruments such as equating CF and secondary traumatic stress in the Pro-QOL provoked considerable debates over their differentiation (J. Berger, Polivka, Smoot, & Owens, 2015). Moreover, the Pro-QOL and STST were considered as more screening test than a diagnostic tool and cannot be substituted for measuring the level of CF among nurses. In addition, these tools measure only some certain aspects of CF (Bride et al., 2007). Another limitation of the existing CF measurement tools is that none of them have been developed based on nurses' experiences, and their appropriateness for nurse populations is questionable (R. Berger & Gelkopf, 2011). Thus, developing a valid and reliable CF measurement tool based on nurses' firsthand experiences proves obviously crucial.

METHODS

This research is a methodological study as a part of larger sequential exploratory mixed-method research, was done in 2015. The three phases of the study were concept identification, item generation, and psychometric evaluation (LoBiondo-Wood & Haber, 2014). These phases are shown in Figure 1 and explained in detail in the following text.

Phase 1: Concept Identification Through the Hybrid Concept Analysis Model

Given the controversies over CF, we employed the hybrid concept analysis approach proposed by Schwartz-Barcott and Kim (2000) to analyze the concept of CF. This approach consists of three steps, namely, theoretical, field work, and final analysis steps (Schwartz-Barcott & Kim, 2000). In theoretical phase, we extensively reviewed the documents published in 1992–2015 on CF and CF measurement tools as selection criteria. Accordingly, the ProQuest, ScienceDirect, Google Scholar, PubMed, and other valid electronic databases were searched using keywords such as *compassion fatigue*, *secondary traumatic stress*, *vicarious traumatization*, *professional burnout*, *validation studies*, and *nursing*.

Overall, 2,781 articles were obtained. After removal of the repetitious and non-English articles, 1,180 papers remained. Again, after reviewing the titles and abstracts of the

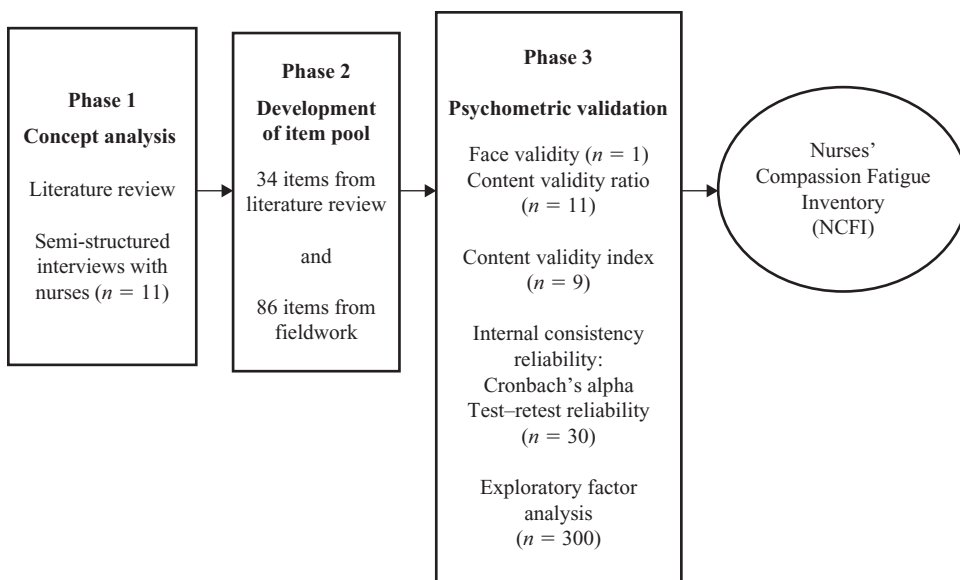


Figure 1. The phases of the study.

articles, 180 articles met the inclusion criteria, and finally, 55 relevant papers were selected after analysis.

The second step, that is, the field work, was made concurrently with the theoretical phase. A purposeful sample of 11 nurses who aged 27–40 years was recruited with maximum variation (gender, years of work experiences, working in different wards) as key informants. Semi-structured interviews were made to explore their CF-related experiences. Interview questions, which had been developed by research team, were the following: “How is your interaction with the patients in the workplace and of what quality is it?” “Have you ever got tired of making a deep relation or sympathy with the patient? Why? Please explain more.” The length of the interviews was 30–45 min. Data collection was stopped when no new information was obtained. Research team analyzed the data by using qualitative content analysis approach in several group sessions to reach the consensus (Graneheim & Lundman, 2004).

Trustworthiness. To increase credibility, the researchers had sufficient close interaction with the nurses as participants. Other measures of increasing credibility included selection of diverse samples, choice of the best meaning units, categories, themes, and demonstration of their best quotations from the interview transcriptions. To increase dependability, use was made of revision of extracted codes by member checks and external checks or auditing. The researchers tried to increase confirmability through keeping all the documentations at all stages of the study and providing detailed reports. Furthermore, they increased transferability via considering the highest diversity in sample selection (Elo & Kyngäs, 2008; Graneheim & Lundman, 2004).

Phase 2: Item Generation

In this phase, based on the categories and subcategories generated in the previous phase, a comprehensive item pool was generated. Because the formal and common language of

TABLE 1. Sample of Item Generation in Theoretical Phase

Category	Subcategory	Example of Code	Example of Meaning Unit in Theoretical Phase
Integrity treat	Physical problem	Sleep disturbance	Stamm (2010) described compassion fatigue/secondary traumatic stress as being associated with “secondary exposure to extremely stressful events” (p. 12) and including fear, sleep difficulties, or avoidance (Yoder, 2010).

The developed item:

I suffer from sleep disorders (such as sleeplessness, difficulty in falling asleep, etc.).

Iranian people is Persian, the instrument items were written in Persian (Farsi). The example of item generation in theoretical and fieldwork phase is shown in Tables 1 and 2, respectively.

Phase 3: Psychometric Evaluation

In this phase, we assessed the psychometric properties of the NCFI, that is, its face, content, and construct validity as well as reliability.

Face Validity Assessment. We assessed the face validity of the NCFI both qualitatively and quantitatively. For this purpose, 10 nurses working in hospitals assess the instrument for clarity of words in each item (qualitative) and evaluate the importance of each item by

TABLE 2. Sample of Item Generation in Fieldwork Phase

Category	Subcategory	Examples of Codes	Example of Meaning Unit in Fieldwork Phase
Hard conditions at workplace	Exposure to patients in pain, suffering, and distress	Care for refractory patients	In the ICU, I always have patients who have undergone major surgery or have sustained multiple traumata. They may be critically ill or have poor prognosis. All these make me physically and mentally exhausted.
		Care for critically ill patients	
		Care for poor prognosis patients	
		Psychosomatic fatigue because of	
		Care for patients in difficult conditions	

The developed item:

Providing care to patients with end-stage and poor prognosis conditions makes me emotionally fatigue.

Note. ICU = intensive care unit.

using a 5-point Likert-type scale (quantitative) to calculate the impact score of each item (Polit, Beck, & Owen, 2007). During qualitative assessment of face validity, the NCFI was given to 10 hospital nurses, and they were asked to identify and resolve ambiguities in the wordings of the items and difficulties in understanding them. After making amendments proposed by the nurses, we performed quantitative face validity assessment by calculating impact score. In this step, 10 hospital nurses were invited to evaluate the importance of each item by using a 5-point Likert-type scale as follows: 5 = *completely important*, 4 = *fairly important*, 3 = *moderately important*, 2 = *relatively important*, and 1 = *not important*. The impact score of each item was calculated by multiplying the relative frequency of nurses who had scored that Item 4 or 5 by the mean importance score of that item. An impact score of greater than 1.5 was considered as appropriate (Broder, Mcgrath, & Cisneros, 2007). Finally, the items were edited by research team.

Content Validity Assessment. Assessment of content validity is the most important step in assessing the validity of a scale. In this step, content validity ratio (CVR) and content validity index (CVI) were calculated. For CVR calculation, 15 experts in nursing, instrument development, and psychology were invited to rate the necessity of items. Nine out of 15 experts responded. Based on the Lawshe's (1975) table of minimum values for CVI, when the number of raters is nine, items with a CVR greater than .78 are appropriate (Ayre & Scally, 2014; Lawshe, 1975). On the other hand, for calculating CVI, 10 experts were invited to determine the relevancy of the items to the construct of CF. Then, item-level content validity index (I-CVI) of each item was calculated through dividing the number of experts who had scored that Item 3 or 4 by the total number of the experts, that is, 10. Items with an I-CVI of .79 or more, .70–.79, and less than .69 were respectively included, revised, and excluded (Hyrkäs, Appelqvist-Schmidlechner, & Oksa, 2003). Scale-level content validity index (S-CVI) was also calculated by using the S-CVI average (S-CVI/Ave) technique. Accordingly, the CVIs of all items were summed and divided by the number of items. Polit and Beck (2010) recommended that S-CVI/Ave values of greater than .90 are acceptable for establishing great content validity.

Construct Validity Assessment. Exploratory factor analysis and convergent validity approaches were used for construct validity assessment. According to Williams, Onsmann, and Brown (2010), at least 300 cases are needed for factor analysis. Using convenience sampling, 350 hospital nurses were recruited from hospitals affiliated to Shahid Beheshti and Kashan Universities of Medical Sciences, in Tehran and Kashan cities in Iran. Inclusion criteria were nurses with bachelor or higher degrees in nursing and a minimum professional experience of 1 year in the hospitals as a clinical nurse. Participants were asked to fill out the NCFI on the self-report basis. Nurses who incompletely filled out the questionnaire were excluded. The factor structure of the questionnaire was extracted by running principal component factor analysis with varimax rotation. Sampling adequacy was ensured through the Kaiser-Meyer-Olkin (KMO) test, whereas appropriateness of the analysis model was assessed by the Bartlett's test of sphericity. A minimum factor loading value of .40 was used to maintain the items in the extracted factors, whereas the scree plot and eigenvalues of greater than 1 were used to determine the number of the NCFI factors (Williams et al., 2010).

For assessing convergent validity, the Copenhagen Burnout Inventory was selected which its validity and reliability were already assessed ($\alpha = .86$) by Yeh, Cheng, Chen, Hu, & Kristensen (2007) and ($\alpha = .89$) by Lin and Lin (2013). This inventory was also filled out by 150 hospital nurses. Pearson correlation analysis was performed to examine the correlation between the scores of the two scales.

Reliability Assessment. The reliability of the NCFI was assessed by using the internal consistency and the temporal stability assessment techniques. The results of internal consistency assessment are reported as Cronbach's alpha. An alpha of greater than .70 was considered as acceptable internal consistency (Bland & Altman, 1997). In this study, Cronbach's alpha was calculated both before and after factor analysis by using the data retrieved from 50 nurses. Moreover, stability assessment was performed through the test–retest technique in which 30 nurses completed the NCFI twice with a 2-week interval. The correlation between the test and the retest scores was examined by using the intra-class correlation coefficient (ICC). ICCs of greater than .75 are acceptable (Marx, Menezes, Horovitz, Jones, & Warren, 2003). On the other hand, item analysis was performed to calculate the correlation of the items with the scale. In ideal conditions, the correlation of each item with the scale should be high and the correlation of the items with each other should be low. Consequently, when the correlation coefficient of two items is greater than .70, one of the items should be excluded. Moreover, when the correlation coefficient of an item with the scale is less than .30, that item should be excluded (Jones et al., 2009). Besides test–retest stability assessment, standard error of measurement (SEM) was also calculated for absolute reliability.

SEM indicates whether the difference in various measurements is a real one or a measurement error. The smaller the SEM, the more reliable the instrument. Comparing SEMs between tools is useful methods for comparing between measurement tools (Atkinson & Nevill, 1998). SEM is a measure of absolute reliability that measures the reliability of the scores of each person in different situations (Overend, Anderson, Sawant, Perryman, & Locking-Cusolito, 2010). It is calculated via the following formula: $SEM = SD\sqrt{1 - r^2}$, where SD is the standard deviation of test scores and r is reliability coefficient (coefficient alpha or ICC; Overend et al., 2010)

Statistical Analysis

All statistical analyses mentioned earlier were performed in SPSS (Version 16.0) at a significance level of less than .05.

Ethical Considerations

The Ethics Committee of Shahid Beheshti University of Medical Sciences, Tehran, Iran, approved the study (with the code of SBMU2.REC.1394.46). The participants were informed about the aims and the importance of the study and ensured about the confidential handling of the study data and the voluntary nature of participation in or withdrawal from the study.

The questionnaires were distributed among the nurses by first author (MS) and were handed back to her after completion. No time constraint was imposed on the nurses for filling out the NCFI even though around 15 min sufficed.

RESULTS

In total, the pool contained 120 items, 34 items of which were extracted from the theoretical step and 86 items from the fieldwork step of the first phase.

The primary item pool were read and revised several times. Then, in a joint meeting among the authors, the items were discussed, and 22 of them were excluded because of

overlaps with some other items. Finally, the 98-item draft of the NCFI was used for psychometric evaluation.

The Results of Face Validity Assessment

During qualitative face validity assessment, nurses noted that eight items were ambiguous and hard to understand. Thus, these items were excluded. For example, majority of nurse believed item of “I have special compassion with patients with specific illness” is ambiguous, and thus, it was deleted.

Moreover, in quantitative face validity assessment, 12 items obtained an impact score of less than 1.5, from which 9 items were excluded and 3 items were revised. Finally, the 81-item questionnaire was edited and reworded by two editing specialists.

The Results of Content Validity Assessment

In this step, 29 items were excluded because of obtaining a CVR of less than .78. Besides, 11 items with an I-CVI of less than .70 were excluded, and 2 items with an I-CVI of .70–.79 were revised. The I-CVI of the revised items were .833 and .916. On the other hand, S-CVI was .90. At the end of this step, the NCFI contained 41 items.

The Results of Reliability Assessment Before Construct Validity Assessment

The Cronbach’s alpha of the 41-item NCFI was .95. Item analysis revealed that six pairs of items each had a high correlation with each other. Thus, one item of each pair (i.e., the Items 30, 34–36, 40, and 42) was excluded. Finally, 35 items remained in the questionnaire. The 35-item NCFI was then subjected to construct validity assessment.

The Results of Construct Validity Assessment

In total, 325 questionnaires were completed (response rate = 92.8%). The characteristics of the nurses are shown in Table 3. The KMO test value was .942, implying that the sample was adequate. Moreover, the result of the Bartlett’s test of sphericity was statistically significant (chi-square value = 7,662.72 and $p < .001$), denoting that the factor analysis model was appropriate. The results of principal component factor analysis with varimax rotation illustrated a six-factor structure for the 35-item NCFI which explained 66.247% of the total variance of CF. The placement of some items was changed because of their greater congruence with the newly allocated factors. For instance, Items 32, 41, 43, and 46 from Factors 3, 1, 2, and 6 were respectively allocated to Factors 1, 6, 6, and 2.

The first factor contained nine items, explained 14.432% of the total variance, and was labeled “limited personal capabilities.” The second factor was labeled “caring infirmity,” consisted of six items, and explained 13.242% of the total variance. The third factor was labeled “psychosomatic disorders,” comprised four items, and was responsible for 10.945% of the total variance. The fourth factor was the six-item “emotional fatigue” factor that explained 10.870% of the total variance. The fifth factor also contained six items, explained 9.453% of the variance, and was labeled “social isolation.” Finally, the sixth factor which was labeled “incompetence in self and family management” consisted of four items and was responsible for 7.305% of the variance. Table 4 shows the factor structure of the NCFI, the factor loadings of the items, and the amount of variance explained by each factor.

TABLE 3. The Participants' Demographic Characteristics

Variables	Mean \pm SD OR Frequency
Age (Years)	21–54 \pm 6.59
Gender	
Female	257 (81.3%)
Male	59 (18.7%)
Marital status	
Married	206 (65.19%)
Single	110 (34.81%)
Employment status	
Provisional official	112 (35.45%)
Permanent official	107 (33.86%)
Post-graduation mandatory service	50 (15.82%)
Under contract	47 (14.87%)
Educational status	
Bachelor's	292 (92.4%)
Master's	24 (7.6%)
Working ward	
Intensive care units	120 (38%)
Internal medicine care	110 (34.8%)
Surgical care	44 (13.9%)
Emergency department	42 (13.3)
Total	316 (100%)

The Results of Reliability Assessment After Construct Validity Assessment

Cronbach's alpha values of the NCFI was .94, and its six factors were .90, .90, .86, .87, .61, and .80, respectively. Moreover, test–retest ICC was .935, indicating acceptable stability of the NCFI. Also, SEM of NCFI was 7.8. The results of which are shown in Table 4.

The Results of Convergent Validity Assessment

To estimate the convergent validity, given the normal distribution of the data, Pearson correlation coefficient was used to calculate the correlation between NCFI scores and Copenhagen Burnout Inventory. The results suggested a moderate correlation between the scores of the two instruments ($p = .002$, $r = .6$).

DISCUSSION

The aim of this study was to develop and evaluate the psychometric properties of a newly developed tool called the *Nurses' Compassion Fatigue Inventory* (NCFI). The

TABLE 4. Factor Structure and the Reliability Information of the NCFI

Item	Factor 1 (9 items)	Factor 2 (6 items)	Factor 3 (4 items)	Factor 4 (6 items)	Factor 5 (6 items)	Factor 6 (4 items)
My self-confidence has decreased (S38).	.705					
I have become depressed (S31).	.665		.432			
I reprimand myself if a patient does not make recovery (S33).	.663					
My ability to cope with problems has decreased (37).	.655	.473				
I am unable to establish appropriate relationships with other people (namely patients, colleagues, and family members) (S39).	.628	.486				
I do not feel happy (S23).	.578					
I am no longer interested in my previous favorable activities (S24).	.564					
I feel that my efficiency in care provision has decreased (S22).	.513					
I have become irritable to the events of daily life (S32).	.472					
My empathy with patients has decreased (S28).		.790				
My motivation for patient care has decreased (S27).		.775				
I am impatient in care delivery (S26).		.755				
I have no desire for helping others (S29).		.738				
I am unable to emotionally support my patients (S46).		.701				
I have become indifferent to my work (S25).		.671	.401			
I suffer from sleep disorders (such as sleeplessness, difficulty in falling asleep, etc.) (S19).			.784			

I suffer from eating disorders (such as polyphagia or anorexia) (S20).	.682	
I feel too tired (S16).	.680	
I have become exhausted (S17).	.668	
I am tired of care provision to critically-ill patients (S3).		.751
Over time, patient care causes me greater levels of mental fatigue (S2).		.734
Providing care to patients with end-stage and poor prognosis conditions makes me emotionally fatigue (S5).		.731
I feel emotionally fatigue due to long-term contact with patients and illnesses.		.702
I feel unhappy that my job obliges me to witness the death of children or young people (S4).		.687
Futile patient care (i.e. a lengthy care with no recovery) makes me emotionally fatigue (S6).		.669
I do not get involved in patients' affairs (S14).		.713
I avoid intimate relationships with colleagues (S12).		.681
I avoid intimate relationships with patients (S11).		.652
I want to be alone (S13).		.622
I do not like to think about others' problems (including patients, friends, and family members) (S7).		.581
I am not interested in wide circles (S15).		.518

(Continued)

TABLE 4. Factor Structure and the Reliability Information of the NCFI (Continued)

Item	Factor 1 (9 items)	Factor 2 (6 items)	Factor 3 (4 items)	Factor 4 (6 items)	Factor 5 (6 items)	Factor 6 (4 items)
I am unable to emotionally support my family members (including parents, children, spouse, etc.) (S45).						.680
I show extreme reactions (either indifference or oversensitivity) to the illnesses of my family members (S44).						.657
My private life has been affected (S43).						.531
I am unable to manage my life conditions (S41).						.486
Adjusted variance	14.43%	13.24%	10.945%	10.870%	9.435%	7.305%
Cumulative variance	14.43%	27.673%	38.619%	49.488%	58.942%	66.247%
Cronbach's alpha	0.90	0.90	0.86	0.87	0.61	0.8
ICC	0.89	0.90	0.84	0.84	0.59	0.78
SEM	3.12	1.74	1.83	2.55	3.25	1.93
Mean \pm Standard deviation	20.06(6.81)	9.55(3.99)	9.06(3.35)	18.95(4.67)	15.21(4.004)	8.6(3.06)

questionnaire was developed based on the contextual definition of CF and the sociocultural context of Iran. Nonetheless, it can be used in other countries after its validation and cross-cultural adaptation according to the immediate setting. The NCFI is an easily applicable, valid, and reliable 35-item self-report questionnaire that can be used by researchers, managers, and nurses.

Study findings revealed that the NCFI has high internal consistency. All NCFI items were strongly correlated with the whole questionnaire, indicating that the questionnaire is a homogenous scale and all its items measure the same construct. The findings also showed that the 35-item NCFI has acceptable reliability and face, content, and construct validity. Exploratory factor analysis revealed a six-factor structure for the questionnaire which explained 66.24% of the total variance of CF. The six factors of the NCFI were “limited personal capabilities,” “caring infirmity,” “psychosomatic disorders,” “emotional fatigue,” “social isolation,” and “incompetence in self and family management.” In agreement with the findings of the previous studies (Coetzee & Klopper, 2010; Owen & Wanzer, 2014), our findings also reflected that CF is a multidimensional concept. However, contrary to the NCFI, most of the existing CF measurement tools comprised only two subscales, that is, Compassion Fatigue and Burnout (Hunsaker, Chen, Maughan, & Heaston, 2015). The Pro-QOL also includes Compassion Fatigue, Burnout, and Compassion Satisfaction subscales (Stamm, 2010). Besides, the STST contains the three subscales of Intrusion, Avoidance, and Arousal (Bride et al., 2007). The items of the STST Avoidance subscale are to some extent compatible with the items of the NCFI Social Isolation subscale. Because of such limitations of the existing CF measurement tools, multiple tools should be used in a single study to provide better understanding of CF (Bride et al., 2007). Using multiple tools is too much time-consuming and not cost-effective. On the other hand, NCFI is a comprehensive multidimensional instrument that can cover most aspects of CF.

One of the strength of the NCFI is that it was generated through both deductive (i.e., reviewing the existing literature) and inductive (i.e., interviewing hospital nurses) methods. An important point is that CF is a multidimensional concept that is influenced by the culture, religion, and predominant cultural values of the immediate context. Thus, deductive–inductive concept analysis is the right approach for exploring CF and developing CF measurement tools. According to Walker and Avant (2005), the items of a scale which is developed through concept analysis can show whether the antecedents of the concept have happened or not. Also, according to Nunnally (1978), concept analysis enables researchers to develop the best and the most comprehensive scale for measuring the intended concept. Contrary to the NCFI, neither Figley’s CF model (Figley, 2002a; Ledoux, 2015) nor the existing CF measurement tools were developed based on nurses’ experiences. Thus, they cannot properly describe nurses’ CF and its contributing factors (Meadors, Lamson, Swanson, White, & Sira, 2009; Sheppard, 2015). In other words, measuring nurses’ CF by the existing tools may not provide credible information about nurses’ CF.

Our findings also revealed a high correlation between the scores of the NCFI and the Copenhagen Burnout Inventory. In line with previous studies (Jenkin & Warren, 2012; Makic, 2015; Watts & Robertson, 2015), this finding denotes the relationship of CF with burnout. Frank and Adkinson (2007) also described CF as the unique form of burnout which affects health care providers. Other scholars also noted that despite clear distinction between CF and burnout, these concepts bear similarities with each other (Potter et al., 2010; Yoder, 2010).

The findings of this study and the use of the NCFI can provide hospital nurses and nursing managers with better insight into the concept of CF and help them adopt effective measures (such as self-care programs) to prevent CF in health care settings. CF prevention

can enhance nurses' job satisfaction, promote their retention, improve care quality, and enhance patient satisfaction and safety.

This study faced several limitations. For example, there was no appropriate culturally adapted tool for measuring Iranian nurses' CF, and thus, we could not assess the concurrent validity of the NCFI. Besides, as the evaluation and validation of an instrument is a lengthy and continuous process, further attempts are still needed for revising the NCFI. The authors hope that they overcome potential shortcomings of the NCFI in future studies. Further studies with large samples are recommended for revising and validating the NCFI. Moreover, interventional studies are needed to prevent and manage nurses' CF.

Description, Administration, and Scoring of the Nurses' Compassion Fatigue Inventory

NCFI is a simple, valid paper-and-pencil self-report 35-item scale that can be easily used by managers, nurses, and researchers. These items are scored on a 5-point Likert-type response scale as follows: 1 = *never*, 2 = *rarely*, 3 = *sometimes*, 4 = *often*, 5 = *always*. The total score of the NCFI is the sum of the item scores and can range from 35 to 175; the higher the score, the severer the CF.

In the scoring phase, using Likert method of Mathematical Logic 3, the test scores were transformed into three classes: low, average, and high (low = 35–81.6, moderate = 81.6–128.2, high = 128.2–175; category interval = 46.6). The amount of time needed for completing the NCFI is as low as 7–15 min.

CONCLUSION

The NCFI is a simple self-report questionnaire that can be easily used by researchers, managers, and nurses. NCFI is developed through hybrid concept analysis (based on reviewing the literature and interviewing hospital nurses) which based on Iranian sociocultural context. NCFI has acceptable reliability and face, content, and construct validity with a six-factor structure. Although the NCFI is a new instrument, it seems to be a useful measure to assess CF.

The Nursing Implication for Practice, Research, and Education

In Nursing Practice. Managers can determine the rate of CF in nurses in organization and promote the personnel's skills in adaptation strategies, coping, and self-care strategies. In this way, they may help to maintain and foster the physical and mental well-being and health of nurses. Moreover, managers may avoid nurses' work leave and nursing force turnover and shortage as a great challenge of the health care systems through creating favorable working environments in clinical nursing setting. On the other hand, nurses can get aware of CF by the use of NCFI and avoid its destructive impacts on themselves, the patients, and the organization via early prompt diagnosis and treatment. In this way, they will maintain their CF ability forever.

In Nursing Education. NCFI increases the awareness and perception of managers, instructors, and personnel about CF. Regarding the ignorance toward this phenomenon, training on CF, its signs and symptoms, and prophylactic methods and its management should be added to the bachelor of science (BS) and master of science (MSc) curriculum to

avoid CF in young nurses and prevent work leave or sick leave leading to the final resolution of nurse shortage. The nursing managers should prioritize the development and implementation of pedagogic programs to promote adaptation skills, communication skills, stress management strategies, and self-care which form the most important preventive measures of CF.

In Nursing Research. Scholars may use NCFI to assess CF in nurses in different wards or sections. In this way, they can prevent this highly debilitating phenomenon with its numerous negative outcomes for nurses, patients, and the organization through identifying the risk factors of CF. Researchers ought to design some studies to determine the efficacy of effective interventions to reduce this destructive and annihilating phenomenon to enhance patient care quality and promote the health and well-being of nurses as the main goals of the health care systems.

REFERENCES

- Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine*, *26*, 217–238.
- Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio revising the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, *47*, 79–86.
- Berger, J., Polivka, B., Smoot, E. A., & Owens, H. (2015). Compassion fatigue in pediatric nurses. *Journal of Pediatric Nursing*, *30*, e11–e17.
- Berger, R., & Gelkopf, M. (2011). An intervention for reducing secondary traumatization and improving professional self-efficacy in well baby clinic nurses following war and terror: A random control group trial. *International Journal of Nursing Studies*, *48*, 601–610.
- Bland, J. M., & Altman, D. G. (1997). Cronbach's alpha. *BMJ*, *314*, 572.
- Boyle, D. A. (2011). Countering compassion fatigue: A requisite nursing agenda. *Online Journal of Issues in Nursing*, *16*, 2.
- Bride, B. E., Radey, M., & Figley, C. R. (2007). Measuring compassion fatigue. *Clinical Social Work Journal*, *35*, 155–163.
- Bride, B. E., Robinson, M. M., Yegidis, B., & Figley, C. R. (2004). Development and validation of the Secondary Traumatic Stress Scale. *Research on Social Work Practice*, *14*, 27–35.
- Broder, H. L., Mcgrath, C., & Cisneros, G. J. (2007). Questionnaire development: Face validity and item impact testing of the Child Oral Health Impact Profile. *Community Dentistry and Oral Epidemiology*, *35*, 8–19.
- Coetzee, S. K., & Klopper, H. C. (2010). Compassion fatigue within nursing practice: A concept analysis. *Nursing & Health Sciences*, *12*, 235–243.
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, *62*, 107–115.
- Figley, C. R. (1995). Compassion fatigue: Toward a new understanding of the costs of caring. In B. H. Stamm (Ed.), *Secondary traumatic stress: Self-care issues for clinicians, researchers, and educators* (pp. 3–28). Baltimore, MD: Sidran Press.
- Figley, C. R. (2002a). Compassion fatigue: Psychotherapists' chronic lack of self care. *Journal of Clinical Psychology*, *58*, 1433–1441.
- Figley, C. R. (2002b). *Treating compassion fatigue*. New York, NY: Routledge.
- Figley, C. R., & Stamm, B. (1996). Psychometric review of compassion fatigue self-test. In B. H. Stamm (Ed.), *Measurement of stress, trauma, and adaptation* (pp. 127–130). Lutherville, MD: Sidran Press.
- Frank, D. I., & Adkinson, L. F. (2007). A developmental perspective on risk for compassion fatigue in middle-aged nurses caring for hurricane victims in Florida. *Holistic Nursing Practice*, *21*, 55–62.
- Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*, 105–112.

- Hunsaker, S., Chen, H. C., Maughan, D., & Heaston, S. (2015). Factors that influence the development of compassion fatigue, burnout, and compassion satisfaction in emergency department nurses. *Journal of Nursing Scholarship, 47*, 186–194.
- Hyrkäs, K., Appelqvist-Schmidlechner, K., & Oksa, L. (2003). Validating an instrument for clinical supervision using an expert panel. *International Journal of Nursing Studies, 40*, 619–625.
- Jenkins, B., & Warren, N. A. (2012). Concept analysis: compassion fatigue and effects upon critical care nurses. *Critical Care Nursing Quarterly, 35*(4), 388–395.
- Joinson, C. (1992). Coping with compassion fatigue. *Nursing, 22*, 116, 118–120.
- Jones, P., Harding, G., Berry, P., Wiklund, I., Chen, W., & Kline Leidy, N. (2009). Development and first validation of the COPD Assessment Test. *European Respiratory Journal, 34*, 648–654.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*, 563–575.
- Ledoux, K. (2015). Understanding compassion fatigue: Understanding compassion. *Journal of Advanced Nursing, 71*, 2041–2050.
- Lin, L. P., & Lin, L. J. (2013). Job burnout amongst the institutional caregivers working with individuals with intellectual and developmental disabilities: Utilization of the Chinese version of the Copenhagen Burnout Inventory survey. *Research in Autism Spectrum Disorders, 6*(7), 777–784.
- LoBiondo-Wood, G., & Haber, J. (2014). *Nursing research: Methods and critical appraisal for evidence-based practice* (8th ed.). St. Louis, MO: Elsevier.
- Lynch, S. H., & Lobo, M. L. (2012). Compassion fatigue in family caregivers: A Wilsonian concept analysis. *Journal of Advanced Nursing, 68*, 2125–2134.
- Makic, M. B. F. (2015). Taking Care of the Caregiver: Compassion Satisfaction and Compassion Fatigue. *Journal of Perianesthesia Nursing, 30*(6), 546–547.
- Marx, R. G., Menezes, A., Horovitz, L., Jones, E. C., & Warren, R. F. (2003). A comparison of two time intervals for test–retest reliability of health status instruments. *Journal of Clinical Epidemiology, 56*, 730–735.
- Meadors, P., Lamson, A., Swanson, M., White, M., & Sira, N. (2009). Secondary traumatization in pediatric healthcare providers: Compassion fatigue, burnout, and secondary traumatic stress. *Omega, 60*, 103–128.
- Melvin, C. S. (2012). Professional compassion fatigue: What is the true cost of nurses caring for the dying? *International Journal of Palliative Nursing, 18*, 606–611.
- Meyer, R. M., Li, A., Klaristenfeld, J., & Gold, J. I. (2015). Pediatric novice nurses: Examining compassion fatigue as a mediator between stress exposure and compassion satisfaction, burnout, and job satisfaction. *Journal of Pediatric Nursing, 30*, 174–183.
- Nunnally, J. (1978). *Psychometric methods*. New York, NY: McGraw-Hill.
- Overend, T., Anderson, C., Sawant, A., Perryman, B., & Locking-Cusolito, H. (2010). Relative and absolute reliability of physical function measures in people with end-stage renal disease. *Physiotherapy Canada, 62*, 122–128.
- Owen, R. P., & Wanzer, L. (2014). Compassion fatigue in military healthcare teams. *Archives of Psychiatric Nursing, 28*, 2–9.
- Polit, D., & Beck, C. (2010). *Essentials of nursing research: appraising evidence for nursing practice*. 2010. Lippincott Williams & Wilkins.
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*, 459–467.
- Potter, P., Deshields, T., Divanbeigi, J., Berger, J., Cipriano, D., Norris, L., & Olsen, S. (2010). Compassion fatigue and burnout: Prevalence among oncology nurses. *Clinical Journal of Oncology Nursing, 14*, E56–E62.
- Sansó, N., Galiana, L., Oliver, A., Pascual, A., Sinclair, S., & Benito, E. (2015). Palliative care professionals' inner life: Exploring the relationships among awareness, self-care, and compassion satisfaction and fatigue, burnout, and coping with death. *Journal of Pain and Symptom Management, 50*, 200–207.
- Schwartz-Barcott, D., & Kim, H. S. (2000). An expansion and elaboration of the hybrid model of concept development. In B. L. Rodgers & K. A. Knafel (Eds.), *Concept development in nursing: Foundations, techniques, and applications* (2nd ed., pp. 161–192). Philadelphia, PA: Saunders.
- Sheppard, K. (2015). Compassion fatigue among registered nurses: Connecting theory and research. *Applied Nursing Research, 28*, 57–59.

- Sinclair, S., Raffin-Bouchal, S., Venturato, L., Mijovic-Kondejewski, J., & Smith-MacDonald, L. (2017). Compassion fatigue: A meta-narrative review of the healthcare literature. *International Journal of Nursing Studies*, *69*, 9–24.
- Stamm B. H. 2005. *The ProQOL manual*. Retrieved July, 16, 2007.
- Stamm, B. H. (2010). *The concise ProQOL manual* (2nd ed.). Pocatello, ID: ProQOL.org.
- Van Mol, M., Kompanje, E., Bakker, J., & Nijkamp, M. (2014). Compassion fatigue and burnout among healthcare professionals in the ICU. *Critical Care*, *18*, P19.
- Walker, L. O., & Avant, K. C. (2005). *Strategies for theory construction in nursing* (4th ed.). Upper Saddle River, NJ: Pearson/Prentice Hall.
- Watts, J., & Robertson, N. (2015). Selecting a measure for assessing secondary trauma in nurses. *Nurse Res*, *23*, 30–35.
- Williams, B., Onsman, A., Brown, T. (2010). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, *8*(3). Retrieved from <http://ro.ecu.edu.au/jephc/vol8/iss3/1>
- Yeh, W. Y., Cheng, Y., Chen, C. J., Hu, P. Y., & Kristensen, T. S. (2007). Psychometric properties of the Chinese version of Copenhagen burnout inventory among employees in two companies in Taiwan. *International Journal of Behavioral Medicine*, *14*(3), 126–133.
- Yoder, E. A. (2010). Compassion fatigue in nurses. *Applied Nursing Research*, *23*, 191–197.

Acknowledgments. Our special thanks to the nurses who participated in the study during the process of scale evaluation. We would also like to state our thanks to the research deputy of Shahid Beheshti University of Medical Sciences for supporting the study.

The article is taken from a nursing PhD dissertation, Shahid Beheshti University of Medical Sciences, Tehran, Iran (Grant No: SBMU2.REC.1394.46).

Correspondence regarding this article should be directed to Mansoureh Zagheri Tafreshi, PhD, Nursing Management Department, School of Nursing and Midwifery, Shahid Beheshti University of Medical Sciences. E-mail: m.z.tafreshi@sbmu.ac.ir