



Normative values for the distress thermometer (DT) and the emotion thermometers (ET), derived from a German general population sample

Andreas Hinz¹ · Alex J. Mitchell² · Csaba L. Dégi³ · Anja Mehnert-Theuerkauf¹

Accepted: 24 September 2018 / Published online: 3 October 2018
© Springer Nature Switzerland AG 2018

Abstract

Purpose The distress thermometer (DT) and the emotion thermometers (ET) are short screening instruments for use in oncological practice. The aim of this study was to provide normative values and to analyze the correlational structure of the ET.

Methods A representative sample of the adult German general population ($N=2437$) completed the ET, the PHQ-4, the FACIT-fatigue scale, and the demoralization scale.

Results The percentages of people above the cutoff (≥ 4) and the mean scores of the five ET scales were as follows: distress: 39.0%, $M=3.15 \pm 2.62$, anxiety: 12.3%, $M=1.36 \pm 1.93$, depression: 16.1%, $M=1.65 \pm 2.11$, anger: 24.5%, $M=2.33 \pm 2.16$, and need for help: 10.7%, $M=1.18 \pm 1.90$. Women reported significantly higher levels of burden than men, with effect sizes between 0.07 (anger) and 0.36 (anxiety). All ET dimensions were interrelated (r between 0.44 and 0.69) and significantly correlated with the other scales (r between 0.36 and 0.68).

Conclusions The normative scores can help qualify assessments of groups of patients. The new four dimensions of the ET provide relevant additional information that is not already covered by the DT.

Keywords Distress · Screening · General population · Normative study

Introduction

Cancer patients often experience clinically significant distress, and there is a need for acceptable brief screening instruments with which this distress can be assessed [1]. One very short instrument is the NCCN distress thermometer (DT) [2–4]. It was designed to assess cancer patients' distress on a 0–10 scale. Several studies have tested the validity [5–7] and the temporal stability [8] of this instrument.

One limitation of the DT is its relative low specificity. In response to this, several authors have suggested using multi-domain screening that targets emotional responses beyond core distress. One example is the development of the

emotion thermometers (ET) [9, 10]. In addition to distress, the ET measures anxiety (AnxT), depression (DepT), anger (AngT), and the need for help (HelpT). All five dimensions are to be rated on a vertical visual analogue scale (thermometer) with the anchors 0 and 10. Several studies have been conducted to test the validity of these new thermometers and to find optimal cutoffs [11–13].

The results obtained with screening instruments in clinical settings can be better interpreted when normative values are available. The difference between the patients' mean scores and those normative scores derived from the general population indicates the burden of the underlying diseases and treatments. Age and gender differences between groups of patients can be accounted for when normative scores are considered. To our knowledge, such normative studies have not been performed so far for the DT and the ET, with the exception of a recent study that adopted the DT to survey a sample of the Dutch general population [14]. The main aim of this study was to provide normative scores derived from a large general population sample. This includes the calculation of age and gender differences. A second aim was to evaluate the degree to which the established DT is

✉ Andreas Hinz
andreas.hinz@medizin.uni-leipzig.de

¹ Department of Medical Psychology and Medical Sociology, University of Leipzig, Philipp-Rosenthal-Str. 55, 04103 Leipzig, Germany

² Department of Cancer and Molecular Medicine, University of Leicester, Leicester, UK

³ Faculty of Sociology and Social Work, Babes-Bolyai University, Cluj-Napoca, Romania

representative of other dimensions of burden as represented by the other ET areas.

Methods

Study participants

This study was part of a German household survey, performed by a demographic consulting company (USUMA, Germany) on behalf of the University of Leipzig, between March and May 2015. A representative sample of the German general population aged 14 years and above was selected. The whole country was separated into 258 areas (sample points) to represent all regions in Germany. These areas were drawn proportionally to the distribution of private households. Once a sample point was selected, streets, houses, households, and household members were chosen randomly with the random route technique. Candidates who did not possess sufficient German language skills were not included. Of the initially selected 4902 addresses, 4844 were valid. The subjects were visited by professional interviewers. Of the contacted persons, 2513 were willing to take part in the study (response rate 51.9%). The study participants were asked to fill in a set of several questionnaires on mental and physical health, including the ET within a face-to-face interview. All adult participants provided their written informed consent to participate in this study and the data to be published. Also, written informed consent from the next of kin, caretakers, or guardians on behalf of the minors/children enrolled in the study was obtained. We excluded persons who were younger than 18 from further analysis and thus arrived at a sample consisting of 2437 people. The Ethics Committee of the Medical Faculty of the University of Leipzig approved the study.

Instruments

The ET consist of five columns (thermometers) with the headings distress, anxiety, depression, anger, and help. The instruction reads as follows: In the first four columns, please mark the number (0–10) that best describes how much emotional upset you have been experiencing in the past week, including today. In the final column please indicate how much you need help for these concerns. In the first four thermometers, the anchors are labeled with “None” (0) and “Extreme” (10), and the corresponding labels for the fifth thermometer (Help) are “Can manage by myself” (0) and “Desperately” (10). Further information on the ET is available at <http://www.psych-oncology.info/ET.htm>.

In addition to the ET, several other instruments were included in the survey. Anxiety and depression were measured with the Patient Health Questionnaire-4 (PHQ-4) [15],

fatigue was assessed with the Functional Assessment of Chronic Illness Therapy Fatigue Scale (FACIT-Fatigue) [16, 17], and demoralization was measured with the demoralization scale (DS) [18, 19].

Statistical analyses

Age and gender differences were tested with two-factorial ANOVAs, using the factors age group (six categories according to Table 1) and gender (two categories). Effect sizes were calculated using Cohen's *d*. Associations between the ET scores and other scales were expressed in terms of Pearson correlations.

Results

Age and gender differences

Sociodemographic characteristics of the study sample are given in Table 1. The mean age was about 50 years, and 55.9% of the sample was female. Table 2 presents ET mean scores for males and females. Using the cutoff ≥ 4 as suggested by [9] for all ET domains, we also calculated the percentage of people with scores above this cutoff (Table 2).

Women reported higher levels of burden in all five DT domains; the effect sizes for the gender differences were as follows: DT: $d=0.16$, AnxT: $d=0.36$, DepT: $d=0.31$, AngT: $d=0.07$, and HelpT: $d=0.25$. The age effects were less clear and did not show a linear trend. The ANOVA results proved that gender had a statistically significant impact on all ET dimensions (all $p < 0.001$). Age group had a statistically significant impact on distress and need for help ($p < 0.001$), and for the other three thermometers there was no statistically significant age effect ($p > 0.05$). The interaction between gender and age group failed to reach the significance level for all ET domains.

Normative scores are given in Table 3, based on the total sample. 19.2% of the general population reported no distress at all ($DT=0$), and further 14.5% reported $DT=1$, resulting in a cumulative score of 33.7% in Table 3. Almost one in two chose the minimal score of 0 in the anxiety thermometer. According to the cutoff ≥ 4 , 39% (100% minus 61.0%) of the general population reported experiencing distress (Table 3).

Correlations among the ET dimensions and associations with the other scales

Table 4 presents the correlations among the ET dimensions (r between 0.44 and 0.69). Among the five ET domains, the DT showed the lowest correlation coefficients with the other ET dimensions (mean correlation: 0.48) and with the other three questionnaires (FACIT-fatigue, PHQ-4, and

Table 1 Sociodemographic characteristics of the sample

	Males (<i>n</i> = 1075)		Females (<i>n</i> = 1362)		Total (<i>n</i> = 2437)	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
Age mean (<i>SD</i>)	49.7	(17.3)	50.0	(17.5)	49.8	(17.4)
Age category (years)						
18–29	176	16.4	212	15.6	388	15.9
30–39	163	15.2	215	15.8	378	15.5
40–49	184	17.1	241	17.7	425	17.4
50–59	211	19.6	255	18.7	466	19.1
60–69	184	17.1	221	16.2	405	16.6
≥ 70	157	14.6	218	16.0	375	15.4
Marital status						
Married, living together	533	49.6	571	41.9	1104	45.3
Married, living apart	21	2.0	33	2.4	54	2.2
Single	331	30.8	337	24.7	668	27.4
Divorced	125	11.6	209	15.3	334	13.7
Widowed	62	5.8	206	15.1	268	11.0
Missing	3	0.3	6	0.4	9	0.4
Education (years)						
≤ 9	387	36.0	460	33.8	847	34.8
10–11	446	41.5	621	45.6	1067	43.8
≥ 12	241	22.4	278	20.4	519	21.3
Missing	1	0.1	3	0.2	4	0.2
Employment						
Working full-time	594	55.3	402	29.5	996	40.9
Working part-time	51	4.7	326	23.9	377	15.5
Unemployed	67	6.2	69	5.1	136	5.6
House wife/man	7	0.7	102	7.5	109	4.5
Retired	297	27.6	389	28.6	686	28.1
Education/training	59	5.5	74	5.4	133	5.5

SD Standard deviation

Demoralization Scale). The highest coefficients were generally found for the Depression thermometer.

Discussion

The first aim of the study was to investigate the degrees to which distress and the other ET domains are present in the general population. According to the DT cutoff ≥ 4 , a relatively high proportion of the respondents (39%) were distressed. The DT mean score was 3.15, which is very similar to the mean of the Dutch normative study ($M = 3.2$) [14]. Most of the studies performed among samples of cancer patients from the US, Ireland, Denmark, and Germany reported higher mean scores, ranging between 3.9 and 6.2 [19–23], while studies from the UK [9, 24] and Singapore [12] found lower mean levels of distress (between 1.4 and 2.9) in their patients' samples.

Of all of the ET domains, the highest mean score was obtained for distress ($M = 3.15$), followed by anger

($M = 2.33$), depression ($M = 1.65$), anxiety ($M = 1.36$), and need for help ($M = 1.18$). The study done among British cancer patients [9] found another sequence: anxiety ($M = 4.8$), distress ($M = 2.9$), depression ($M = 2.3$), and anger ($M = 2.0$). It can be concluded that the difference between cancer patients and the general population is greatest in anxiety and smallest in distress. One possible reason for the relatively high distress scores in our general population study is that healthy people may have taken into account all of their daily hassles when they reported on their distress levels, while patients potentially assumed the DT question pertaining specifically to their disease. Another possible explanation for this result is that patients are often exempt from some of the obligations that normally define their everyday lives (e.g., professional duties). When using the cutoff ≥ 4 for the DT, one should take into account that this is a low criterion since such a degree of distress is also common in the healthy population.

Women reported significantly higher levels of all of the ET dimensions than men, while age differences were less

Table 3 Normative scores (cumulative distribution functions, in %) for the ET dimensions

Score	Distress	Anxiety	Depression	Anger	Help
0	19.2	48.6	41.5	24.4	56.3
1	33.7	67.5	61.6	42.0	72.9
2	48.6	80.6	75.2	61.1	82.8
3	61.0	87.7	83.9	75.5	89.3
4	69.6	91.2	88.9	82.8	92.2
5	79.3	94.5	92.9	90.7	94.6
6	85.6	96.6	95.2	94.1	96.9
7	92.1	98.1	97.3	97.4	98.3
8	97.4	99.6	98.9	99.0	99.1
9	99.1	99.8	99.7	99.6	99.7
10	100.0	100.0	100.0	100.0	100.0

The line between scores 3 and 4 indicates the cutoff ≥ 4

Table 4 Correlations between the ET dimensions and other scales

	Distress	Anxiety	Depression	Anger	Help
Distress	–	.49	.50	.48	.44
Anxiety		–	.69	.50	.68
Depression			–	.57	.67
Anger				–	.47
Help					–
FACIT-fatigue	–.42	–.58	–.63	–.45	–.63
PHQ-4 anxiety	.40	.63	.62	.46	.57
PHQ-4 depression	.36	.52	.64	.41	.54
PHQ-4 total	.41	.62	.68	.47	.60
Demoralization scale	.36	.58	.68	.47	.59

All correlations: $p < 0.001$. FACIT-Fatigue: high scores indicate low levels of fatigue

pronounced. When comparing samples of patients concerning their emotional responses, gender has to be taken into account. It is well known that anxiety is higher in women compared to men [25], but higher scores for women were also found in all of the other domains. Whenever ET mean scores of groups of patients with different gender distributions are compared one with another, the normative scores can be used to quantify the degree to which this difference is due to general gender differences.

The second aim of this investigation was to analyze whether the DT can be considered a generalized measure of patients' burden. The domains of the ET were significantly correlated, with coefficients between 0.44 and 0.69. However, on average, the correlations between the DT and the other ET scales and supplementary scales were lower than those of the other ET dimensions. Although the three instruments used here (FACIT-Fatigue, PHQ-4, and Demoralization Scale) are not at all representative of all domains of

mental problems, the relatively low correlations of the DT indicate that the DT should not be considered a generalized score representing all other thermometers. There is a considerable gain in information when the DT is complemented with the other domains.

Short scales or even one-item scales are generally assumed to be useful for research questions on group-level, while larger scales are preferred for individual decision-making [26]. The DT and ET, however, have been designed as short screening instruments which assist clinicians in identifying patients who have a special need for support. Sensitivity and specificity analyses supported the appropriateness of the ET for this purpose [9, 13]. The five items of the ET have not been designed as items of a 5-item scale. Nevertheless, there is a certain overlap of the ET domains, and combinations of the ET domains have been investigated concerning their power in predicting mental health outcomes [10].

Some limitations of this study should be mentioned. Despite the representativeness of the sample in terms of age and gender, we have no information about the distress level of the non-respondents. Therefore, a certain bias is possible. We used a German sample, and the generalizability to other countries is difficult to assess. Further studies done in other countries are necessary for assessing generalizability. The percentages presented in the normative table only give a summarizing overview of the distribution of the scores since they are aggregated across gender and age groups.

Taken together, the multi-domain extension of the original DT, the ET, provides essential information, while nevertheless remaining an instrument that is short and easy to complete. Although normative scores have already been performed for other ultra-short screening instruments, this study provides ET normative scores that can be used in clinical routine and research.

Compliance with ethical standards

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

Ethical approval The study was approved by the Ethics Committee of the University of Leipzig. All procedures were in accordance with the ethical standards of the national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all participants.

References

- Mitchell, A. J. (2010). Short screening tools for cancer-related distress: A review and diagnostic validity meta-analysis. *Journal of the National Comprehensive Cancer Network*, 8(4), 487–494.

2. Mitchell, A. J. (2007). Pooled results from 38 analyses of the accuracy of distress thermometer and other ultra-short methods of detecting cancer-related mood disorders. *Journal of Clinical Oncology*, 25(29), 4670–4681.
3. National Comprehensive Cancer Network (NCCN). (2003). Distress management. Clinical practice guidelines. *Journal of the National Comprehensive Cancer Network*, 1(3), 344–374.
4. Roth, A. J., Kornblith, A. B., Batel-Copel, L., Peabody, E., Scher, H. I., & Holland, J. C. (1998). Rapid screening for psychologic distress in men with prostate carcinoma. *Cancer*, 82(10), 1904–1908.
5. Recklitis, C. J., Blackmon, J. E., & Chang, G. (2016). Screening young adult cancer survivors for distress with the distress thermometer: Comparisons with a structured clinical diagnostic interview. *Cancer*, 122(2), 296–303.
6. Ma, X., Zhang, J., Zhong, W., Shu, C., Wang, F., Wen, J., et al. (2014). The diagnostic role of a short screening tool—The distress thermometer: A meta-analysis. *Supportive Care in Cancer*, 22(7), 1741–1755.
7. Donovan, K. A., Grassi, L., McGinty, H. L., & Jacobsen, P. B. (2014). Validation of the distress thermometer worldwide: State of the science. *Psycho-Oncology*, 23(3), 241–250.
8. Leclair, T., Carret, A.-S., Samson, Y., & Sultan, S. (2016). Stability and repeatability of the distress thermometer (DT) and the Edmonton Symptom Assessment System-Revised (ESAS-r) with parents of childhood cancer survivors. *PLoS ONE*, 11(7), e0159773.
9. Mitchell, A. J., Baker-Glenn, E. A., Park, B., Granger, L., & Symonds, P. (2010). Can the distress thermometer be improved by additional mood domains? Part II. What is the optimal combination of emotion thermometers? *Psycho-Oncology*, 19(2), 134–140.
10. Mitchell, A. J., Baker-Glenn, E. A., Granger, L., & Symonds, P. (2010). Can the distress thermometer be improved by additional mood domains? Part I. Initial validation of the emotion thermometers tool. *Psycho-Oncology*, 19(2), 125–133.
11. Mitchell, A. J., Morgan, J. P., Petersen, D., Fabbri, S., Fayard, C., Stoletniy, L., et al. (2012). Validation of simple visual-analogue thermometer screen for mood complications of cardiovascular disease: The emotion thermometers. *Journal of Affective Disorders*, 136(3), 1257–1263.
12. Beck, K. R., Tan, S. M., Lum, S. S., Lim, L. E., & Krishna, L. K. (2016). Validation of the emotion thermometers and hospital anxiety and depression scales in Singapore: Screening cancer patients for distress, anxiety and depression. *Asia-Pacific Journal of Clinical Oncology*, 12(2), 9.
13. Schubart, J. R., Mitchell, A. J., Dietrich, L., & Gusani, N. J. (2015). Accuracy of the emotion thermometers (ET) screening tool in patients undergoing surgery for upper gastrointestinal malignancies. *Journal of Psychosocial Oncology*, 33(1), 1–14.
14. van Oers, H. A., Schepers, S. A., Grootenhuis, M. A., & Haverman, L. (2017). Dutch normative data and psychometric properties for the distress thermometer for parents. *Quality of Life Research*, 26(1), 177–182.
15. Löwe, B., Wahl, I., Rose, M., Spitzer, C., Glaesmer, H., Wingenfeld, K., et al. (2010). A 4-item measure of depression and anxiety: Validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *Journal of Affective Disorders*, 122(1–2), 86–95.
16. Yellen, S. B., Cella, D. F., Webster, K., Blendowski, C., & Kaplan, E. (1997). Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. *Journal of Pain and Symptom Management*, 13(2), 63–74.
17. Cella, D., Lai, J.-S., Chang, C.-H., Peterman, A., & Slavin, M. (2002). Fatigue in cancer patients compared with fatigue in the general United States population. *Cancer*, 94(2), 528–538.
18. Mehnert, A., Vehling, S., Hocker, A., Lehmann, C., & Koch, U. (2011). Demoralization and depression in patients with advanced cancer: Validation of the German version of the demoralization scale. *Journal of Pain and Symptom Management*, 42(5), 768–776.
19. Kissane, D. W., Wein, S., Love, A., Lee, X. Q., Kee, P. L., & Clarke, D. M. (2004). The demoralization scale: A report of its development and preliminary validation. *Journal of Palliative Care*, 20(4), 269–276.
20. Faller, H., Weis, J., Koch, U., Brähler, E., Härter, M., Keller, M., et al. (2017). Utilization of professional psychological care in a large German sample of cancer patients. *Psycho-Oncology*, 26(4), 537–543.
21. Bidstrup, P. E., Mertz, B. G., Dalton, S. O., Deltour, I., Kroman, N., Kehlet, H., et al. (2012). Accuracy of the Danish version of the ‘distress thermometer’. *Psycho-Oncology*, 21(4), 436–443.
22. Lazenby, M., Dixon, J., Bai, M., & McCorkle, R. (2014). Comparing the distress thermometer (DT) with the patient health questionnaire (PHQ)-2 for screening for possible cases of depression among patients newly diagnosed with advanced cancer. *Palliative & Supportive Care*, 12(1), 63–68.
23. Ryan, D. A., Gallagher, P., Wright, S., & Cassidy, E. M. (2012). Sensitivity and specificity of the distress thermometer and a two-item depression screen (Patient Health Questionnaire-2) with a ‘help’ question for psychological distress and psychiatric morbidity in patients with advanced cancer. *Psycho-Oncology*, 21(12), 1275–1284.
24. Rogers, S. N., Mepani, V., Jackson, S., & Lowe, D. (2017). Health-related quality of life, fear of recurrence, and emotional distress in patients treated for thyroid cancer. *British Journal of Oral & Maxillofacial Surgery*, 55(7), 666–673.
25. Löwe, B., Decker, O., Müller, S., Brähler, E., Schellberg, D., Herzog, W., et al. (2008). Validation and standardization of the generalized anxiety disorder screener (GAD-7) in the general population. *Medical Care*, 46(3), 266–274.
26. Ziegler, M., Kemper, C. J., & Kruey, P. (2014). Short scales—five misunderstandings and ways to overcome them. *Journal of Individual Differences*, 35(4), 185–189.