

Assessing the Validity and Reliability of the International Anxiety and Depression Questionnaire in Two Bereaved National Samples

Philip Hyland^{1*}, Enya Redican², Thanos Karatzias³, Mark Shevlin²

¹Department of Psychology, Maynooth University, County Kildare, Ireland; ²Department of Psychology, Ulster University, Northern Ireland;

³Department of Health and Social Care, Napier University, Edinburgh, Scotland

ABSTRACT

Background: The International Anxiety Questionnaire (IAQ) and International Depression Questionnaire (IDQ) have been developed as self-report measures of ICD-11 Generalized Anxiety Disorder (ICD-11 GAD) and ICD-11 Single Episode Depressive Disorder (ICD-11 DD). This study tested the psychometric properties of the IAQ and IDQ in two samples of bereaved adults from the United Kingdom (UK) and the Republic of Ireland.

Objective: Confirmatory Factor Analysis (CFA) was used to test the combined dimensionality and measurement invariance of the IAQ and IDQ across the UK (n=1,012) and Irish (n=1,011) samples. Differential item functioning was tested using Multiple Indicator Multiple Cause (MIMIC) modelling based on country, age, and sex while convergent validity was assessed using bivariate correlations. Prevalence rates of ICD-11 GAD and ICD-11 DD were also estimated.

Results: The CFA results supported a correlated two-factor model in both samples. The UK had higher latent variable means for depression and anxiety, older adults had lower latent variable means for depression and anxiety, and females had higher latent variable means for anxiety. The MIMIC model showed that the IDQ item "Had recurrent thoughts of death or suicide" showed DIF and the effect was small. Internal reliability of the scales was high while scores on the IAQ and IDQ were highly correlated with scores on external measures of anxiety, depression, and prolonged grief disorder. The IAQ and IDQ effectively distinguished between individuals with and without a history of mental health treatment. The prevalence of ICD-11 GAD was 18.6% and 16.1% and ICD-11 DD was 13.8% and 10.5% in the UK and Irish samples, respectively.

Conclusion: Findings of the study provide support for the validity, measurement invariance, and reliability of the IAQ and IDQ among two bereaved national samples.

Keywords: ICD-11; Generalized anxiety disorder; Single episode depressive disorder; Validity; Reliability; Measurement invariance; Prevalence

INTRODUCTION

Data from the Global Burden of Diseases (GBD) project shows that anxiety and depression are the two most common mental health disorders, with depression being the second leading cause of years lost to disability and anxiety being the eighth [1]. The eleventh version of the International Classification of Diseases [2] came into effect on 1 January 2022 for all WHO member states and included updated diagnostic guidelines for anxiety and depressive disorders. Generalised Anxiety Disorder (ICD-11 GAD: code 6B00) is defined by the presence of general apprehension or excessive worry occurring for more days than not, for at least several months, along with other problems (e.g., muscle tension, abdominal discomfort, difficulty concentrating, irritability) that

together cause significant distress or impairment. Single Episode Depressive Disorder (ICD-11 DD: code 6A70) is defined by the presence of depressed mood or diminished interest in activities occurring for most of the day, nearly every day, for at least two weeks, along with other problems (e.g., feelings of worthlessness, hopelessness, fatigue, recurrent thoughts of death) that together cause significant distress or impairment. The ICD-11 descriptions of GAD and DD are now the de facto diagnostic models of anxiety and depression.

There are many empirically supported self-report measures of anxiety (e.g., the Beck Anxiety Inventory [3], the GAD-7 [4], the State-Trait Anxiety Inventory [5]) and depression (e.g., the Centre for Epidemiological Studies-Depression scale [6], the Patient Health Questionnaire-9 [7], the Beck Depression Inventory-

Correspondence to: Philip Hyland, Department of Psychology, Maynooth University, County Kildare, Ireland, E-mail: philip.hyland@mu.ie

Received: 16-Feb-2023, Manuscript No. JDA-23-21823; **Editor assigned:** 20-Feb-2023, PreQC No. JDA-23-21823 (PQ); **Reviewed:** 06-Mar-2023, QC No. JDA-23-21823; **Revised:** 13-Mar-2023, Manuscript No. JDA-23-21823 (R); **Published:** 20-Mar-2023, DOI:10.35248/2167-1044.23.12.498.

Citation: Hyland P, Redican E, Karatzias T, Shevlin M (2023) Assessing the Validity and Reliability of the International Anxiety and Depression Questionnaire in Two Bereaved National Samples. J Dep Anxiety.12:498.

Copyright: © 2023 Hyland P, et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

II [8], the Hospital Anxiety and Depression Scale [9]) that are exceptionally useful for research and clinical practice, but none were designed to capture the symptoms and diagnostic requirements of ICD-11 GAD and DD. Recently developed the International Anxiety Questionnaire (IAQ) and the International Depression Questionnaire (IDQ) to do exactly that [10]. Using a nationally representative sample of adults from the general population of the United Kingdom, Shevlin et al. found that the latent structure of each measure was unidimensional, the items tapped the underlying dimensions at clinically meaningful levels, that each scale produced scores with high internal reliability, each had the capacity to discriminate between individuals with and without a history of treatment for mental health problems, and scores on both measures correlated positively and strongly with established measures of anxiety and depression symptoms. Furthermore, application of the ICD-11 diagnostic guidelines yielded prevalence estimates of 7.1% for GAD, and 7.4% for DD. Initial psychometric testing has therefore been promising, but further research is required.

In this study, we further tested the reliability and validity of the IAQ and IDQ scores in two samples of bereaved adults from the United Kingdom and the Republic of Ireland. Similar to the initial validation study of the IAQ and IDQ, the aims are to (1) describe the distribution of item and scale level scores and (2) assess construct validity using factor analytic methods. This study extends the earlier study by using data collected from a large sample of bereaved participants across two countries. We recruited bereaved participants because research has shown that the experience of the loss of a loved one increases the probability of the onset of multiple mental health problems including depression [11,12] and anxiety [13,14]. This study also addressed the important psychometric issue of measurement invariance; this refers to the extent that a scale measures the same construct, in the same way, for different people. The ability to make valid comparison of anxiety and depression, for example between different age groups or countries, is based on the assumption that the items contained within these scales operate equivalently for these different groups of interest, or that the item performance is 'invariant' [10]. There has been extensive invariance testing of other measures of anxiety and depression such as the DASS-21 [15] and the PHQ-9/GAD-7 [10].

In this study we hypothesised that (1) a correlated 2-factor model of the IAQ and the IDQ would provide acceptable fit to the sample data, (2) that scores on the IAQ and IDQ would have high internal reliability, (3) that there would be no differential item functioning associated with age, gender, or nationality (UK/Ireland), (4) that scores on the IAQ and IDQ would be strongly and positively correlated with an independent measure of anxiety and depression symptoms as well as a measure of prolonged grief disorder symptoms, and (5) the IAQ and IDQ scores would distinguish between those individuals with and without a history of treatment for mental health problems.

MATERIALS AND METHODS

Participants and procedures

Data were collected from a sample of bereaved adults from the United Kingdom (UK: N=1,012) and the Republic of Ireland (Ireland: N=1,011) using identical procedures. The survey company Qualtrics was employed to recruit participants in each nation from existing, actively managed, double-opt-in research panels via

email, SMS, or in-app notifications. Inclusion criteria were that respondents were aged 18 years or older, were residing in the UK or Ireland, respectively, could complete the survey in English, and had answered 'Yes' to the following question that screened for lifetime bereavement: "During your life have you known anyone who has died (e.g., a partner, parent, child, close friend)?" An a priori power analysis was conducted to determine the optimal sample size needed to detect prolonged grief disorder with an assumed prevalence rate of 2.4% among bereaved adults, with a precision of 1%, and a confidence level of 95%. This resulted in a required sample size of N=900. Given the size of the respective populations of the UK and Ireland and the available panel members in each nation, we set our target sample sizes for 1,000 participants in both the UK and Ireland. The UK data were collected between 19 April and 13 August 2022, and the Irish data were collected from 21 April 2022 and 12 September 2022. Ethical approval for the collection of all data was provided by the research ethics committee at Ulster University (Reference number: FCPSY-22-026-A). Quota sampling was used to select participants in a manner that resulted in the composition of the UK and Irish samples being closely aligned to the sex, age, and regional distributions of the respective populations. Key demographic information for each sample is presented in Table 1.

Table 1: Demographic characteristics of the UK and Irish samples.

Ireland (N=1011)	%	UK (N=1012)	%
Sex		Sex	
Female	52.5	Female	51.3
Male	47.3	Male	47.9
Age		Age	
18-24	9.2	18-24	9.5
25-34	20.3	25-34	20.2
35-44	20.1	35-44	19.9
45-54	18.8	45-54	18.8
55+	31.7	55+	31.7
Place of birth		Place of birth	
Ireland	76.9	UK	94.6
Living location		Living location	
City	20.5	City	23.6
Suburb	22.8	Suburb	25.6
Town	25.5	Town	32.4
Rural area	31.2	Rural area	18.4
Income		Income	
0-€19,999	24.9	0-£19,999	36.5
€20,000-€39,999	38.9	£20,000-£39,999	44
€40,000-€59,999	21.3	£40,000-£59,999	13.7
€60,000-€79,000	9.6	£60,000-£79,000	4.2
€80,000 or more	5.3	£80,000 or more	1.7
Ethnicity		Ethnicity	
Irish	76.5	British	79.8
British/Irish	11.5	British/Irish	12
Indian	0.8	Indian	1.1
Pakistani	1.7	Pakistani	1.7
Chinese	2	Chinese	0.8
Other Asian	0.4	Other Asian	0.3
African	1.6	African	0.7
Other ethnic group	8.4	Afro-Caribbean	0.1
Arab	0.2	Other ethnic group	3.4

Bangladeshi	0.1	Arab	0.2
Education		Education	
No Qualifications	0.7	No Qualifications	3.7
Finished mandatory schooling	6.3	GCSE or similar	23
Finished secondary school	22.2	A-level or similar	20.3
Technical qualification	15.2	Technical qualification	22.8
Undergraduate degree	20.4	Undergraduate degree	22.8
Diploma	12.3	Diploma	4.2
Postgraduate degree	21.7	Postgraduate degree	14.4
Other qualifications	1.3	Other qualification	2.2
Employment		Employment	
Full-time	52.3	Full-time	49.5
Part-time	16.3	Part-time	15.6
Unemployed	10.9	Unemployed	12.4
Retired	13	Retired	15.9
Student	4.3	Student	2.4
Disabled	3.2	Disabled	4.2
Religion		Religion	
Christian	71.7	Christian	55
Muslim	1.8	Muslim	3.2
Jewish	0.3	Jewish	0.6
Hindu	0.7	Hindu	1
Buddhist	0.8	Buddhist	0.4
Atheist	13.2	Atheist	24.8
Agnostic	7.1	Agnostic	8.6
Other religion	4.5	Other religion	6.3
		Sikh	0.1

ICD-11 GAD: The IAQ is a self-report measure designed to capture all diagnostic requirements for ICD-11 GAD [10]. Participants answer eight questions that assess the two 'essential' (questions 1 and 2) and six 'accompanying' (questions 3 to 8) symptoms of GAD, based on the following instruction: 'Over the last several months, how frequently have you had the following feelings, thoughts, and behaviours?' These questions are answered using a five-point Likert scale where 0='Never', 1='Only a few days', 2='Half the days', 3='Most days', and 4='Every day'. Symptom endorsement is based on responses of 3 or 4 on the Likert scale. There is an additional question measuring functional impairment ('Have these experiences caused problems in personal, family, social, educational, occupational, or other important areas of your life?') that is answered on a 'Yes' or 'No' basis. The IAQ can be used to measure symptom severity or to identify probable diagnostic status. The severity scoring method involves summing responses to the eight questions, producing possible scores ranging from 0 to 32. The diagnostic criteria for ICD-11 GAD requires that four or more symptoms be endorsed with at least one from questions 1 or 2 (i.e., the essential symptoms), and functional impairment is also endorsed.

ICD-11 DD: The IDQ is a self-report measure designed to capture all diagnostic requirements for ICD-11 DD [10]. Participants answer nine questions assessing the two 'essential' (questions 1 and 2) and seven 'accompanying' (questions 3 to 8) symptoms of DD based on the following instruction: 'Over the last two weeks, how frequently have you had the following feelings, thoughts, and behaviours?'

These questions are answered using a five-point Likert scale where 0='Never', 1='Only a few days', 2='Half the days', 3='Most days', and 4='Every day'. Symptom endorsement is based on responses of 3 or 4 on the Likert scale. There is an additional question measuring functional impairment ('Have these experiences caused problems in personal, family, social, educational, occupational, or other important areas of your life?') that is answered on a 'Yes' or 'No' basis. The IDQ can be used to measure symptom severity or to identify probable diagnostic status. The severity scoring method involves summing responses to the nine questions, producing possible scores ranging from 0 to 36. The diagnostic criteria for ICD-11 DD requires that five or more symptoms be endorsed with at least one being from question 1 or 2 (i.e., the essential symptoms), and functional impairment is also endorsed.

Anxiety and depression symptoms: The Patient Health Questionnaire-4 [16] was used as a concurrent measure of anxiety and depression symptoms. The PHQ-4 includes the two core items measuring depression symptoms from the PHQ-9 and the two core items measuring generalized anxiety symptoms from the GAD-7. As with the parent scales, the PHQ-4 includes the stem question: 'Over the last 2 weeks, how often have you been bothered by the following problems?' A four-point Likert response scale is used where 0='not at all', 1='several days', 2='more than half the days', and 3='nearly every day', and possible scores range from 0-12. Previous research has shown that the PHQ-4 produces reliable and valid scores in general population samples [17], and the internal reliability of the scale scores in the UK ($\alpha=.93$) and Irish ($\alpha=.93$) samples were excellent.

Prolonged grief symptoms: The International Prolonged Grief Disorder Scale [18] is a self-report measure aligned to the ICD-11 description of Prolonged Grief Disorder. The IPGDS includes two items measuring the core symptoms of longing for the deceased and pre-occupation with the deceased, and ten items measuring different forms of emotional pain associated with bereavement (e.g., intense feelings of sorrow, anger, guilt, being unable to accept the loss). Participants indicate the frequency of these symptoms over the past week on a five-point Likert-scale ranging from 0 ('Not at all') to 4 ('Extremely'). Scores can range from 0-48 with higher scores reflecting higher symptoms of prolonged grief. The psychometric properties of the IPGDS scale scores have been supported in several national samples [18], and the internal reliability of the scale scores in the UK ($\alpha=.94$) and Irish ($\alpha=.92$) samples were excellent.

Treatment for mental health problems: Participants were asked to provide information about their current or past treatment for a mental health problem. They received the following statement: 'Mental health difficulties are very common. It will help us understand our survey results if you would tell us whether you currently or have in the past received treatment (medication or talking therapies) for these kind of difficulties.' Three response options were provided including 'I have never received treatment for mental health problems', 'I have received treatment for mental health problems in the past', and 'I am currently receiving treatment for mental health problems'.

Data analysis

The analyses were conducted in five phases. First, descriptive statistics for the individual items and the summed scores on the IAQ and the IDQ were calculated, and cross-country differences were tested. Second, a Confirmatory Factor Analysis (CFA) model of the IAQ and the IDQ indicators was estimated to establish the

fit of a baseline model for the UK and Ireland separately. The model specified two correlated latent variables, with the IAQ items loading on an 'Anxiety' latent variable and the IDQ items loading on a 'Depression' latent variable. The data from both countries were then combined and tests of configural and metric invariance were conducted: configural invariance tests that the latent structure (i.e., a correlated two-factor model) is consistent across the groups, and metric invariance adds constraints to assess for the equality of factor loadings across the groups. Scalar invariance was not assessed as differences in the intercepts were assessed as part of the DIF analysis.

Third, a MIMIC model based on the exogenous predictor variables of country, age, and sex was specified to test for DIF on the IAQ/IDQ items. The MIMIC models provide information on: (1) the factor loadings for the IAQ/IDQ measurement model; (2) the associations between the exogenous variables and the latent variables with the regression coefficients reflecting the mean differences at the level of the latent variable across different levels of the exogenous variables; and (3) direct effects between the predictor variables and the IAQ/IDQ items, adjusting for variability on the latent variables. The presence of any direct effects is indicative of DIF. As the country and sex variables are binary in nature the regression coefficients for these are reported as 'y-standardised', where the latent variable is standardised but the predictor variables retain their binary scoring; this makes the interpretation easier as the effect reflects the mean difference between the 2 levels of the binary variable in terms of a standard deviation of a standard normal distribution. For age, the fully standardised effects are reported.

Whether a direct effect was to be included was determined using Modification Indices (MIs) [19] and the Standardized Expected Parameter Changes (SEPCs) [20,21] values. MIs indicate which path, if added to the model, would significantly improve model fit should it be freely estimated. Improved model fit is indicated by a reduction of 3.84 or more in the chi-square (this is the critical value for the chi-square for one degree of freedom, $p < .05$). In the present study, a more conservative value of 10 was used to avoid the addition of small insignificant parameters, and this is reflected in Mplus only reporting MIs greater than 10. The SEPC indicates the estimated value of a fixed parameter (in this case fixed to zero) if it were estimated, that is, the expected standardised regression coefficient. The MIs are influenced by sample size [21], and with a very large sample as in the present study this is likely to indicate that parameters with very small absolute values should be added to the model. Due to this, it has been recommended that the selection of which parameters should be added to the model be based on a combination of MIs and SEPCs [22]. Thus, in this study, a direct effect from the predictor to an IAQ/IDQ item would be added if the MI was greater than 10 and the SEPC was greater than .20. A process followed whereby the path with the largest MI/SEPC was freely estimated in the model and the model was re-estimated. This continued until there were no MIs/SEPCs greater than 10/.20.

All analyses were conducted in Mplus 8.1 [23] and all models were estimated using robust Maximum Likelihood Estimation (MLE) [24]. Numerous fit statistics were used to evaluate the goodness of fit for each model: the Chi-square, the Comparative Fit Index (CFI) [25], and the Tucker-Lewis Index (TLI) [26]. A non-significant chi-square and CFI and TLI values $\geq .90$ and $\geq .95$ were considered as good and excellent model fit. Additionally, the Root Mean Square Error of Approximation (RMSEA) [27] was reported, where a value less than .05 indicated close fit and values up to .08 indicated reasonable errors of approximation. The same cut-off values can

be used for the Standardized Root Mean Square Residual (SRMR) [28]. To compare the configural and metric models of invariance the criteria proposed were used: less than .010 change in CFI, less than .015 in RMSEA, and less than .030 for the SRMR [29].

Fourth, bivariate associations between the IAQ and IDQ summed scores and summed scores on the PHQ-4 (depression/anxiety symptoms) and the IPGDS (prolonged grief symptoms) were assessed using a Pearson correlation test. Finally, factorial between groups Analysis of Variance (ANOVA) tests, with Bonferroni post-hoc tests, were used to compare mean IAQ and IDQ scores across the two-level country variable (UK/Ireland) and three-level mental health treatment variable (Never, Past, Current). Effect sizes are reported as eta-squared values (η^2) where values up to .06 indicate a 'small effect', values from .06 to .13 indicate a 'medium' effect, and values of .14 and above indicate a 'large' effect [30].

RESULTS

Endorsement rates results

Mean item and scale scores for the IDQ and IAQ are presented in Table 2. The IDQ items with the highest mean scores in the UK sample were item 9 'experienced reduced energy or fatigue' ($M=1.55$, $SD=1.36$), item 3 'Have difficulty concentrating' ($M=1.39$, $SD=1.27$), and item 2 'Experienced less interest or pleasure from normal activities' ($M=1.28$, $SD=1.19$). The IDQ items with the highest means scores in the Irish sample were item 9 ($M=1.40$, $SD=1.24$), item 3 ($M=1.24$, $SD=1.22$), and item 8 'moved slower or felt more restless' ($M=1.07$, $SD=1.15$). The UK sample had higher mean scores on all IDQ items as compared to the Irish sample.

The IAQ items with the highest mean scores in the UK sample were item 2 'Worried a lot about different things' ($M=1.77$, $SD=1.28$), item 8 'Experienced sleep disturbances' ($M=1.69$, $SD=1.35$), and item 1 'Felt nervous or anxious' ($M=1.61$, $SD=1.28$). The IAQ items with the highest mean scores for the Irish sample were item 2 ($M=1.68$, $SD=1.24$), item 1 ($M=1.47$, $SD=1.22$), and item 7 'Been easily annoyed by different things' ($M=1.45$, $SD=1.16$). The UK sample had higher mean scores on all IAQ items as compared to the Irish sample except for item 2 and item 6 'Had difficulty concentrating'.

The mean summed IDQ and IAQ scores were significantly higher for the UK sample ($IDQ=11.19$, $SD=9.89$; $IAQ=12.32$, $SD=8.99$) as compared to the Irish sample ($IDQ=9.19$, $SD=8.61$; $IAQ=10.99$, $SD=9.19$), although there was no difference in the levels of functional impairment (Table 2).

CFA and reliability results

The CFA fit statistics in Table 2 show that the correlated two-factor model was acceptable in both samples on all fit statistics except the chi-square. The chi-square was significant for all models however this should not lead to rejection of these models as the power of chi-square tests is positively related to sample size. The standardised factor loadings were all positive, high, and statistically significant ($p < .001$), and these are reported in Table 3. The configural and metric models of invariance also indicated adequate model fit based on the differences in the CFI, RMSEA and SRMR ($\Delta CFI=.002$, $\Delta RMSEA=.001$, $\Delta SRMR=.003$) (Table 3).

The composite reliability (ω) estimates for the IAQ and IDQ were high (UK IAQ $\omega=.95$, UK IDQ $\omega=.96$, Ireland IAQ $\omega=.95$, Ireland IDQ $\omega=.95$).

The data from the UK and Ireland were combined and binary variables representing country and gender, and a continuous variable representing age, was added as predictors of the anxiety and depression latent variables. The standardised regression coefficients from the country variable to the latent variable indicated that there was significant difference (UK higher than Ireland) in the factor means for depression ($\beta=.217$, $p<.001$) and anxiety ($\beta=.158$, $p<.001$). There was a significant negative effect for age on depression ($\beta=-.022$, $p<.001$) and anxiety ($\beta=-.024$, $p<.001$). Females had significantly higher latent mean scores for anxiety ($\beta=.187$, $p<.001$) than males. Collectively, these variables explained 14.1% ($p<.001$) of the variance in the depression latent variable and 18.1% ($p<.001$) for the anxiety latent variable.

The largest MI and SEPC was a direct effect between the variable representing gender and item 6 of the IDQ item ('Had recurrent thoughts of death or suicide?': MI=31.743, SEPC=-.251). This direct effect was added, and the model was re-estimated. No other MI/SPEC met the criteria for adding additional direct effects. The final model estimates show that the standardised direct effect from gender to IDQ item 6 was $\beta=-.230$ ($p<.001$) indicating that males scored higher on this item than females, when the level of depression was constant. The overall effect was small, increasing the percentage of variance explained in that item from 48.0% to 49.3%. Thus, this direct effect explained 1.3% of the variation of that item.

Associations with external variables

As shown in Table 5, the IAQ and IDQ summed scores were strongly, positively, and significantly ($ps<.001$) correlated with the summed PHQ-4 scores, and the IPGDS (prolonged grief disorder symptoms) scores in the UK and Irish samples (Tables 4 and 5).

A three-way factorial Analysis of Variance (ANOVA) was conducted to examine the influence of country and mental health treatment-seeking on mean depression and anxiety scores. For depression scores, there was a significant main effect for country ($F [1, 2022]=11.57$, $p<.001$, $\eta^2=.00$), and mental health help seeking ($F [2, 2022]=173.92$, $p<.001$, $\eta^2=.14$) and no significant interaction ($F [2, 2022]=285.069$, $p=.142$). For anxiety scores, there was no main effect for country ($F [1, 2022]=198.428$, $p=.07$), a significant main effect for mental health help seeking ($F [2, 2022]=26222.34$, $p<.001$), and the interaction was not significant ($F [2, 2022]=231.65$, $p=.152$). The mean plots are shown in Figure 1 and it is clear that

the mean IDQ and IAQ scores are highest for those who reported currently seeking mental health treatment, lowest for those who never sought treatment, and the intermediate groups were those who reported having previously sought help.

Prevalence estimates

In the UK sample, 18.6% (95% CI=16.2%, 21.0%) met criteria for ICD-11 GAD, and 13.8% (95% CI=11.7%, 16.0%) met criteria for ICD-11 DD. Of those who met requirements for ICD-11 GAD or ICD-11 DD, 47.7% (95% CI=41.1%, 54.4%) met requirements for both disorders, 15.3% (95% CI=10.5%, 20.1%) met requirements for ICD-11 DD only, and 36.9% (95% CI=30.5%, 43.3%) met requirements for ICD-11 GAD only. Significantly more females than males met criteria for ICD-11 GAD (23.7% vs. 12.8%, $\chi^2 [1]=19.88$, $p<.001$; OR=2.12 [95% CI=1.52, 2.96]) but no gender differences were observed for ICD-11 DD (14.8% vs. 12.0%, $\chi^2 [1]=1.78$, $p=.182$; OR=1.28 [95% CI=.89, 1.85]). Those who met the diagnostic requirements for ICD-11 GAD were also significantly younger than those that did not ($M=37.46$ [SD=13.06] years vs. $M=47.33$ [SD=15.87] years: $t (325.81) =8.95$, $p<.001$, $d=.64$). Likewise, those who met the diagnostic requirements for ICD-11 DD were significantly younger than those that did not ($M=38.09$ [SD 12.78] years vs. $M=46.68$ [SD=15.99] years: $t (215.463) =7.12$, $p<.001$, $d=.55$).

In the Irish sample, 16.1% (95%CI=13.9%, 18.4%) of the sample met criteria for ICD-11 GAD, and 10.5% (95%CI=8.6%, 12.4%) met criteria for ICD-11 DD. Of those who met requirements for ICD-11 GAD or ICD-11 DD, 49.4% (95% CI=42.1%, 56.8%) met requirements for both disorders, 9.4% (95%CI=5.1%, 13.8%) met requirements for ICD-11 DD only, and 41.1% (95% CI=33.9%, 48.4%) met requirements for ICD-11 GAD only. Significantly more females than males met criteria for ICD-11 GAD (14.3% vs. 6.3%, $\chi^2 [1]=17.28$, $p<.001$; OR=2.95 [95% CI=1.60, 3.88]) and ICD-11 DD (22.6% vs. 9.0%, $\chi^2 [1]=34.37$, $p<.001$; OR=2.49 [95% CI=2.03, 4.29]). Those who met the diagnostic requirements for ICD-11 GAD were also significantly younger than those that did not ($M=37.07$ [SD=11.97] years vs. $M=46.63$ [SD=15.64] years: $t (1009)=7.39$, $p<.001$, $d=.63$). Likewise, those who met the diagnostic requirements for ICD-11 DD were significantly younger than those that did not ($M=37.95$ [SD=12.71] years vs. $M=45.92$ [SD=15.59] years: $t (1009)=5.07$, $p<.001$, $d=.52$).

Table 2: Mean and endorsement rates (%) for IDQ and IAQ.

Item	IDQ				Item	IAQ			
	Mean (SD)		t (df)	p		Mean (SD)		t (df)	p
UK	Ire	UK			Ire	UK	Ire		
1	1.25 (1.20)	1.05 (1.12)	3.93 (2012.30)	<.001	1	1.61 (1.28)	1.47 (1.22)	2.55 (2015.68)	0
2	1.28 (1.19)	1.03 (1.08)	4.83 (2001.79)	<.001	2	1.77 (1.28)	1.68 (1.24)	1.63 (2018.81)	0
3	1.39 (1.27)	1.24 (1.22)	2.81 (2017.66)	0	3	1.50 (1.28)	1.35 (1.21)	2.82 (2014.84)	0
4	1.20 (1.31)	0.95 (1.20)	4.47 (2004.19)	<.001	4	1.22 (1.28)	1.01 (1.18)	3.72 (2007.44)	<.001
5	1.22 (1.34)	0.98 (1.21)	4.20 (2002.67)	<.001	5	1.46 (1.29)	1.20 (1.21)	4.75 (2012.55)	<.001
6	0.79 (1.19)	0.52 (0.97)	5.61 (1942.11)	<.001	6	1.51 (1.30)	1.40 (1.24)	1.94 (2015.58)	0
7	1.24 (1.29)	0.95 (1.13)	5.46 (1984.33)	<.001	7	1.55 (1.27)	1.45 (1.16)	1.96 (2004.84)	0
8	1.26 (1.30)	1.07 (1.15)	3.44 (1990.84)	<.001	8	1.69 (1.35)	1.42 (1.22)	4.63 (2002.27)	<.001

9	1.55 (1.36)	1.40 (1.24)	2.71 (2003.95)	0					
Total	11.19 (9.89)	9.19 (8.61)	4.85 (1983.71)	<.001	Total	12.32 (8.99)	10.99 (9.19)	3.48 (2008.74)	<.001
FI	.38 (.49)	.35 (.48)	1.46 (2020.41)	0.1	FI	.39 (.48)	.36 (.48)	1.13 (2020.70)	0

Note: FI=Functional Impairment.

Table 3: Fit statistics for the correlated two-factor model and tests of invariance for the IAQ/IDQ Items.

Model	χ^2	df	p	CFI	TLI	RMSEA (90% CI)	SRMR
CFA							
UK	750	118	<.001	0.9	0.9	.073 (.068-.078)	0
Ireland	843	118	<.001	0.9	0.9	.078 (.073-.083)	0
Invariance							
Configural	1594	236	<.001	0.9	0.9	.075 (.072-.079)	0
Metric	1657	251	<.001	0.9	0.9	.074 (.071-.078)	0
Difference	40.2	15	<.001	0		0	0

Note: χ^2 =Chi-square Goodness of Fit statistic; df=degrees of freedom; p=Statistical significance; CFI=Comparative Fit Index; TLI=Tucker Lewis Index; RMSEA (90% CI)=Root-Mean-Square Error of Approximation with 90% confidence intervals; SRMR=Standardized Root-Mean Square Residual; ω =omega reliability coefficient.

Table 4: Standardised factor loadings for IAQ/IDQ confirmatory factor analysis for each country.

Item	Ireland	UK
Depression		
Felt down or depressed for most of the day?	0.849	0.836
Experienced less interest or pleasure from normal activities for most of the day?	0.853	0.848
Have had difficulty concentrating?	0.831	0.854
Had feelings of worthlessness or guilt?	0.861	0.858
Felt hopeless?	0.863	0.871
Had recurrent thoughts of death or suicide?	0.704	0.66
Have had changes in appetite or sleep?	0.82	0.815
Moved slower or felt more restless?	0.865	0.853
Experienced reduced energy or fatigue?	0.829	0.826
Anxiety		
Felt nervous or anxious?	0.869	0.847
Worried a lot about different things?	0.862	0.852
Felt physically tense or agitated?	0.899	0.885
Felt your heart racing, difficulty breathing, stomach discomfort, or dry mouth?	0.838	0.815
Felt 'on edge'?	0.863	0.882
Had difficulty concentrating?	0.855	0.878
Been easily annoyed by different things?	0.84	0.824
Experienced sleep disturbances?	0.759	0.741
Factor correlation	0.897	0.876

Table 5: Bivariate correlations for the IAQ and IDQ.

Samples	PHQ-4	IPGDS
UK sample (N=1,012)		
IAQ summed scores	0.82***	0.60***
IDQ summed scores	0.82***	0.69***
Irish sample (N=1,011)		
IAQ summed scores	0.83***	0.54***
IDQ summed scores	0.82***	0.61***

Note: Statistical significance=* p<0.05, ** p<0.01, *** p<0.0010.

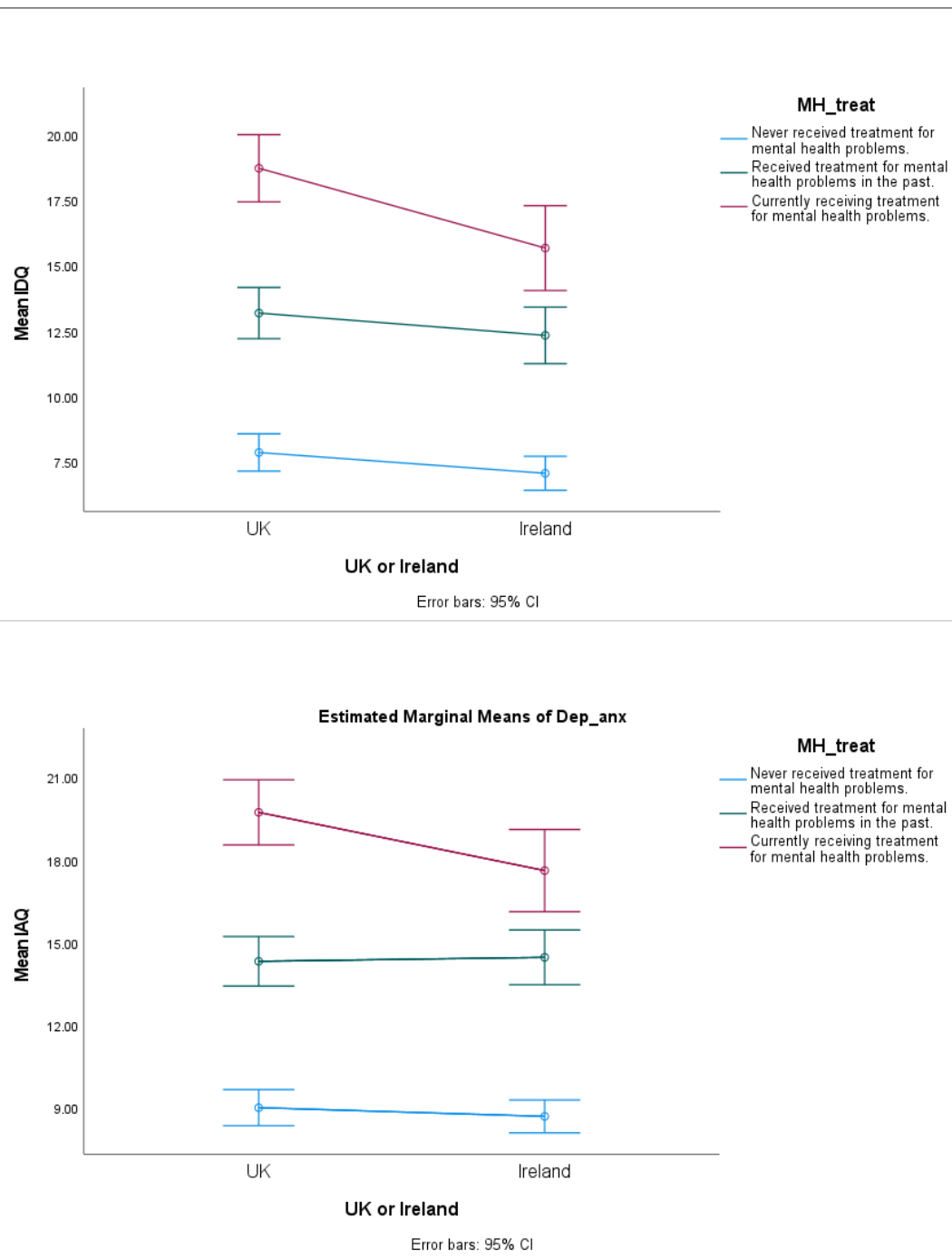


Figure 1: Mean plots for IDQ and IAQ scores by country and mental health treatment status. **Note:** (—) Never received treatment for mental health problems, (—) Received treatment for mental health problems in the past, (—) Currently receiving treatment for mental health problems.

DISCUSSION

The current study’s objectives were to examine the: (1) combined dimensionality of the IAQ and the IDQ, (2) internal reliability of the IAQ and IDQ scores, (3) differential item functioning according to age, gender, and nationality, (4) convergent validity of the IAQ and IDQ scores, (5) ability of the IAQ and IDQ scores to distinguish between people who have and have not previously received treatment for mental health issues, and to (6) estimate prevalence rates of ICD-11 GAD and ICD-11 DD and determine whether these prevalence rates varied according to sex and age.

Extending findings from the initial development and validation

study of the IAQ and IDQ, the current study demonstrated that a correlated two-factor model, in which all IDQ items loaded on a ‘Depression’ latent variable and all IAQ items loaded on an ‘Anxiety’ latent variable provided the best fit to the data in both samples. Similar to, all factor loadings were strong and statistically significant and both the IAQ and IDQ demonstrated excellent internal consistency. Significant differences in latent variable means for depression and anxiety were observed according to country, age, and gender. Specifically, the UK sample had higher latent variable means for depression and anxiety, older adults had lower latent variable means for depression and anxiety, and females had higher latent variable means for anxiety.

The UK sample generally reported higher IDQ/IAQ scores at the item and total scale level, and the MIMIC model also indicated that the latent means were higher. Recent research reported no difference in population levels of depression and anxiety between UK and Ireland [10]. The difference may be attributable to cultural differences in relation to bereavement and loss. In Ireland, it is customary for the community to come together to support the grieving individuals and share their burdens [31]. Specifically, it is common to have a wake for at least two to three days in the family home where close ones can come to mourn their loss, and this is typically followed by a funeral, a burial or cremation service, and then a repast. Different mourning customs generally apply in the UK, where the deceased is kept at a mortuary until a funeral ceremony and a repast takes place. It could be argued that the traditions implemented in Ireland are somewhat more “community-focused”, with it being well-established that social support is strongly linked to bereavement outcomes [32,33]. Moreover, Ireland is more ethnically and religiously homogenous than the UK such that the Irish population is predominantly Christian, or influenced by Christian traditions related to death, whereas the UK is comprised of more varied religious-based or religiously-influenced traditions related to death. Future research may benefit from examining how cultural and ethnic differences in grieving and loss in the UK and Ireland relate to bereavement outcomes, and how cultural and ethnic customs related to death influence psychological responses to bereavement.

Regarding the DIF analysis, after controlling for the overall level of depression, males were found to score higher on the IDQ item ‘Had recurrent thoughts of death or suicide’; however the size of the effects were small and unlikely to contribute to incorrect inferences surrounding gender differences in IDQ scores. Overall, our findings indicate that both the IDQ and IAQ are measuring depression and anxiety, respectively, in the same manner, across country, sex, and age among the bereaved population. These findings largely mirror a prior study examining the measurement invariance of the PHQ-9 and GAD-7 in the general populations of the UK and Ireland [10]. Supporting the convergent validity of the IAQ and IDQ, findings from the present study highlighted a significant association between the latent IAQ and IDQ scores and scores on the PHQ-4, a brief measure of GAD and MDD as defined by DSM-IV (APA.). These results reflect those of Shevlin et al. who also found that summed scores on the IAQ and IDQ were strongly associated with scores on DSM-IV based measures of GAD and MDD, respectively. The strong association between the IAQ and IDQ scores and an independent measure of prolonged grief symptoms (i.e., IPGDS) aligns with findings from a recent systematic review which highlighted a high degree of co-occurrence among the symptoms of anxiety, depression, and prolonged grief disorder [34]. Should the IAQ and IDQ be operating in their intended manner, it would be expected that both measures could accurately distinguish between people who have and have not previously received treatment for mental health issues. Indeed, similar to Shevlin et al. the current study found significant differences in IAQ and IDQ scores depending on mental health treatment-seeking status such that those who were currently receiving mental health treatment had the highest average IAQ and IDQ scores, while those who had never received mental health treatment had the lowest average IAQ and IDQ scores. There was no interaction effect between country and mental health treatment seeking status on average IAQ and IDQ scores, suggesting that the IAQ and IDQ effectively discriminates among different levels of

anxiety and depression symptom severity irrespective of country of residence.

The final aim of the present study was to determine the prevalence of ICD-11 GAD and ICD-11 DD and whether prevalence rates varied according to sex and age. Findings demonstrate that the prevalence of ICD-11 GAD was 18.6% and 16.1% in the UK and Irish samples, respectively and the prevalence of ICD-11 DD was 13.8% and 10.5% in the UK and Irish samples, respectively. These rates are higher than those observed by Shevlin et al. in their large community sample of adults from the UK where the prevalence of ICD-11 GAD and ICD-11 DD were 7.1% and 7.4%, respectively. It is well-established that the loss of a loved one can trigger the onset or worsening of depression and anxiety [12,35-37], and hence it is likely that the bereaved nature of the participants in the present study explains these higher rates. The co-occurrence of depression and anxiety disorders is well documented [38-40], and the results of the current study show that more people in the UK and Irish samples met the criteria for both ICD-11 GAD and ICD-11 DD rather than for either disorder alone. These findings support recommendations from who suggested that the IAQ and IDQ are appropriate measures for the assessment of ICD-11 mixed depressive and anxiety disorder (MDAD; 6A73). Contrary to previous research [39], which found that more people met the criteria for both ICD-11 DD and ICD-11 GAD than for either disorder alone, the current study demonstrated that the percentage of people who met the criteria for ICD-11 GAD alone was comparable to the percentage of people who met the criteria for both disorders. It would seem that among those who have experienced a loss, “pure” anxiety is equally as prevalent as combined anxiety and depression. Indeed, it is widely acknowledged that anxiety is a common response to bereavement due to the separation from a significant other, confrontation with one’s mortality, and exposure to stressors such as financial adversity.

Sex differences in internalizing disorders are well-established [41], with findings from the current study highlighting a higher prevalence of ICD-11 GAD among females in both samples and a higher prevalence of ICD-11 DD among females in the Irish sample only. Different factors have been proposed to account for sex differences in internalizing disorders including genetic, neurobiological, neurodevelopmental, environmental, and psychological. The absence of such an effect for ICD-11 DD in the UK sample is unsurprising given that prior research reported no sex differences in ICD-11 GAD nor ICD-11 DD in the UK general population [39]. Consistent with existing research [42], and our earlier findings regarding mean IAQ and IDQ scores, findings from the current study illustrated a higher prevalence of ICD-11 GAD and ICD-11 DD among younger adults as compared to older adults.

This study has a number of limitations. First, participants were recruited using a non-probability sampling method and hence the degree to which the samples are representative of the UK and Irish bereaved populations is uncertain. That being said, the composition of the final samples reflected the sex, age, and regional distributions of the respective nations. Second, further research is required in clinical samples where the prevalence of ICD-11 GAD and ICD-11 DD are likely to be much higher. Given that one focus of the ICD-11 is on improving the clinical utility of psychiatric diagnoses globally, replication of this study’s procedures across low- and middle-income countries is crucial.

CONCLUSION

In conclusion, this study provides further evidence that the IAQ and IDQ are psychometrically sound measures of ICD-11 GAD and DD. Our results show that the IAQ and IDQ generate reliable and valid scores irrespective of sex, age, and nationality. Our results also show that ICD-11 GAD and DD are prevalent in a substantial minority of bereaved people in the UK and Irish populations, underscoring the mental health effects of bereavement. It should be noted that a clinician-administered measure of ICD-11 GAD and DD has not yet been developed. Third, despite these samples being comprised of bereaved adults, it was not possible to ascertain whether ICD-11 GAD and ICD-11 DD stemmed from the bereavement specifically or from other factors. Finally, the samples used in the current study were drawn from affluent, English-speaking Western European nations.

REFERENCES

- GBD 2019 Mental Disorders Collaborators. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet Psychiatry*. 2022;9(2):137-50.
- World Health Organisation (2022). ICD-11 for mortality and morbidity statistics.
- Beck AT, Epstein N, Brown G, Steer R. Beck anxiety inventory. *J Consult Clin Psychol*. 1993.
- Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med*. 2006;166(10):1092-7.
- Spielberger CD. State-trait anxiety inventory: Bibliography, 2nd edn. Palo Alto.
- Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psychol Meas*. 1977 ;1(3):385-401.
- Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med*. 2001;16(9):606-13.
- Venta A, Sharp C, Hatkevich C, Rotenberg KJ. Emotional trust in mothers serves as a buffer against suicide attempts in inpatient adolescents with depressive symptoms. *J Soc Clin Psychol*. 2017;36:221-37.
- Montgomery SA, Åsberg MA. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134(4):382-9.
- Shevlin M, Butter S, McBride O, Murphy J, Gibson-Miller J, Hartman TK, et al. 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*. 2022;22(1):154.
- Pham S, Porta G, Biernesser C, Walker Payne M, Iyengar S, Melhem N, et al. The burden of bereavement: Early-onset depression and impairment in youths bereaved by sudden parental death in a 7-year prospective study. *Am J Psychiatry*. 2018;175(9):887-96.
- Zisook S, Kendler KS. Is bereavement-related depression different than non-bereavement-related depression?. *Psychol Med*. 2007;37(6):779-94.
- Shear MK, Skritskaya NA. Bereavement and anxiety. *Curr Psychiatry Rep*. 2012;14:169-75.
- Thimm JC, Kristoffersen AE, Ringberg U. The prevalence of severe grief reactions after bereavement and their associations with mental health, physical health, and health service utilization: a population-based study. *Eur J Psychotraumatol*. 2020; 11(1):1844440.
- Scholten S, Velten J, Bieda A, Zhang XC, Margraf J. Testing measurement invariance of the Depression, Anxiety, and Stress Scales (DASS-21) across four countries. *Psychol Assess*. 2017;29(11):1376.
- Kroenke K, Spitzer RL, Williams JB, Löwe B. An ultra-brief screening scale for anxiety and depression: the PHQ-4. *Psychosomatics*. 2009;50(6):613-21.
- Löwe B, Wahl I, Rose M, Spitzer C, Glaesmer H, Wingenfeld K, et al. A 4-item measure of depression and anxiety: validation and standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J Affect Disord*. 2010;122(1-2):86-95.
- Killikelly C, Merzhvynska M, Zhou N, Stelzer EM, Hyland P, Rocha J, et al. Examination of the new ICD-11 prolonged grief disorder guidelines across five international samples. *Clinical Psychology in Europe*. 2021;3(1):1-6.
- Sörbom D. Model modification. *Psychometrika*. 1989;54(3):371-84.
- Chou CP, Bentler PM. Invariant standardized estimated parameter change for model modification in covariance structure analysis. *Multivariate Behav Res*. 1993;28(1):97-110.
- Saris WE, Satorra A, Sörbom D. The detection and correction of specification errors in structural equation models. *Sociol Methodol*. 1987:105-29.
- Kaplan D. Model modification in covariance structure analysis: Application of the expected parameter change statistic. *Multivariate Behav Res*. 1989;24(3):285-305.
- Muthén LK, Muthén B. Mplus user's guide: Statistical analysis with latent variables, user's guide. Muthén & Muthén; 2017.
- Yuan KH, Bentler PM. 5. Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociol Methodol*. 2000;30(1):165-200.
- Bentler PM. Comparative fit indexes in structural models. *Psychol Bull*. 1990;107(2):238.
- Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973;38(1):1-0.
- Steiger JH. Structural model evaluation and modification: An interval estimation approach. *Multivariate Behav Res*. 1990 Apr 1;25(2):173-80.
- Jöreskog KG, Sörbom D. LISREL 8: User's reference guide. Scientific Software International; 1996.
- Chen FF. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Modeling*. 2007;14(3):464-504.
- Cohen J. Statistical power analysis for the behavioral sciences. Routledge.
- Ronan M. Funerals in the time of Coronavirus. *Ir J Sociol*. 2021; 29(2):236-40.
- Cao X, Yang C, Wang D. The impact on mental health of losing an only child and the influence of social support and resilience. *OMEGA-Journal of Death and Dying*. 2020;80(4):666-84.
- Scott HR, Pitman A, Kozhuharova P, Lloyd-Evans B. A systematic review of studies describing the influence of informal social support on psychological wellbeing in people bereaved by sudden or violent causes of death. *BMC psychiatry*. 2020;20:1-20.
- Komischke-Konnerup KB, Zachariae R, Johannsen M, Nielsen LD, O'Connor M. Co-occurrence of prolonged grief symptoms and symptoms of depression, anxiety, and posttraumatic stress in bereaved adults: A systematic review and meta-analysis. *J Affect Disord*. 2021; 4:100140.
- Niederkrotenthaler T, Laido Z, Kirchner S, Braun M, Metzler H, Waldhör T, et al. Mental health over nine months during the SARS-CoV2 pandemic: Representative cross-sectional survey in twelve waves between April and December 2020 in Austria. *J Affect Disord*. 2022;296:49-58.
- Peisah C, Latif E, Wilhelm K, Williams B. Secrets to psychological success: why older doctors might have lower psychological distress and burnout than younger doctors. *Aging Ment Health*. 2009;13(2):300-7.

37. Zisook S, Iglewicz A, Avanzino J, Maglione J, Glorioso D, Zetumer S, et al. Bereavement: course, consequences, and care. *Curr Psychiatry Rep.* 2014;16:1-0.
38. Möller HJ, Bandelow B, Volz HP, Barnikol UB, Seifritz E, Kasper S. The relevance of 'mixed anxiety and depression' as a diagnostic category in clinical practice. *Eur Arch Psychiatry Clin Neurosci.* 2016 ;266:725-36.
39. Shevlin M, Hyland P, Butter S, McBride O, Hartman TK, Karatzias T, et al. The development and initial validation of self-report measures of ICD-11 depressive episode and generalized anxiety disorder: The International Depression Questionnaire (IDQ) and the International Anxiety Questionnaire (IAQ). *Journal of Clinical Psychology.* 2022.
40. Shevlin M, Hyland P, Nolan E, Owczarek M, Ben-Ezra M, Karatzias T. ICD-11 'mixed depressive and anxiety disorder' is clinical rather than sub-clinical and more common than anxiety and depression in the general population. *Br J Clin Psychol.* 2022;61(1):18-36.
41. Boyd A, Van de Velde S, Vilagut G, De Graaf R, Florescu S, Alonso J, et al. Gender differences in mental disorders and suicidality in Europe: results from a large cross-sectional population-based study. *J Affect Disord.* 2015;173:245-54.
42. Hobbs MJ, Anderson TM, Slade T, Andrews G. Relationship between measurement invariance and age-related differences in the prevalence of generalized anxiety disorder. *J Affect Disord.* 2014;152:306-12.