Journal of Affective Disorders 330 (2023) 188–197

Contents lists available at ScienceDirect

Journal of Affective Disorders

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

Research paper

The Traumatic Grief Inventory-Clinician Administered: A psychometric evaluation of a new interview for ICD-11 and DSM-5-TR prolonged grief disorder severity and probable caseness

Lonneke I.M. Lenferink [a,b,c,*], Minita Franzen [b], Peter M. ten Klooster [c], Christine Knaevelsrud [d], Paul A. Boelen [a,e], Carina Heeke [d]

[a] Department of Clinical Psychology, Faculty of Social Sciences, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, the Netherlands
[b] Department of Clinical Psychology and Experimental Psychopathology, Faculty of Behavioral and Social Sciences, University of Groningen, the Netherlands
[c] Department of Psychology, Health, & Technology, Faculty of Behavioural, Management, and Social Sciences, University of Twente, Drienerlolaan 5, 7522 NB Enschede, the Netherlands
[d] Freie University Berlin, Department of Clinical-Psychological Intervention, Habelschwerdter Allee 45, 14195 Berlin, Germany
[e] ARQ National Psychotrauma Center, Diemen, the Netherlands

ARTICLE INFO

Keywords:
Prolonged grief
Assessment
Screening
ICD-11
DSM
DSM-5-TR

ABSTRACT

Background: There is a need for an interview-based measure to assess Prolonged Grief Disorder (PGD) included in the text revision of the fifth Diagnostic and Statistical Manual for Mental Disorder (DSM-5-TR) and 11th edition of the International Classification of Disease (ICD-11). We evaluated the psychometric properties of the Traumatic Grief Inventory-Clinician Administered (TGI-CA); a new interview measuring DSM-5-TR and ICD-11 PGD severity and probable caseness.

Methods: In 211 Dutch and 222 German bereaved adults, the: (i) factor structure, (ii) internal consistency, (iii) test-retest reliability, (iv) measurement invariance across subgroups (e.g., differing in language), (v) prevalence of probable caseness, (vi) convergent validity, and (vii) known-groups validity were examined.

Results: Confirmatory factor analyses (CFAs) showed acceptable fit for the unidimensional model for DSM-5-TR and ICD-11 PGD. Omega values indicated good internal consistency. Test-retest reliability was high. Multi-group CFAs demonstrated configural and metric invariance for DSM-5-TR and ICD-11 PGD criteria for all group-comparisons; for some we found support for scalar invariance. Rates of probable caseness for DSM-5-TR PGD were lower than ICD-11 PGD. Optimal agreement in probable caseness was reached when increasing the number of accessory symptoms for ICD-11 PGD from 1+ to 3+. Convergent and known-groups validity was demonstrated for both criteria-sets.

Limitations: The TGI-CA was developed to assess PGD severity and probable caseness. Clinical diagnostic interviews for PGD are needed.

Conclusions: The TGI-CA seems a reliable and valid interview for DSM-5-TR and ICD-11 PGD symptomatology. More research in larger and more diverse samples is needed to further test its psychometric properties.

1. Introduction

Longing for the deceased and preoccupation with the deceased are common reactions to the death of a loved one. Most people adapt to bereavement without professional support (Lenferink et al., 2020; Nielsen et al., 2019). When grief reactions last and disrupt daily functioning, a grief disorder could be considered. The definition, conceptualization, and assessment of a grief disorder are topics of debate among scholars (see Lenferink et al., 2021). Currently, six different diagnostic criteria-sets for pathological or disturbed grief are used in practice and research (Boelen and Lenferink, 2020; Eisma et al., 2022). In short, a first set of criteria, called Prolonged Grief Disorder (PGD), was proposed by Prigerson et al. (2009) for inclusion in the fifth edition of the Diagnostic and Statistical Manual for Mental Disorder (DSM-5) and the 11th edition of the International Classification of Diseases (ICD-11). A second criteria-set, named Complicated Grief was proposed by Shear et al.

* Corresponding author at: Department of Clinical Psychology, Faculty of Social Sciences, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, the Netherlands.
E-mail address: l.i.m.lenferink@uu.nl (L.I.M. Lenferink).

https://doi.org/10.1016/j.jad.2023.03.006
Received 10 June 2022; Received in revised form 6 February 2023; Accepted 6 March 2023
Available online 11 March 2023

(2011). However, these criteria-sets proposed by Prigerson and Shear were not included in these classifications systems. Instead, Persistent Complex Bereavement Disorder was included in DSM-5 (APA, 2013). PCBD was included as condition for further study and reflects a combination of the proposals by Prigerson et al. (2009) and Shear et al. (2011). Then, a draft proposal for ICD-11 PGD criteria was published by Maercker et al. (2013), followed by slightly different PGD criteria published in the final ICD-11 (WHO, 2018). Recently, a text revision of the DSM-5 (DSM-5-TR) was published including PGD (APA, 2022).

While the most recent criteria for grief disorder in ICD-11 and DSM-5-TR are both called PGD, the content of the criteria slightly differ (Eisma et al., 2022). A first notable difference is that ICD-11 PGD can be diagnosed in adults and children six months after the loss. DSM-5-TR PGD can be diagnosed 12 months after the loss (and after six months for children). An important similarity is that both diagnostic sets include separation distress, in the form of yearning or longing for the deceased and/or preoccupation with the deceased. At least one of these should be present to meet DSM-5-TR and ICD-11 criteria. Then the DSM-5-TR PGD proposal proceeds with eight additional symptoms of which at least three should be present, including identity disruption, disbelief, avoidance of reminders, intense emotional pain, difficulty reintegrating (social) activities, numbness, life is meaningless, and loneliness. The ICD-11 PGD criteria include ten additional symptoms, described as "sadness, guilt, anger, denial, blame, difficulty accepting the death, feeling one has lost a part of one's self, an inability to experience positive mood, emotional numbness, difficulty in engaging with social or other activities". The presence of at least one of these 10 additional symptoms, together with separation distress, indicates probable 'caseness' for ICD-11 PGD (WHO, 2018). Furthermore, in DSM-5-TR and ICD-11 these responses need to cause significant impairment in daily functioning and to exceed expected social, cultural, and religious norms. The difference in content between these two recent criteria-sets and the difference in diagnostic scoring rules likely explains the difference in prevalence rates. For instance, in a representative German bereaved sample, prevalence rates of 3.3 % and 4.2 % were found when using these diagnostic scoring rules for DSM-5-TR and ICD-11, respectively (Rosner et al., 2021).

Measures of the diagnostic agreement and validity of the ICD-11 PGD and DSM-5-TR PGD criteria-sets were evaluated in at least six prior studies (Boelen and Lenferink, 2020, 2021; Kokou-Kpolou et al., 2022; Lenferink et al., 2022; Prigerson et al., 2021; Rosner et al., 2021). One study used exploratory factor analyses and found that ICD-11 and DSM-5-TR items were best represented as a unidimensional construct (Kokou-Kpolou et al., 2022). Two studies compared the statistical fit of a two-factor model (as proposed in ICD-11 and DSM-5-TR with one factor representing separation distress symptoms and a second factor representing additional cognitive, emotional, behavioral symptoms) with a one-factor model for ICD-11 PGD and DSM-5-TR PGD items, showing that the two-factor model of both criteria-sets did not show a significant better fit than the one-factor model (Boelen and Lenferink, 2020; Lenferink et al., 2022). Prigerson et al. (2021) and Boelen and Lenferink (2021) only evaluated the factor structure of DSM-5-TR PGD items and found support for a unidimensional construct. Support was also found for longitudinal measurement invariance (MI) of PGD DSM-5-TR items (Boelen and Lenferink, 2021). In the study from Boelen and Lenferink (2020), optimal agreement with DSM-TR-PGD (3+ accessory symptoms) was reached when 2+ accessory symptoms were present for ICD-11 PGD. Rosner et al. (2021) found optimal agreement between the two criteria-sets for DSM-5-TR PGD (3+ accessory symptoms) and ICD-11 PGD (4+ accessory symptoms). Lenferink et al.'s (2022) study showed similar prevalence rates of ICD-11 PGD (1+ accessory symptoms) and DSM-5-TR PGD (3+ accessory symptoms). These studies indicate that items from both criteria-sets have a unidimensional factor structure and that agreement in diagnostic rates increased when increasing the number of accessory symptoms for ICD-11 PGD.

In addition, people meeting criteria for self-rated DSM-5-TR PGD (3+ accessory symptoms) and ICD-11 PGD (1+ accessory symptom), have been found to report higher concurrent psychopathology levels (i.e., depression, anxiety, agoraphobia, and suicidal ideation) than people not meeting the criteria. The cases also differed in terms of established background and loss-related correlates of PGD (Boelen and Lenferink, 2020). Symptom severity levels of DSM-5-TR PGD and ICD-11 PGD were significantly related to concurrent PTSD and depression levels and established background and loss-related correlates of PGD (Lenferink et al., 2022). Furthermore, DSM-5-TR PGD levels predicted PTSD, depression, and quality of life levels over time (Prigerson et al., 2021). These findings support convergent and known-groups validity of DSM-5-TR and ICD-11 PGD.

While these prior studies provided preliminary support for the diagnostic performance of DSM-5-TR PGD and ICD-11 PGD, they are limited in several respects. First, all studies used data collected before the release of the PGD criteria in ICD-11 and DSM-5-TR. Except for two studies (i.e., Lenferink et al., 2022; Kokou-Kpolou et al., 2022), they all used outdated measures that were not designed to measure PGD as defined in DSM-5-TR and ICD-11. Furthermore, all these prior studies relied on self-reported questionnaire data. There was one exception: in one of the three existing datasets used by Prigerson et al. (2021) interview-based data were used (Prigerson et al., 2009). Interview-based assessments have several advantages. First, interviews may result in more accuracy in detecting cases, because surveys tend to overestimate symptom severity (cf. Lim et al., 2018). Second, interview-based assessment usually yield less missing data and non-responses compared with survey-based assessment. Third, interviews allow for explanation of questions and answers, which may lead to more measurement accuracy (Marta-Pedroso et al., 2007). Consequently, there is an urgent need to develop and validate a structured interview that may be used to screen for both DSM-5-TR PGD *and* ICD-11 PGD. Such an interview would enable direct comparison of the clinical utility of the PGD criteria-sets as defined in DSM-5-TR and ICD-11 (Lenferink et al., 2021). Accordingly, we developed the Traumatic Grief Inventory-Clinician Administered (TGI-CA). The TGI-CA is an interview-version of the 22 item self-report version of the TGI, the TGI-SR+ (Lenferink et al., 2022), designed to assess PGD severity and probable caseness in terms of DSM-5-TR and ICD-11 PGD by asking bereaved people to rate how often they experienced each PGD symptom during the previous month on 5-point Likert scales. The TGI-CA is therefore not a clinical diagnostic interview allowing to set a formal diagnosis of PGD but, instead, only allows to screen for probable PGD.

The main aim of this study was to evaluate the psychometric properties of the TGI-CA. We first examined the factor structure. When comparing a two-factor model with a one-factor model, we did not expect a relevant improvement in fit for the 2-factor model over the unidimensional model for DSM-5-TR PGD (Boelen and Lenferink, 2020, 2021; Prigerson et al., 2021) and ICD-11 PGD (Boelen et al., 2018, 2019b; Killikelly et al., 2020). Different from what we planned in our preregistration (https://osf.io/63wkr), we did not evaluate psychometric properties of TGI-CA assessing PGD as defined by Prigerson et al. (2009) and DSM-5 PCBD (APA, 2013), as these criteria appeared to be outdated at the time of data-collection. Moreover, we expected the unidimensional factors to have good internal consistency reflected by McDonald's omega >0.70 (Hayes and Coutts, 2020) and temporal stability as indicated by strong associations between symptom levels over time.

We expected to demonstrate MI of the PGD items (for DSM-5-TR and ICD-11 separately) tapped with the TGI-CA across subgroups of bereaved people differing in terms of language, gender, age, educational level, kinship to the deceased, time since loss, and cause of death. Furthermore, we examined probable rates of DSM-5-TR and ICD-11 PGD. In addition, convergent validity was investigated by examining associations between probable caseness of DSM-5-TR and ICD-11 PGD and concurrently assessed levels of posttraumatic stress disorder (PTSD), depression, and functional impairment. We expected that probable

caseness (vs. probable non-caseness) was associated with more severe PTSD, depression, and functional impairment (Boelen and Lenferink, 2020; Lenferink et al., 2022; Prigerson et al., 2021). Lastly, background and loss-related correlates of DSM-5-TR and ICD-11 PGD were examined for evaluating known-groups validity. We expected to find higher DSM-5-TR and ICD-11 PGD levels for women (vs. men), older (vs. younger) bereaved people, lower (vs. higher) educated people, people who lost a spouse/child (vs. another loved one), more recently (vs. remotely) bereaved, and people whose loved one died due to an unnatural (vs. natural) cause (Boelen and Lenferink, 2020; Heeke et al., 2019; Lundorff et al., 2017).

2. Methods

2.1. Procedure

This cross-sectional study was part of a longitudinal study called 'TGI-CA Assessment after Loss in Europe (TALE) project' (https://osf.io/a6hmc/). Dutch and German-speaking adults whose spouse, family member, or friend died at least six months earlier were eligible to participate. People were excluded when they were suicidal and/or had been diagnosed with a psychotic disorder (assessed with single items). Various recruitment strategies were used, including advertisement on social media, via convenience sampling (by the researchers reaching out to their social network), and by snowball sampling (asking participants to refer other to this study). Participants did not receive financial compensation. However, first-year psychology students could participate in exchange for course credits. This study was preregistered before start of data-collection (https://osf.io/63wkr) and approved by local ethics committees in the Netherlands (Ethics Committee Psychology, University of Groningen) and Germany (Ethics Committee of Freie University Berlin). Participants gave written informed consent. Data collection started in November 2019 and ended in September 2020. The dataset is available via doi:10.17605/OSF.IO/A6HMC. Participation consisted of a telephone interview that lasted about 30–45 min. Interviews were conducted by 14 undergraduate and graduate psychology students. They received a six-hour training by one or more of the authors who are experts in disturbed grief and interviewing (LL, CH, MF, and PB). The training addressed the theoretical background and phenomenology of disturbed grief, PTSD, and depression and involved practicing interviewing skills. Monthly supervision took place with the interviewers and the research team.

2.2. Participants

In total, 433 people participated. Table 1 summarizes background characteristics of the participants. About half of the sample was Dutch-speaking (n = 211, 48.7 %). For the total sample, four out of five participants were women. Participants were 43 years old on average with the Dutch sample being slightly, but significantly younger than the German sample. Compared with the German sample, significantly more people in the Dutch sample had a university degree (55.5 % vs. 44.6 %). About four out of 10 people in the total sample had lost a partner or child and three quarter lost their loved one due to a natural cause. On average, the loss occurred six years ago; for the Dutch sample the death took place significantly more recently compared with the German sample.

Table 1
Characteristics of participants.

	Dutch sample (n = 211)	German sample (n = 222)	Total (N = 433)	Differences between samples
Gender, N (%)				$\chi^2(1, 433) = 0.02$, $p = .896$
Man	40 (19.0)	41 (18.5)	81 (18.7)	
Woman	171 (81.0)	181 (81.5)	352 (81.3)	
Other	0 (0)	(0)	0 (0)	
Age, M (SD)	41.3 (16.7)	44.8 (16.9)	43.1 (16.9)	$t(431) = -2.16$, $p = .031$
Level of education, N (%)				$\chi^2(1, 433) = 5.01$, $p = .024$
Primary school	2 (0.9)	1 (0.5)	3 (0.7)	
High school	51 (24.2)	62 (27.9)	113 (26.1)	
Vocational education	41 (19.4)	60 (27.0)	101 (23.3)	
University	117 (55.5)	99 (44.6)	216 (49.9)	
Deceased relative is my… N (%)				$\chi^2(1, 433) = 0.31$, $p = .576$
Partner/spouse	64 (30.3)	55 (24.8)	119 (27.5)	
Child	16 (7.6)	35 (15.8)	51 (11.8)	
Parent	63 (29.9)	67 (30.2)	130 (30.0)	
Sibling	4 (1.9)	12 (5.4)	16 (3.7)	
Grandparent	39 (18.5)	34 (15.3)	73 (16.9)	
Friend	11 (5.2)	8 (3.6)	19 (4.4)	
Other	14 (6.6)	11 (5.0)	25 (5.8)	
Time since loss in years, M (SD)	5.28 (6.6)	7.2 (9.3)	6.24 (8.2)	$t(431) = -2.39$, $p = .017$
Cause of death				$\chi^2(1, 433) = 0.26$, $p = .608$
Natural cause	165 (78.2)	169 (76.1)	334 (77.1)	
Suicide	32 (15.2)	25 (11.3)	57 (13.2)	
Accident	9 (4.3)	18 (8.1)	27 (6.2)	
Homicide	1 (0.5)	2 (0.9)	3 (0.7)	
Other	4 (1.9)	8 (3.6)	12 (2.8)	
Expectedness of death…, M (SD)	3.1 (1.6)	3.2 (1.7)	3.2 (1.6)	$t(430) = -0.57$, $p = .567$
Number of losses				$\chi^2(1, 432) = 30.82$, $p < .001$
1	125 (59.2)	73 (33.0)	198 (45.7)	
2	44 (20.9)	53 (24.00)	97 (22.5)	
3	28 (13.3)	48 (21.7)	76 (17.6)	
4	11 (5.2)	36 (16.3)	47 (10.9)	
5 or more	3 (1.4)	11 (5.0)	14 (3.2)	

Note. The following variables were dichotomized: relationship to the deceased (0 = other than child/spouse, 1 = child/spouse), cause of death (0 = natural, 1 = unnatural), and number of losses (0 = single loss, 1 = multiple losses).

2.3. Measures

2.3.1. PGD symptoms

DSM-5-TR and ICD-11 PGD symptoms were assessed with the TGI-CA. This interview was developed in Dutch and German language for the purpose of this study. The TGI-CA is based on the 18-item Traumatic Grief Inventory-Self Report, originally developed and validated in Dutch (TGI-SR; Boelen and Smid, 2017; Boelen et al., 2019a) and subsequently translated to German (Comtesse and Rosner, 2017). Because the TGI-SR was developed before the release of the ICD-11 and DSM-5-TR it did not capture all PGD criteria. The TGI-SR was therefor recently extended by four items. This 22-item self-report questionnaire, called the Traumatic Grief Inventory-Self Report Plus (TGI-SR+) has been shown to be a reliable and valid instrument (Lenferink et al., 2022). Two changes were made in the TGI-CA compared to the TGI-SR+. First, in the TGI-CA, items were phrased as questions rather than statements (as in the TGI-SR+). Second, in the instructions and items, we replaced the wording "deceased loved one" with the first name (e.g., "Albert") or relationship (e.g., "your husband") of the deceased person. English, Dutch, and German translation of TGI-CA are freely available (https://osf.

io/a6hmc/). Participants were instructed to state how often they experienced each symptom during the previous month, on 5-point Likert scales with anchors 1 = never, 2 = seldom, 3 = sometimes, 4 = often, 5 = always. See Supplemental Table 1 for TGI-CA items.

To meet DSM-5-TR PGD criteria, at least one out of two symptoms of the B Criterion, at least three out of eight symptoms of the C Criterion, and the functional impairment criterion should be endorsed (APA, 2013). A symptom is considered endorsed when rated with 4 or 5. To fulfill ICD-11 PGD criteria, the following scoring rule was used: the presence of at least one out of two symptoms of the B criterion, at least one out of ten symptoms of the C criterion, and endorsement of the functional impairment criterion (WHO, 2018). Again, a score of at least 4 was used for symptom endorsement.

2.3.2. PTSD levels

The Dutch (Boeschoten et al., 2014) and German (Krüger-Gottschalk et al., 2017) Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5) was used to assess PTSD symptoms related to the death of their loved one (Weathers et al., 2013). It encompasses 20 items, scored on 5-point Likert scales (0 = not at all, 1 = a little bit, 2 = moderately, 3 = quite a bit, 4 = extremely). An example item is: "*In the past month, how much were you bothered by feeling jumpy or easily startled?*" While "in the past month, how much were you bothered by…" is included in the instruction in the original measure, we included it in each item during the interviews. McDonald's omega levels were high in the current study; ω = 0.91 for Dutch and ω = 0.90 for German sample.

2.3.3. Depression levels

Depression levels were assessed using the Dutch (van Steenbergen-Weijenburg et al., 2010) and German (Gräfe et al., 2004) translation of the Patient Health Questionnaire-9 (PHQ-9; Kroenke et al., 2001). The PHQ-9 assesses depression symptoms experienced during the past two weeks. Items (e.g., "Over the last 2 weeks, how often have you been bothered by any of the following problems? Feeling down, depressed, or hopeless") were rated on 4-point Likert scales (0 = not at all, 1 = on some days, 2 = on more than half of the days, 3 = almost every day). Item scores were summed to represent a total score ranging from 0 to 27. We rephrased the items such that they were questions instead of statements (e.g. "Over the last 2 weeks, how often have you been bothered by feeling down, depressed, or hopeless?"). The PHQ-9 demonstrated acceptable internal consistency in the Dutch (ω = 0.77) and German (ω = 0.82) sample.

2.3.4. Functional impairment

Functional impairment was measured with the 5-item Work and Social Adjustment Scale (WSAS; Mundt et al., 2002) in Dutch (de Graaf et al., 2009) and German (Lutz et al., 2019). People rated to what extent the death of their loved one impaired them in their (i) work, (ii) household chores, (iii) social activities, (iv) leisure activities, and (v) close relationships, on a scale from 1 = not at all through 9 = severely. The answer option "not applicable" was added to the item referring to work. The WSAS demonstrated good internal consistency in both samples (Dutch ω = 0.81 and German ω = 0.80).

2.4. Statistical analyses

2.4.1. Descriptive statistics

We started with testing for differences between the language groups (0 = Dutch, 1 = German) in terms of the background and loss-related characteristics. Chi-square tests were used for dichotomized variables (gender (0 = man, 1 = woman), educational level (0 = other than university, 1 = university), kinship (0 = other than child/spouse, 1 = child spouse), cause of death (0 = natural, 1 = unnatural), and number of losses (0 = single loss, 1 = multiple loss)). *t*-tests were used for continuous variables (i.e., age in years, time since loss in years, and expectedness of death).

2.4.2. Confirmatory factor analyses

Confirmatory factor analyses (CFAs) were performed in Mplus version 8.4 (Muthén and Muthén, 1998–2019) to examine the factor structure for DSM-5-TR PGD and ICD-11 PGD items. Kurtosis values for each TGI-CA item were <10 and skewness <3; robust maximum likelihood estimation was used. The following fit statistics were evaluated, whereby Comparative Fit Index (CFI) and Tucker Lewis Index (TLI) values > 0.90 represent acceptable fit and >0.95 excellent fit. Values below 0.10 of the Root Mean Square Error of Approximation (RMSEA) with 90 % confidence intervals (CI) and Standardized Root Mean Square Residual (SRMR) indicated acceptable fit and below 0.05 excellent fit. Lower Akaike, Bayesian, and Sample-Size adjusted Bayesian information criteria (AIC, BIC, and SS-BIC) were preferred when comparing the one- and two-factor models for ICD-11 and DSM-5-TR separately (Kline, 2005). To statistically test the difference in fit of the nested one-and two-factor models, Satorra-Bentler scales Chi-Square tests were used (Muthén & Muthén, 2023). Less than 2 % of the TGI-CA items were missing and handled with full information maximum likelihood estimation.

2.4.3. International consistency

Internal consistencies of the DSM-5-TR PGD and ICD-11 PGD items were examined using Hayes' add-on for SPSS to calculate McDonald's omega (ω). Values >0.70 represent acceptable internal consistency (Hayes and Coutts, 2020).

2.4.4. Test-retest reliability

Temporal stability was examined by correlating DSM-5-TR and ICD-11 PGD sum scores assessed at the first time-point with sum scores using data from 289 people who completed the TGI-CA twelve months later. For more details about all measures used in follow-up interview see https://osf.io/a6hmc/.

2.4.5. Measurement invariance

MI across groups was tested for DSM-5-TR and ICD-11 separately following the guidelines of Van de Schoot et al. (2012). Testing MI consists of comparing the fit of one model with a more constrained model, using the CONFIGURAL METRIC SCALAR COMMAND in Mplus. A difference in CFI value of ≤0.02 and a non-significant χ^2 value ($p > .05$) demonstrates invariance for the more constrained model (Chen, 2007; Gloster et al., 2021; Putnick and Bornstein, 2016). Because concerns have been raised that the Chi-square difference test is too strict in favoring less constrained models (Meade et al., 2008; Milfont and Fischer, 2010), we relied on the difference in CFI values for model comparisons.

Examination of MI constitutes of three steps. The first step is examination of a model in which factor loadings and intercepts are allowed to vary freely across the groups (Model 1: configural invariance). Support for configural invariance implies that the number of factors is similar across groups and the underlying factor structure is adequate in both groups. In the second step, factor loadings were constrained to be equal across groups (Model 2: metric invariance), which assumes that the items contribute equally to the factor across groups. In the third step, the factor loadings and intercepts were constrained to be equal across groups (Model 3: scalar invariance). When supported, this indicates that the items contribute equally to the factor across groups and that the levels of the underlying items (intercepts) are equal across groups. Multigroup CFAs were conducted among language groups (Dutch vs. German), gender (men vs. women), age (younger than 45 years vs. older than 45 years), educational level (lower than university vs. university), kinship to the deceased (child/spouse vs other), time since loss (less vs. more than two years bereaved), and cause of death (natural vs. unnatural).

A complementary method to test for MI, which was pointed out to us by one of the reviewers and therefore not described in our preregistration, is to examine the presence of differential item functioning

(DIF; Elis, 1989). Assessing DIF entails testing whether individual item scores differ across groups (e.g., Dutch vs. German) while controlling for the overall latent variable (e.g., PGD). Following the procedures used in prior research (Shevlin et al., 2022), DIF was tested by specifying a Multiple Indicators Multiple Causes (MIMIC) model, because it allows among others for examination of multiple grouping variables in one factor model and can be used to identify the source of any invariance at the item level (Rubio et al., 2008). The MIMIC model was built by regressing the latent variable on each of the grouping variables. Furthermore, the direct effect for each item of the latent variable on the grouping variables was fixed to zero. Modification indices (MIs, indicating the expected change in chi-square for one degree of freedom) and standardized expected parameter change values (SEPCs, indicating the estimated value of the regression parameter) were used for determining which direct effect should be included to improve model fit. A direct path was included when its MI was >10 (which are displayed by default in MPlus) and SEPC >0.20 (as recommended by Kaplan (1989)). In an iterative process, the path with the largest MI and SEPC values was freed and the model was re-estimated. This was continued until the model did not contain paths with MI and SEPC values >10 and 0.20, respectively.

2.4.6. Probable prevalence rates

Frequencies of probable caseness of PGD was determined using diagnostic scoring rules for DSM-5-TR (APA, 2022) and ICD-11 (WHO, 2018), which requires ≥1 B criterion symptom for DSM-5-TR and ICD-11, and ≥3 C criterion symptoms for DSM-5-TR PGD and ≥1 for ICD-11 PGD, plus endorsement of functional impairment criterion (i.e., D criterion; see Supplemental Table 1). Chi-square tests were used to compare probable prevalence rates of PGD between DSM-5-TR and ICD-11.

Following prior research (Boelen and Lenferink, 2020; Rosner et al., 2021), we also calculated probable ICD-11 PGD rates when increasing the number of required C criterion symptoms for ICD-11 PGD, from 2+ through 6+. Pairwise agreement between DSM-5-TR PGD and ICD-11 PGD (with different numbers of Criterion C symptoms) was evaluated using Kappa statistics with 95 % confidence intervals (CIs).

2.4.7. Convergent validity

Differences in levels of PTSD, depression, and functional impairment between people with and without probable PGD in DSM-5-TR and ICD-11 (considered separately) were tested using Mann-Whitney tests for non-normally distributed outcomes. The Kolmogorov-Smirnov test indicated that symptom levels were non-normally distributed. None of the PTSD and depression items were missing. On the functional impairment measure, 15 participants (3.5 %) responded 'not applicable' to the item referring to impairment in work and one person (0.2 %) missed an answer for the item referring to impairment related to family and relations. We replaced these answers with item means. Effect sizes r were calculated for differences in symptom levels between probable cases and non-cases by dividing the z-scores by the square root of the total number of participants. Effect sizes $r < 0.30$ were considered small, ≥0.30 and <0.50 medium, and ≥0.50 large (Rosenthal, 1991).

2.4.8. Known-groups validity

To test whether the summed 10 DSM-5-TR items and 12 ICD-11 PGD items differed between gender (women vs. men), age (in years), educational level (lower than university vs. university), kinship to the deceased (loss of child/spouse vs other), time since loss (in years), and cause of loss (natural vs. unnatural loss), Mann-Whitney tests were conducted for dichotomized variables and Spearman's correlations for continuous outcomes.

3. Results

3.1. Dimensionality of the TGI-CA

Table 2 shows the results of the CFAs. All fit indices were acceptable for the one- and two-factor model for DSM-5-TR PGD. The fit of the two-factor model was slightly better as indicated by lower AIC, BIC, and SS-BIC values and a significant χ^2 difference test ($\Delta\chi^2 = 6.31$, $p < .05$). However, a very strong correlation was found between the two factors ($r = 0.94$, $p < .001$) indicating that factors were not meaningfully distinct. The more parsimonious one-factor model was therefore retained.

For the unitary model of ICD-11 PGD, the RMSEA and SRMR values were acceptable; the CFI and TLI values were below 0.90. Again the fit of the two-factor model slightly improved indicated by lower AIC, BIC, and SS-BIC values and a significant χ^2 difference test ($\Delta\chi^2 = 8.41$, $p < .01$), but its TLI value was still below 0.90. Again, a strong association was found between the two factors ($r = 0.88$, $p < .001$), pointing to the preference of the more parsimonious one-factor model for ICD-11 PGD. Additionally, MIs for the one-factor model indicated that the error-terms of item-pairs C8 ("difficulties experiencing positive feelings") and C9 ("feeling emotionally numb") and C8 and C10 ("difficulty moving on") were strongly correlated. Considering the item content, this likely stemmed from non-random measurement error. When correlating these error-terms, all fit indices were acceptable. See Table 3 for factor loadings.

3.2. Internal consistency

McDonald's omega values were 0.89 for the 10 items representing DSM-5-TR criteria and 0.90 for the 12 ICD-11 PGD items.

3.3. Test-retest reliability

Association between DSM-5-TR PGD levels at the first time-point were strongly associated with symptom levels 12 months later ($\rho(289) = 0.82$, $p \leq .001$). For ICD-11 PGD, similar association was found ($\rho(289) = 0.81$, $p \leq .001$).

3.4. Measurement invariance

Findings from the multigroup CFAs are shown in Supplemental Material Tables 2 and 3. For DSM-5-TR PGD, configural, metric, and scalar invariance were demonstrated for the group-comparisons regarding language, gender, age, educational level, and cause of loss. This means that across these subgroups, the underlying factor structure of the DSM-5-TR PGD items was similar, items contributed equally to their factor, and item means showed similar patterns. Regarding multigroup CFAs for the subgroups in terms of kinship to the deceased and time since loss, we found support for configural and metric invariance, but not for scalar invariance. This indicates that the underlying factor structure of the items was similar across the groups and that the items contributed equally to the factors, but patterns of item means were not equal between these groups.

For ICD-11 PGD, we demonstrated configural, metric, and scalar invariance across gender, educational level, and cause of loss. Thus, underlying factor structure, contribution of item to the factor, and pattern of item means were similar across these groups. For the other group-comparisons, we found support for configural and metric invariance, but not for scalar invariance. This indicates that the underlying factor structure and contribution of items to the factor but not the item means were equal between these groups.

Findings from MIMIC models for DSM-5-TR PGD, indicated that DIF was detected for four (out of 70 possible) item to grouping variable paths. Fit indices of DIF models are shown in Supplemental Material Table 4. The largest MI and SEPC values were found for age and the item representing PGD B2 symptom (MI = 15.33, SEPC = 0.380). After re-

Table 2
Fit indices confirmatory factor analysis.

	χ^2 (df)	*p*-Value	CFI	TLI	RMSEA (90 % CI)	SRMR	AIC	BIC	SS-BIC
DSM-5-TR PGD									
1-Factor model	105.21 (35)	<.001	0.95	0.93	0.068 (0.053–0.083)	0.044	11,719.89	11,842.02	11,746.81
2-Factor model	98.79 (34)	<.001	0.95	0.93	0.066 (0.051–0.082)	0.043	11,713.88	11,840.08	11,741.70
ICD-11 PGD									
1-Factor model	302.99 (54)	<.001	0.89	0.87	0.086 (0.074–0.098)	0.056	13,501.69	13,648.24	13,534.00
2-Factor model	217.76 (53)	<.001	0.90	0.87	0.085 (0.073–0.097)	0.054	13,492.40	13,643.01	13,525.60
1-Factor model with correlated error terms	168.36 (52)	<.001	0.93	0.91	0.072 (0.060–0.084)	0.050	13,427.27	13,581.96	13,461.37

Note. AIC = Akaike Information Criterion; BIC = Bayesian Information Criterion; CFI = Comparative Fit Index; CI = confidence interval; DSM-5-TR = 5th text revised edition of the Diagnostic and Statistical Manual of Mental Disorders. ICD-11 = 11th edition of the International Classification of Diseases; PGD = Prolonged Grief Disorder; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual; SS-BIC = sample-size adjusted Bayesian information criterion; TLI = Tucker Lewis Index.

Table 3
Standardized factor loadings of the 1-factor models of DSM-5-TR PGD and ICD-11 PGD (*N* = 433).

	Factor loading	SE
DSM-5-TR PGD		
B1	0.730	0.028
B2	0.563	0.040
C1	0.728	0.029
C2	0.621	0.037
C3	0.311	0.056
C4	0.786	0.023
C5	0.633	0.037
C6	0.660	0.037
C7	0.763	0.024
C8	0.682	0.032
ICD-11 PGD		
B1	0.704	0.032
B2	0.589	0.037
C1	0.784	0.022
C2	0.580	0.044
C3	0.672	0.038
C4	0.649	0.036
C5	0.385	0.050
C6	0.756	0.030
C7	0.722	0.029
C8	0.636	0.037
C9	0.649	0.040
C10	0.605	0.040

Note. DSM-5-TR = 5th text revised edition of the Diagnostic and Statistical Manual of Mental Disorders; ICD-11 = 11th edition of the International Classification of Diseases; PGD = Prolonged Grief Disorder; SE = standard error.

estimating the model including this direct path, the R-squared value for this item increased from 0.320 to 0.345 ($\Delta R^2 = 0.025$), indicating that DIF accounted for 2.5 % of the variance in this particular item. Furthermore, the next largest direct path was found between time since loss and the PGD C1 symptom (MI = 12.95, SEPC = 0.422), followed by language groups and the item representing symptom C2 (MI = 11.87, SEPC = −0.339). Including these direct paths minimally increased the explained variance ($\Delta R^2 = 0.029$ and $\Delta R^2 = 0.020$, resp.).

For ICD-11 PGD, findings from MIMIC models showed DIF for seven (out of 85 possible) direct paths. The largest MI/SEPC values were found for the path between language and the item assessing symptom C5 (MI = 20.12, SEPC = −0.324). Adding this path increased R^2 for this item by 0.025. The next largest paths were found between age and item assessing symptom B2 (MI = 18.93, SEPC = 0.422, $\Delta R^2 = 0.033$), language and item assessing C4 symptom (MI = 14.58, SEPC = −0.376, $\Delta R^2 = 0.026$), kinship and item tapping C7 symptom (MI = 13.93, SEPC = 0.435, $\Delta R^2 = 0.013$), gender and C8 symptom (MI = 11.964, SEPC = −0.357, $\Delta R^2 = 0.022$), age and C8 symptom (MI = 10.662, SEPC = −0.259, $\Delta R^2 = 0.019$), and time since loss and item measuring C4 symptom (MI = 10.914, SEPC = −0.349, $\Delta R^2 = 0.011$).

3.5. Probable PGD caseness

Twenty-one people (4.8 %) met criteria for probable DSM-5-TR PGD. Twenty-seven people (6.2 %) met criteria for probable ICD-11 PGD; 20 people (4.6 %) met criteria for both DSM-TR PGD criteria and ICD-11 PGD when using the scoring rule of ≥1 accessory symptom. When increasing the number of accessory symptoms for ICD-11 PGD, the highest diagnostic agreement with DSM-5-TR PGD was reached, when 3+ accessory symptoms were used for ICD-11 PGD (see Table 4).

3.6. Convergent validity

Mann-Whitney tests showed that participants meeting probable DSM-5-TR criteria reported significantly higher PTSD, depression, and functional impairment levels than those not meeting the criteria. For ICD-11 PGD similar differences were found (see Table 5). Effect sizes of differences were small to medium.

3.7. Known-groups validity

As anticipated, DSM-5-TR PGD and ICD-11 PGD levels were significantly higher for women (vs. men), more recently (vs. less recently) bereaved people, people who lost a child/spouse (vs. other loved one), and those whose loved one died due to an unnatural (vs natural) cause of death (see Table 6). Educational level was not significantly related to PGD levels.

4. Discussion

We evaluated the psychometric properties of the TGI-CA, an

Table 4
Prevalence rates of ICD-11 PGD with increasing number of required accessory symptoms and pairwise agreement with DSM-5-TR PGD (*N* = 433).

	Caseness		Diagnostic agreement with DSM-5-TR PGD	
ICD-11 PGD caseness with increasing number of accessory symptoms	N	%	κ	95 CI
1+	27	6.2	0.82	0.70–0.94
2+	24	5.5	0.88	0.78–0.98
3+	20	4.6	0.92	0.84–1.10
4+	14	3.2	0.79	0.64–0.94
5+	14	3.2	0.79	0.64–0.94
6+	11	2.5	0.68	0.49–0.87

Note. DSM-5-TR = 5th text revised edition of the Diagnostic and Statistical Manual of Mental Disorders; ICD-11 = 11th edition of the International Classification of Diseases; PGD = Prolonged Grief Disorder.

Table 5
Differences in posttraumatic stress, depression, and functional impairment levels between cases and non-cases ($N = 433$).

	Median	IQR	U	*z*	*p*-Value	*r*
Posttraumatic stress disorder levels						
DSM-5-TR PGD non-cases	30.00	25.00–39.00	914.00	−6.10	<.001	−0.29
DSM-5-TR PGD cases	55.00	46.50–62.00				
ICD-11 PGD non-cases	30.00	25.00–39.00	1534.00	−6.27	<.001	−0.30
ICD-11 PGD cases	51.00	40.00–60.00				
Depression levels						
DSM-5-TR PGD non-cases	14.00	12.00–18.00	1012.50	−5.94	<.001	−0.29
DSM-5-TR PGD cases	23.00	19.50–27.00				
ICD-11 PGD non-cases	14.00	12.00–18.00	1644.50	−6.11	<.001	−0.29
ICD-11 PGD cases	22.00	18.00–26.00				
Functional impairment levels						
DSM-5-TR PGD non-cases	16.52	10.00–23.00	1339.50	5.34	<.001	−0.26
DSM-5-TR PGD cases	30.00	26.52–34.00				
ICD-11 PGD non-cases	16.00	10.00–23.00	1620.00	−6.14	<.001	−0.29
ICD-11 PGD cases	30.00	25.00–35.00				

Note. DSM-5-TR = 5th text revised edition of the Diagnostic and Statistical Manual of Mental Disorders; ICD-11 = 11th edition of the International Classification of Diseases; IQR = interquartile range; PGD = Prolonged Grief Disorder.

Table 6
Sociodemographic and loss-related correlates of prolonged grief ($N = 433$).

	DSM-5-TR PGD	Test statistic	ICD-11 PGD	Test statistic
Gender, median (IQR)				
Men	16.00 (13.00–23.50)	U = 10,949.50, z = −3.26, *r* = −0.16	18.00 (14.50–27.50)	U = 11,389.00, z = −2.826, r = −0.16
Women	19.00 (15.00–26.00)		22.00 (17.00–29.00)	
Education level, median (IQR)				
Lower than university	19.00 (14.50–25.00)	U = 22,933.00, z = −0.39, *r* = −0.01	21.00 (16.00–29.00)	U = 23,227.50, z = −0.16, *r* = −0.02
University	18.00 (14.00–26.00)		20.00 (16.00–29.00)	
Time since loss (in years)		$\rho = 0.37$		$\rho = -0.37$
Kinship to the deceased, median (IQR)				
Other than spouse/child	18.00 (14.00–23.00)	U = 14,689.50, z = −6.03, *r* = −0.29	20.00 (16.00–25.00)	U = 15,214.00, z = −5.62, *r* = −0.27
Spouse/child	22.00 (17.00–30.00)		25.00 (18.00–34.25)	
Cause of death, median (IQR)				
Natural	18.00 (14.00–25.00)	U = 13,445.00, z = −2.83, *r* = −0.14	20.00 (16.00–27.25)	U = 13,326.50, z = −2.94, r = −0.14
Unnatural	21.00 (16.00–30.00)		24.00 (17.00–34.00)	

Note. DSM-5-TR = 5th text revised edition of the Diagnostic and Statistical Manual of Mental Disorders; ICD-11 = 11th edition of the International Classification of Diseases; IQR = interquartile range; PGD = Prolonged Grief Disorder. * $p < .05$, ** $p < .01$, *** $p < .001$.

interview-based tool for PGD severity and probable caseness in terms of DSM-5-TR and ICD-11 PGD. This validation study was conducted among 211 Dutch and 222 German bereaved people. Psychometric properties of the TGI-CA were evaluated in terms of the factor structure, internal consistency, test-retest reliability, MI across subgroups, agreement in prevalence rates, convergent validity, and known-groups validity.

Our first main finding was that CFAs indicated that the unidimensional model for DSM-5-TR PGD and ICD-11 PGD showed acceptable fit. While the two-factor model (with separation distress and the accessory symptoms representing distinct factors) showed a slightly better statistical fit than the one-factor model, correlations between the two factors were very strong ($r = 0.94$ for DSM-5-TR PGD and $r = 0.88$ for ICD-11 PGD), pointing to no meaningful distinction between the two factors. The most parsimonious one-factor models were therefore retained. This aligns with prior research showing a satisfactory fit for the unidimensional models for DSM-5-TR PGD (Boelen and Lenferink, 2020, 2021; Prigerson et al., 2021) and ICD-11 PGD (Boelen et al., 2018, 2019a, 2019b; Killikelly et al., 2020) using instruments that did not assess all DSM-5-TR and ICD-11 PGD criteria. Recent studies evaluating psychometric properties of the self-report questionnaire version (i.e., TGI-SR+) of the TGI-CA also support the unidimensional factor structure of DSM-5-TR and ICD-11 PGD in exploratory and confirmatory factor analyses (Kokou-Kpolou et al., 2022; Lenferink et al., 2022). As expected, the unidimensional factors had good internal consistency, which is comparable to values found in a validation study using the TGI-SR+ (Lenferink et al., 2022). The TGI-CA demonstrated good temporal stability based on strong associations that were found using data from 289 people who completed the TGI-CA twice with a one year time-interval.

The second main finding was that for all multi-group CFAs, we found support for configural and metric invariance for both DSM-5-TR and ICD-11 PGD criteria. This means that, for both criteria-sets, the underlying factor structure and the contribution of each item to the factor is equal between Dutch vs. German speaking participants, men vs. women, younger vs. older people, lower vs. higher educated people, people who lost their child or spouse vs. other loved one, more recently vs. remotely bereaved people, and people who experienced a natural vs. unnatural loss. For some comparisons we also found support for scalar invariance, meaning that the groups can be compared on latent PGD scores. For DSM-5-TR PGD, this was true for language, gender, age, educational level, and cause of loss, but not for kinship to the deceased, and time since loss. For ICD-11 PGD, scalar invariance was supported for gender, educational level, and cause of loss, but not for language, age, kinship to the deceased, and time since loss. Thus, some caution is warranted when comparing the scores for these groups for which we could only establish partial MI. In addition, MIMIC models were estimated to test for DIF. After controlling for overall PGD levels, we found that some of the grouping variables (i.e., gender, age, language, kinship, or time since loss) led to difference in scores on three DSM-5-TR PGD items and five ICD-11 PGD items. The effect sizes were however small (i.e., additional explained variance when including the path ranged between 1.3 and 3.3

%), which indicated that it is unlikely that these differences substantially affect conclusions drawn about differences in PGD levels across these groups.

Our third main finding was that probable prevalence rates of DSM-5-TR PGD were somewhat lower than ICD-11 PGD (4.8 % vs. 6.2 %). A previous study using a representative German sample found similar rates, i.e., 3.3 % for DSM-5-TR and 4.2 % for ICD-11 PGD (Rosner et al., 2021). As indicated by prior research (Boelen and Lenferink, 2020; Eisma et al., 2020), the diagnostic scoring rule of 1+ accessory symptoms for ICD-11 might be too lenient, leading to overestimation of caseness. Indeed, when increasing the number of accessory symptoms for ICD-11 to 3+ accessory symptoms optimal agreement was found with DSM-5-TR PGD. Earlier studies also found that increasing the number of accessory symptoms for ICD-11 PGD led to higher agreement. For instance, Boelen and Lenferink (2020) found optimal agreement using 2+ and Rosner et al. (2021) found optimal agreement using 4+ accessory symptoms. Discrepancies between the number of accessory symptoms for ICD-11 might be due use of different instruments to assess DSM-5-TR and ICD-11 PGD across these studies.

Last, as expected, our findings supported convergent and known-groups validity for both criteria-sets. People who met criteria for DSM-5-TR PGD and ICD-11 PGD reported significantly higher PTSD, depression and functional impairment levels. This aligns with prior research (Boelen and Lenferink, 2020; Lenferink et al., 2022; Prigerson et al., 2021). Furthermore, women, more recently bereaved people, those who lost a spouse or child, and those whose loved one died due to an unnatural cause (e.g., suicide) reported more severe DSM-5-TR and ICD-11 PGD symptoms. These risk factors for PGD have also been identified earlier (see Heeke et al., 2019; Lobb et al., 2010; Lundorff et al., 2017). Unexpectedly, we did not find differences in PGD levels as function of educational level. A large Dutch population-based study showed that people who have lower education reported higher PGD levels than higher educated people (Newson et al., 2011). Newson et al.'s study operationalized educational level as a continuous score on a scale from 1 (primary education) to 6 (university). We used a dichotomized indicator (lower than university vs. university). The difference in operationalization of educational level may explain the difference in results.

When interpreting our findings, several limitations need to be taken into account. We assessed PGD levels by conducting telephone interviews. Answering questions while you have never met a person in real life, may have impacted the results. For instance, it may have lowered the threshold to share thoughts and feelings leading to more accurate answers because there was no dependent relationship between the interviewer and interviewee (unlike a therapist-client relationship). On the other hand, non-verbal communication was not possible to assess, which may have led interviewers to miss cues for asking some further questions (e.g., when someone answers "not at all", while the interviewer observes serious doubt about that answer). Moreover, the TGI-CA was developed for use in research and clinical practice to assess PGD severity and probable caseness. We now evaluated the psychometric properties of the TGI-CA in a non-clinical setting of which the majority of the participants did not experience elevated PGD levels. Further studies testing the psychometric properties of the TGI-CA in clinical samples are needed, preferably studies in which the TGI-CA can be compared with clinical diagnostic interviews for PGD in terms of the accuracy in detecting PGD caseness. However, till date this is not possible because clinical diagnostic interviews for PGD are not yet available. Furthermore, we examined associations between TGI-CA and PTSD and depression. Fixed ordering of these measures (i.e., we first examined depression, then PGD, followed by PTSD), may have inflated correlations between these measures. Moreover, examining associations between the TGI-CA and another PGD measure, such as the International Prolonged Grief Disorder Scale for the ICD-11 (IPGDS; Killikelly et al., 2020), would have provided important additional evidence for convergent validity (Campbell and Fiske, 1959). In addition, the wording of some TGI-CA items slightly deviate from the wording used in ICD-11 and DSM-5-TR. For instance, the TGI-CA item "In the past month, did you have intrusive thoughts or images related to the death of…" is assumed to represent "Preoccupation with thoughts or memories of the deceased person" in DSM-5-TR (APA, 2022) and 'persistent preoccupation with the deceased' in ICD-11 (WHO, 2018), but slightly differs from how these symptoms are worded. To what extent this affected our results remains to be studied, for instance by comparing the performance of this TGI-CA items with similar items used in other PGD measures, such as the item "I am preoccupied with thoughts about the deceased or circumstances of the death" in the IPGDS. Moreover, this study builds upon a body of literature comparing clinical usefulness of PGD criteria between the DSM and ICD classification system (Boelen and Lenferink, 2020; Bonanno and Malgaroli, 2020; Cozza et al., 2016; Haneveld et al., 2022; Mauro et al., 2017; Maciejewski et al., 2016). It is, however, noteworthy that these systems did not intend to be identical (First et al., 2015, 2021), which likely explains differences across findings using different criteria-sets. Lastly, our sample size was relatively small for some of the analyses regarding MI. For instance, the majority of our sample consisted of women who experienced natural losses. Pending replication of our findings in larger samples, our results regarding MI should be interpreted with caution.

To conclude, we demonstrated that the TGI-CA is a reliable and valid interview to assess DSM-5-TR and ICD-11 PGD severity and probable caseness. We demonstrated this by showing a satisfactory fit for unidimensional models of DSM-5-TR and ICD-11 PGD, which has also been found in prior research. Furthermore, (partial) MI was shown across subgroups of bereaved people. Optimal agreement between the prevalence rates of the two criteria-sets for PGD was reached when increasing the number of accessory symptoms for ICD-11 PGD. Based on this study, using at least 3+ accessory symptoms for ICD-11 PGD is advised. Our results also supported convergent and known-groups validity of the TGI-CA. These first results indicate that the TGI-CA is a valid and reliable tool that can be used by clinicians and researcher to assess PGD severity and probable caseness in terms of DSM-5-TR and ICD-11.

Role of the funding source

This research did not receive a grant from any funding agency, commercial or not-for-profit sectors.

CRediT authorship contribution statement

LL conceptualized he study. LL, CH, PB, and MF trained and supervised the interviewers. LL analyzed the data and wrote manuscript. PtK offered statistical support with MIMIC modeling. All authors reviewed and edited the draft of the manuscript.

Conflict of interest

None.

Data availability

This study was preregistered (https://osf.io/63wkr). The dataset is available via the Open Science Framework (doi:10.17605/OSF.IO/A6HMC). The Traumatic Grief Inventory-Clinician Administered is freely available in English, German, and Dutch via https://osf.io/a6hmc/.

Acknowledgments

We would like to express our gratitude to the following people for their support in recruiting participants and collecting the data by telephone interviews: Stance Hollander, Marleen van den Munckhof, Ria Vranken, Noa Becker, Hendrik Hofmann, Philip Frohn, Nina Heeres, Marianne Hoekstra, Marjel Buiter, Judith Kramer Jil Maiß, Annalena

Korn, Judith Walther, and Leah Fisches. We also thank Heidi Müller, Iris Hilgeman, Sarah Schoenmakers, Sanne ten Broeke, Ewout Poelmann, Daan van Heertum, and Faber van der Klift for their help with recruitment.

Funding statement

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jad.2023.03.006.

References

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th ed. American Psychiatric Publishing, Arlington, VA.

American Psychiatric Association, 2022. Diagnostic and Statistical Manual of Mental Disorders Fifth Edition Text Revision (DSM-5-TR). American Psychiatric Publishing, Arlington, VA.

Boelen, P.A., Djelantik, A.A.A.M.J., de Keijser, J., Lenferink, L.I.M., Smid, G.E., 2019a. Further validation of the traumatic grief inventory-self report (TGI-SR): a measure of persistent complex bereavement disorder and prolonged grief disorder. Death Stud. 43 (6), 351–364.

Boelen, P.A., Lenferink, L.I.M., 2020. Comparison of six proposed diagnostic criteria sets for disturbed grief. Psychiatry Res. 285, 112786.

Boelen, P.A., Lenferink, L.I.M., 2021. Prolonged grief disorder in DSM-5-TR: Early predictors and longitudinal measurement invariance. Aust. N. Z. J. Psychiatry. https://doi.org/10.1177/00048674211025728, 48674211025728. Advance online publication.

Boelen, P.A., Lenferink, L.I.M., Nickerson, A., Smid, G.E., 2018. Evaluation of the factor structure, prevalence, and validity of disturbed grief in DSM-5 and ICD-11. J. Affect. Disord. 240, 79–87. https://doi.org/10.1016/j.jad.2018.07.041.

Boelen, P.A., Lenferink, L.I.M., Smid, G.E., 2019b. Further evaluation of the factor structure, prevalence, and concurrent validity of DSM-5 criteria for persistent complex bereavement disorder and ICD-11 criteria for prolonged grief disorder. Psychiatry Res. 3, 206–210. https://doi.org/10.1016/j.psychres.2019.01.006.

Boelen, P.A., Smid, G.E., 2017. The traumatic grief inventory self-report version (TGI-SR): introduction and preliminary psychometric evaluation. J. Loss Trauma 22 (3), 196–212. https://doi.org/10.1080/15325024.2017.1284488.

Boeschoten, M.A., Bakker, A., Jongedijk, R.A., Olff, M., 2014. PTSS Checklist Voor de DSM-5 (PCL-5). Arq Psychotrauma Expert Group, Diemen, the Netherlands.

Bonanno, G.A., Malgaroli, M., 2020. Trajectories of grief: comparing symptoms from the DSM-5 and ICD-11 diagnoses. Depress. Anxiety 37 (1), 17–25. https://doi.org/10.1002/da.22902.

Campbell, D.T., Fiske, D.W., 1959. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol. Bull. 56 (2), 81–105. https://doi.org/10.1037/h0046016.

Chen, F.F., 2007. Sensitivity of goodness of fit indexes to lack of measurement invariance. Struct. Equ. Model. 14 (3), 464–504. https://doi.org/10.1080/10705510701301834.

Comtesse, H., Rosner, R., 2017. Deutsche Übersetzung des Traumatic Grief Inventory Self-Report. Unveröffentlichtes Manuskript, Katholische Universität Eichstätt-Ingolstadt.

Cozza, S.J., Fisher, J.E., Mauro, C., et al., 2016. Performance of DSM-5 persistent complex bereavement disorder criteria in a community sample of bereaved military family members. Am. J. Psychiatr. 173 (9), 919–929. https://doi.org/10.1176/appi.ajp.2016.15111442.

de Graaf, L.E., Gerhards, S.A., Arntz, A., Riper, H., Metsemakers, J.F., Evers, S.M., Huibers, M.J., 2009. Clinical effectiveness of online computerised cognitive-behavioural therapy without support for depression in primary care: randomised trial. Br. J. Psychiatry 195 (1), 73–80. https://doi.org/10.1192/bjp.bp.108.054429.

Eisma, M.C., Janshen, A., Lenferink, L.I.M., 2022. Content overlap analyses of ICD-11 and DSM-5 prolonged grief disorder and prior criteria-sets. Eur. J. Psychotraumatol. https://doi.org/10.1080/20008198.2021.2011691.

Eisma, M.C., Rosner, R., Comtesse, H., 2020. ICD-11 prolonged grief disorder criteria: turning challenges into opportunities with multiverse analyses. FrontiersPsychiatry 11 (752). https://doi.org/10.3389/fpsyt.2020.00752.

Elis, B., 1989. Differential item functioning: implication for test translation. J. Appl. Psychol. 74 (6), 912–921. https://doi.org/10.1037/0021-9010.74.6.912.

First, M.B., Reed, G.M., Hyman, S.E., Saxena, S., 2015. The development of the ICD-11 clinical descriptions and diagnostic guidelines for mental and behavioural disorders. World Psychiatry 14, 82–90.

First, M.B., Wolfgang, G., Maj, M., Stein, D., Kogan, C.S., Saunders, J.B., Reed, G.M., 2021. An organization- and category-level comparison of diagnostic requirements for mental disorders in ICD-11 and DSM-5. World Psychiatry 20 (1), 34–51. https://doi.org/10.1002/wps.20825.

Gloster, A.T., Block, V.J., Klotsche, J., Villanueva, J., Rinner, M.T.B., Benoy, C., Walter, M., Karekla, M., Bader, K., 2021. Psy-flex: a contextually sensitive measure of psychological flexibility. J. Context. Behav. Sci. 22, 13–23. https://doi.org/10.1016/j.jcbs.2021.09.001.

Gräfe, K., Zipfel, S., Herzog, W., Löwe, B., 2004. Screening psychischer Störungen mit dem "Gesundheitsfragebogen für patienten (PHQ-D)": ergebnisse der deutschen validierungsstudie = screening for psychiatric disorders with the patient health questionnaire (PHQ) results from the German validation study. Diagnostica 50 (4), 171–181.

Haneveld, J., Rosner, R., Vogel, A., Kersting, A., Rief, W., Steil, R., Comtesse, H., 2022. Same name, same content? Evaluation of DSM-5-TR and ICD-11 prolonged grief criteria. J. Consult. Clin. Psychol. 90 (4), 303–313. https://doi.org/10.1037/ccp0000720.

Hayes, A.F., Coutts, J.J., 2020. Use omega rather than Cronbach's alpha for estimating reliability. But…. Commun. Methods Meas. 14, 1–24. https://doi.org/10.1080/19312458.2020.1718629.

Heeke, C., Kampisiou, C., Niemeyer, H., Knaevelsrud, C., 2019. A systematic review and meta-analysis of correlates of prolonged grief disorder in adults exposed to violent loss. Eur. J. Psychotraumatol. 10 (1), 1583524. https://doi.org/10.1080/20008198.2019.1583524.

Kaplan, D., 1989. Model modifcation in covariance structure analysis: application of the expected parameter change statistic. Multivar. Behav. Res. 24 (3), 285–305. https://doi.org/10.1207/s15327906mbr2403_2.

Killikelly, C., Zhou, N., Merzhvynska, M., Stelzer, E.-M., Dotschung, T., Rohner, S., Sun, L.H., Maercker, A., 2020. Development of the international prolonged grief disorder scale for the ICD-11: measurement of core symptoms and culture items adapted for Chinese and German-speaking samples. J. Affect. Disord. 277, 568–576. https://doi.org/10.1016/j.jad.2020.08.057.

Kline, R.B., 2005. Principles and Practice of Structural Equation Modeling, (2nd ed.). Guilford Press.

Kokou-Kpolou, C.K., Lenferink, L.I.M., Brunnet, A.E., Park, S., Megalakaki, O., Boelen, P., Cénat, J.M., 2022. The ICD-11 and DSM-5-TR prolonged grief criteria: validation of the traumatic grief inventory-self report plus using exploratory factor analysis and item response theory. Clin. Psychol. Psychother. 29 (6), 1950–1962. https://doi.org/10.1002/cpp.2765.

Kroenke, K., Spitzer, R.L., Williams, J.B.W., 2001. The PHQ9 validity of a briefdepression severity measure. J. Gen. Intern. Med. 16, 606–613. https://doi.org/10.1046/j.1525-1497.2001.016009606.x.

Krüger-Gottschalk, A., Knaevelsrud, C., Rau, H., Dyer, A., Schäfer, I., Schellong, J., Ehring, T., 2017. The German version of the posttraumatic stress disorder checklist for DSM-5 (PCL-5): psychometric properties and diagnostic utility. BMC Psychiatry 17.

Lenferink, L., Boelen, P., Smid, G., Paap, M., 2021. The importance of harmonising diagnostic criteria sets for pathological grief. Br. J. Psychiatry 219 (3), 473–476. https://doi.org/10.1192/bjp.2019.240.

Lenferink, L.I.M., Nickerson, A., De Keijser, J., Smid, G.E., Boelen, P.A., 2020. Trajectories of grief, depression, and posttraumatic stress in disaster-bereaved people. Depress. Anxiety 37, 35–44. https://doi.org/10.1002/da.22850.

Lenferink, L.I.M., Eisma, M.C., Smid, G.E., de Keijser, J., Boelen, P.A., 2022. Valid measurement of DSM-5 persistent complex bereavement disorder and DSM-5-TR and ICD-11 prolonged grief disorder: the traumatic grief inventory-self report plus (TGI-SR+). Compr. Psychiatry 112, 152281. https://doi.org/10.1016/j.comppsych.2021.152281.

Lim, G., Tam, W., Lu, Y., Ho, C., Zhang, M., Ho, R., 2018. Prevalence of depression in the community from 30 countries between 1994 and 2014. Sci. Rep. 8 (1), 2861. https://doi.org/10.1038/s41598-018-21243-x.

Lobb, E.A., Kristjanson, L.J., Aoun, S.M., Monterosso, L., Halkett, G.K., Davies, A., 2010. Predictors of complicated grief: a systematic review of empirical studies. Death Stud. 34 (8), 673–698. https://doi.org/10.1080/07481187.2010.496686.

Lundorff, M., Holmgren, H., Zachariae, R., Farver-Vestergaard, I., O'Connor, M., 2017. Prevalence of prolonged grief disorder in adult bereavement: a systematic review and meta-analysis. J. Affect. Disord. 212, 138–149. https://doi.org/10.1016/j.jad.2017.01.030.

Lutz, W., Neu, R., Rubel, J.A., 2019. Evaluation und Effekterfassung in der Psychotherapie. Hogrefe, Göttingen.

Maciejewski, P.K., Maercker, A., Boelen, P.A., Prigerson, H.G., 2016. 'Prolonged grief disorder' and 'persistent complex bereavement disorder', but not 'complicated grief', are one and the same diagnostic entity: an analysis of data from the Yale bereavement study. World Psychiatry 15, 266–275. https://doi.org/10.1002/wps.20348.

Maercker, A., Brewin, C.R., Bryant, R.A., Cloitre, M., van Ommeren, M., Jones, L.M., et al., 2013. Diagnosis and classification of disorders specifically associated with stress: proposals for ICD-11. World Psychiatry 12 (3), 198–206. https://doi.org/10.1002/wps.20057.

Marta-Pedroso, C., Freitas, H., Domingos, T., 2007. Testing for the survey mode effect on contingent valuation data quality: a case study of web based versus in-person interviews. Ecol. Econ. 62 (3–4), 388–398.

Mauro, C., Shear, M., Reynolds, C., et al., 2017. Performance characteristics and clinical utility of diagnostic criteria proposals in bereaved treatment-seeking patients. Psychol. Med. 47 (4), 608–615. https://doi.org/10.1017/S0033291716002749.

Meade, A.W., Johnson, E.C., Braddy, P.W., 2008. Power and sensitivity of alternative fit indices in tests of measurement invariance. J. Appl. Psychol. 93 (3), 568–592.

Milfont, T.L., Fischer, R., 2010. Testing measurement invariance across groups: applications in cross-cultural research. Int. J. Psychol. Res. 3 (3), 111–130.

Mundt, J.C., Marks, I.M., Shear, M.K., Greist, J.H., 2002. The work and social adjustment scale: a simple measure of impairment in functioning. Br. J. Psychiatry 180, 461–464.

Muthén, L.K., Muthén, B.O., 1998–2019. Mplus User's guide. Muthén & Muthén, Los Angeles, CA.
Muthén, L.K., Muthen, B.O., 2023. Chi-Square Difference Testing Using the Satorra-Bentler Scaled Chi-Square. Accessed on March 10 via: https://www.statmodel.com/chidiff.shtml.
Newson, R.S., Boelen, P.A., Hek, K., Hofman, A., Tiemeier, H., 2011. The prevalence and characteristics of complicated grief in older adults. J. Affect. Disord. 132 (1–2), 231–238. https://doi.org/10.1016/j.jad.2011.02.021.
Nielsen, M.K., Carlsen, A.H., Neergaard, M.A., Bidstrup, P.E., Guldin, M.B., 2019. Looking beyond the mean in grief trajectories: a prospective, population-based cohort study. Soc. Sci. Med. 232, 460–469. https://doi.org/10.1016/j.socscimed.2018.10.007.
Prigerson, H.G., Boelen, P.A., Xu, J., Smith, K.V., Maciejewski, P.K., 2021. Validation of the new DSM-5-TR criteria for prolonged grief disorder and the PG-13-revised (PG-13-R) scale. World Psychiatry 20 (1), 96–106. https://doi.org/10.1002/wps.20823.
Prigerson, H.G., Horowitz, M.J., Jacobs, S.C., Parkes, C.M., Aslan, M., Goodkin, K., et al., 2009. Prolonged grief disorder: psychometric validation of criteria proposed for DSM-V and ICD-11. PLoS Med. 6 (8), e1000121 https://doi.org/10.1371/journal.pmed.1000121.
Putnick, D.L., Bornstein, M.H., 2016. Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. Dev. Rev. 41, 71–90. https://doi.org/10.1016/j.dr.2016.06.004.
Rosenthal, R., 1991. Meta-analytic Procedures for Social Research, 2nd ed. Sage, Newbury Park, CA.
Rosner, R., Comtesse, H., Vogel, A., Doering, B.K., 2021. Prevalence of prolonged grief disorder. J. Affect. Disord. 287, 301–307. https://doi.org/10.1016/j.jad.2021.03.058.
Rubio, D.M.G., Berg-Weger, M., Tebb, S.S., Rauch, S.M., 2008. Validating a measure across groups. The use of MIMIC models in scale development. J. Soc. Serv. Res. 29 (3), 53–67. https://doi.org/10.1300/J079V29N03_03.
Shear, M.K., Simon, N., Wall, M., Zisook, S., Neimeyer, R., Duan, N., et al., 2011. Complicated grief and related bereavement issues for DSM-5. Depress Anxiety 28 (2), 103–117. https://doi.org/10.1002/da.20780.
Shevlin, M., Butter, S., McBride, O., et al., 2022. Measurement invariance of the patient health questionnaire (PHQ-9) and Generalized Anxiety Disorder Scale (GAD-7) across four European countries during the COVID-19 pandemic. BMC Psychiatry 22 (154). https://doi.org/10.1186/s12888-022-03787-5.
van de Schoot, R., Lugtig, P., Hox, J., 2012. A checklist for testing measurement invariance. Eur. J. Dev. Psychol. 9 (4), 486–492. https://doi.org/10.1080/17405629.2012.686740.
van Steenbergen-Weijenburg, K.M., de Vroege, L., Ploeger, R.R., et al., 2010. Validation of the PHQ-9 as a screening instrument for depression in diabetes patients in specialized outpatient clinics. BMC Health Serv. Res. 10, 235. https://doi.org/10.1186/1472-6963-10-235.
Weathers, F.W., Litz, B.T., Keane, T.M., Palmieri, P.A., Marx, B.P., Schnurr, P.P., 2013. The PTSD Checklist for DSM-5 (PCL-5). National Center for PTSD at www.ptsd.va.gov. Scale available from the.
World Health Organization (WHO), 2018. ICD-11 Prolonged Grief Disorder Criteria. WHO. Retrieved from: https://icd.who.int/browse11/l-m/en#/http://id.who.int/icd/entity/1183832314. Accessed on December 3, 2021.