



Research paper

Evaluation of the Patient Health Questionnaire (PHQ) Stress Scale over the course of the pandemic in a large German general population sample

Julia Petersen^{a,*}, Christoph Kasinger^a, Nora Hettich-Damm^a, Elmar Brähler^{a,b},
Katja Petrowski^c, Alexander K. Schuster^d, Karl J. Lackner^e, Jochem König^f, Manfred E. Beutel^a

^a Department of Psychosomatic Medicine and Psychotherapy, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany

^b Department of Medical Psychology and Medical Sociology, Medical Faculty, University of Leipzig, Leipzig, Germany

^c Medical Psychology and Medical Sociology, Department of Psychosomatic Medicine and Psychotherapy, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany

^d Department of Ophthalmology, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany

^e Institute of Clinical Chemistry and Laboratory Medicine, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany

^f Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center of the Johannes Gutenberg-University Mainz, Mainz, Germany

ARTICLE INFO

Keywords:

Psychometrics

Patient health questionnaire

Psychological stress

Confirmatory factor analysis

Mental health trajectories

Pandemic stress

Financial stress

ABSTRACT

Background: The Patient Health Questionnaire stress scale (PHQ stress) is a brief 10-item measure of psychosocial stressors, including work, health, relationship, and financial concerns. This study examines changes in psychosocial stress during the COVID-19 pandemic and evaluates the scale's temporal stability and measurement invariance.

Methods: Longitudinal population-based data ($N = 4412$) were collected at three time points during the pandemic (baseline, 4 months later, and 1.5 years after baseline). Internal consistency was assessed using McDonald's omega, and temporal stability via Pearson correlations. Confirmatory factor analysis (CFA) was conducted to verify the unidimensional factor structure. A sensitivity analysis included pre-pandemic data (1.5 years before baseline).

Results: PHQ stress scores increased across the four time points. Internal consistency remained acceptable ($\omega = 0.77\text{--}0.79$), with moderate temporal stability ($r = 0.631\text{--}0.697$). CFA supported a stable one-factor structure ($CFI = 0.947$, $TLI = 0.932$, $RMSEA = 0.048$, $SRMR = 0.047$).

Conclusion: The PHQ stress scale demonstrated adequate reliability and temporal stability in a longitudinal context. It effectively captured variations in stress influenced by major external events such as the pandemic and economic inflation.

1. Introduction

Stress is a prominent focus within medical research, consistently being associated with both mental and physical health conditions, including depression, anxiety, sleep disorders, and cardiovascular disease (Contrada, 2025; Hammen et al., 2009; Kalmbach et al., 2018). It often arises from social and environmental factors such as personal relationships, workplace demands, financial strain, and other external influences that impact mental and emotional well-being. Numerous studies have highlighted how the COVID-19 pandemic has intensified stress levels across populations, leading to higher rates of depression, anxiety, psychological distress, as well as posttraumatic stress disorder

(Hettich et al., 2022; Jafri et al., 2022). Many individuals experienced acute stress due to health uncertainties, disruptions to daily life, and social isolation resulting from lockdowns (Brooks et al., 2020; Duby et al., 2022; Lahav, 2020). Concerns about contracting the virus and worries for loved ones' health were significant contributors to heightened anxiety (Autenrieth et al., 2024; Mercader Rubio et al., 2022). Financial strain also played a major role, as pandemic-related business closures, reduced working hours, and job losses created widespread financial insecurity (Dragano et al., 2022; Petersen et al., 2022; Richter et al., 2024). Economic inflation, exacerbated by the onset of the Ukraine war, further increased these challenges, especially for disadvantaged socioeconomic groups (Movsisyan et al., 2024). Certain

* Corresponding author at: Department of Psychosomatic Medicine and Psychotherapy, University Medical Center of the Johannes Gutenberg-University Mainz, Untere Zahlbacher Straße 8, 55131 Mainz, Germany.

E-mail address: Julia.Petersen@unimedizin-mainz.de (J. Petersen).

<https://doi.org/10.1016/j.jad.2025.119831>

Received 6 February 2025; Received in revised form 30 June 2025; Accepted 2 July 2025

Available online 5 July 2025

0165-0327/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

populations experienced a disproportionate impact from pandemic-related stress. Low-income households, in particular, faced greater vulnerability to financial stress, which often led to increased rates of depression and anxiety (Argabright et al., 2022; Wilson et al., 2020). Women and younger people also experienced higher levels of distress due to intensified responsibilities at work and home, particularly as caregiving demands increased with school closures and family illness (Dunatchik et al., 2021; Hettich-Damm et al., 2023; Thébaud et al., 2024).

As mental health is affected by the long-term stress of an era-defining crisis, it is highly important to have validated instruments of high sensitivity to changes using population-based epidemiological surveys. There are a multitude of stress questionnaires, for example, the Perceived Stress Scale (PSS-10; (Cohen et al., 1983; Klein et al., 2016)), that measures the perception of stress and how unpredictable, uncontrollable, and overloaded individuals experience their lives. The Patient Health Questionnaire – Stress Scale (PHQ stress), on the other hand, is a screening instrument for psychosocial stress inquiring how much respondents are affected by ten major psychosocial stressors, including concerns about weight and health, problems with partner, and financial worries. It is a subscale of the Patient Health Questionnaire (PHQ), which also screens for depressiveness, anxiety, eating and somatoform disorders. Despite the intensive psychometric evaluation of the other scales of the PHQ, to the best of our knowledge, there has only been one psychometric study performed on the psychosocial stress subscale: A large representative study indicated that the PHQ stress scale, which provides a one-dimensional total stress score, is a valid, practical and reliable self-report instrument for assessing the severity of psychosocial stress (Petrowski et al., 2024). However, there is a lack of data regarding its temporal stability and its test-retest validity in a German sample during the pandemic and over extended periods.

The large data sets from the prospective Gutenberg Health Study (GHS) and the Gutenberg COVID-19 Study (GCS) collected at one pre-pandemic time point and three separate time points during the pandemic and after the onset of the Ukraine war, offer a unique opportunity to attain a new and sound assessment of the temporal stability (test-retest-reliability) of psychosocial stress. Therefore, the objectives of the study are to

- (1) determine the trajectory of psychosocial stress over the course of the pandemic;
- (2) evaluate the psychometric properties of the PHQ stress scale including test-retest reliability and measurement invariance;
- (3) analyse associations with sociodemographic and psychological factors.

2. Methods

2.1. Data

Our primary data source is the Gutenberg COVID-19 Study (GCS), which was initiated following the outbreak of SARS-CoV-2. The participants who had previously taken part in the Gutenberg Health Study (GHS, (Wild et al., 2012)) before the pandemic were invited for the study of the epidemiology of COVID-19. Participants had been recruited from the city of Mainz and the county of Mainz-Bingen area through random sampling from the resident registration office, stratified by sex, age, and place of residence. The GCS baseline examination occurred between October 2020 and April 2021 (T1). The first follow-up took place from March 2021 to June 2021 (T2), and the second follow-up from May 2022 to November 2022 (T3). During these sessions at the study center, participants completed a computer-assisted personal interview and provided biomaterial samples. They also received questionnaires beforehand to be completed in preparation.

The GCS included $N = 8121$ individuals from the GHS and an additional $N = 2129$ newly recruited participants (aged 25–44). For this

study, we only included respondents with PHQ stress data available from all three GCS examination points, resulting in a sample of $N = 4412$ individuals. A sensitivity analysis was conducted on a subsample of $N = 2118$ individuals who had valid PHQ stress data at all four time points, including additional pre-pandemic data from the GHS 10-year follow-up (2017–2020; T0).

Throughout the study design, execution, and analysis, we adhered to Good Clinical Practice (GCP), Good Epidemiological Practice (GEP), and the ethical guidelines of the Declaration of Helsinki. We also followed the Federal Data Protection Act requirements. The Ethics Committee of the Rhineland-Palatinate Medical Association and the Data Protection Officer of the Johannes Gutenberg-University Medical Center Mainz reviewed all relevant documentation for both the GHS and the GCS, granting their approval. Additionally, the Rhineland-Palatinate data protection commissioner approved the sample selection through citizens' registration offices.

2.2. Measures

2.2.1. Patient Health Questionnaire Stress Scale (PHQ stress)

The Patient Health Questionnaire stress scale (Gräfe et al., 2004; Petrowski et al., 2024) is a 10-item scale that inquires about major psychosocial stressors, concerns about health and weight, work-related stress, financial worries, low sexual desire, lack of a confidant and other factors, rated on a 4-point scale (0 = not affected to 3 = very affected). The items are summarized into a sum score that ranges from 0 to 30, with higher values indicating higher psychosocial stress load. The items can be found in Table 2.

2.2.2. Psychological variables

We assessed depressiveness by the PHQ-9 (Kliem et al., 2024; Kocalevent et al., 2013; Löwe et al., 2004)), symptoms of anxiety using the Generalized Anxiety Disorder Screener (GAD-2, (Kroenke et al., 2007; Löwe et al., 2010; Spitzer et al., 2006; Wicke et al., 2022)), and somatic symptom burden using the SSS-8 (Somatic Symptom Scale; (Gierk et al., 2014; Kliem et al., 2021)) that asks about the eight most common somatic symptoms.

2.2.3. Health related variables

We also inquired about the respondents' subjective mental and physical health condition by asking them to rate their health status on a 4-point scale that ranged from 1 = very good to 4 = bad.

2.2.4. Sociodemographic variables

For sociodemographic variables, we assessed the participants' age, sex (assigned at birth), migration background, educational degree, employment status (no current employment, irregular or part-time, fulltime), having a partner (yes/no), as well as their socioeconomic status, calculated according to the index of Lampert and colleagues (Lampert and Kroll, 2009). Additionally, we categorized at-risk-of-poverty by using the definition of relative poverty by the European Union Statistics on Income and Living Conditions (EU-SILC, (Eurostat, n. d.)). According to this definition, a person is at-risk-of-poverty if his or her net equalized income is lower than 60% of the median net equalized income of all German households. We calculated the net equalized income by dividing the monthly income of a household by a weighted household size (the first adult contributed a unit of 1.0 to the weight, every additional adult from the age of 14 added an additional 0.5, and every child under 14 added a unit of 0.3 to the weight). Since the median net equalized income in 2019 was 1790€, we used a threshold of <1074€ for the categorization of people at-risk-of-poverty.

2.3. Statistical procedure

Absolute and relative frequencies were calculated for categorical variables, while means and standard deviations were determined for

continuous variables to describe the sample characteristics. Mean differences were expressed using Cohen's d effect sizes, while their significance was assessed using alpha-adjusted t -tests. McDonald's total omega (ω_t) was used to determine the internal stability of the sum scores. Temporal stability of both the sum score and individual items was estimated using Pearson's correlation coefficient. Discriminatory power coefficients were estimated through the corrected correlation between individual items and the total score.

We tested factor validity using confirmatory factor analysis (CFA) to test if all ten items of the PHQ stress loaded onto one factor. This test was performed for all three GCS time points separately using the weighted least squares mean and variance adjusted (WLSMV) estimator. To assess the goodness-of-fit we used the Comparative Fit Index (CFI), the Tucker Lewis Index (TLI), the Standardized Root Mean Squared Residual (SRMR), as well as the Root Mean Square Error of Approximation (RMSEA). A good fit is indicated by a RMSEA and SRMR of <0.050 as well as a CFI and TLI of >0.950 (Hu and Bentler, 1999; Van De Schoot et al., 2012).

We then tested for measurement invariance across time. Meredith and Teresi's (2006) sequential strategy was used for this process. We assessed configural, weak, strong, and strict invariance within this framework, by adding more restrictions as the models progressed. Initially, configural invariance allows for variations in the loadings, intercepts, and variances of the latent construct among different groups or over time, making it completely unrestricted. When assuming weak measurement invariance, we constrained the loadings to be equal across groups or time points. To test for strong measurement invariance, both factor loadings and item intercepts were constrained to equality across time points. Finally, strict measurement invariance involved assuming equal factor loadings, item intercepts, and residual variances across time points. Each invariance model was then compared to a more stringent model. The chi-square test is the most commonly used method for testing model fit, but it can be influenced by large sample sizes, leading to the false rejection of reasonable models. Therefore, we relied on the four fit indices mentioned above to compare the model fits. Full invariance can be assumed if the difference in CFI between all invariance models does not exceed the threshold of $\Delta CFI \geq 0.010$, as proposed by Chen (2007).

Finally, we conducted a sensitivity analysis to detect significant pandemic-related differences, particularly by evaluating changes in psychosocial stress from pre-pandemic to pandemic time points. Test-retest stability between these time points was assessed using Pearson correlations. All analyses were conducted using R version 4.1.1 (packages: psych (Revelle, 2023), lsr (Navarro, 2015), lavaan (Rosseel, 2012), semTools (Jorgensen et al., 2022)).

3. Results

3.1. Sample characteristics

As can be seen in Table 1, the sample of this study was 50.1% ($N = 2211$) female with a mean age of 53.49 ($SD = 13.86$). Almost one - fifth had a migration background, approximately half worked fulltime, and 2.9% were at-risk-of-poverty.

3.2. PHQ stress mean scores and item characteristics

We found that the PHQ stress score increased steadily and significantly across the time points with an initial value of $M = 4.07$ ($SD = 3.14$) at T1, a score of $M = 4.33$ ($SD = 3.34$) at T2 and a mean of 4.55 ($SD = 3.45$) at T3 (see Table 3). However, the effect sizes of these increases were relatively small, ranging from $d = 0.07$ to $d = 0.15$. The internal consistency of the PHQ stress scale was good across all time points. The largest increases between T1 and T2 were found for the items asking about the concerns for weight and physical appearance ($d = 0.15$) and lack of a confidant ($d = 0.24$). Between T2 and T3, the most relevant

Table 1

Sample characteristics at T1, $N = 4412$.

	Available data	Sample	ω_t
Sex (female, yes %)	4412	2211 (50.1%)	
Age (M, SD)	4412	53.49 (13.86)	
Age group (%)	4412		
Young (25 – 39 years)		842 (19.1%)	
Middle-aged (40 – 59 years)		2001 (45.3%)	
Old (60+ years)		1569 (35.6%)	
Migration background (yes, %)	4407	878 (19.9%)	
Partnership (yes, %)	3260	2984 (91.5%)	
Employment	4272		
Retired or unemployed		1012 (23.7%)	
irregularly or part-time		1149 (26.9%)	
fulltime		2111 (49.4%)	
SES (M, SD)	4260	15.42 (3.91)	
Equalized income (M, SD)	4196	3075.52 (1827.82)	
At-risk-of-poverty (yes, %)	4196	122 (2.9%)	
Subjective physical condition (M, SD)	4410	1.93 (0.54)	
Subjective mental condition (M, SD)	4410	2.00 (0.60)	
PHQ-Stress (M, SD)	4412	4.07 (3.14)	0.77
PHQ-9 (M, SD)	4303	4.33 (3.77)	0.86
GAD-2 (M, SD)	3553	0.82 (1.12)	0.75
SSS-8 (M, SD)	4285	5.64 (4.54)	0.82

Note. SES = Socioeconomic status; PHQ = Patient Health Questionnaire; GAD = Generalized Anxiety Disorder; SSS-8 = Somatic Symptom Scale; ω_t = McDonald's omega.

item for the increase was financial issues or concerns ($d = 0.19$). The overall increase between T1 and T3 can be mostly attributed to the items inquiring about financial issues ($d = 0.18$) and having no one to talk to ($d = 0.18$). Concerning temporal stability, we found correlation coefficients ranging from 0.63 to 0.70 for the sum score and ranging from 0.22 (something bad that happened recently) to 0.65 (stress at work or at school) for the individual items.

3.3. Measurement invariance

Confirmatory factor analyses confirmed the one factor structure of the PHQ stress scale. With the exception of the TLI, which was close to the threshold of good fit (>0.950), all fit indices demonstrated a good fit of the model. Full measurement invariance across time was able to be established by not surpassing the threshold of $\Delta CFI \geq 0.010$, as proposed by Chen (2007). All results can be found in Table 3. As individual items/stressors did not apply to the specific subgroups (i.e., men vs. women, young vs. old) in the same way (see Supplementary Table 5), we were unable to establish measurement invariance across these subgroups. We, therefore, only reported measurement invariance for the total sample.

3.4. Associations with sociodemographic and other psychological factors

Correlation analyses for PHQ stress at all three time points indicated positive associations with depressiveness (PHQ-9), anxiety (GAD-2), somatic symptom burden (SSS-8), female sex, being at-risk-of-poverty, bad subjective and mental health (see Table 4). Negative associations were found for age and socioeconomic status. The highest correlations were found for depressiveness, with the coefficient ranging from $r = 0.63$ to $r = 0.66$.

3.5. Sensitivity analysis

Lastly, we investigated a subsample of respondents with pre-pandemic data (see Table 5). We found a (insignificant) decrease between the pre-pandemic time point (T0) and the first pandemic time point (T1). Psychosocial stress was found to steadily increase during all pandemic measurement time points with only the value of T3 being

Table 2
PHQ stress items and score characteristics.

Item / score	T1		T2		T3		Δ_{2+1}		Δ_{3+2}		Δ_{3+1}		ES ₁₂₋		ES ₁₃₋		r _{tt} t1		r _{tt} t2		r _{tt} t3		r _{tt} t3-t1	
	M (SD)		M (SD)		M (SD)		M (SD)		M (SD)		M (SD)		t1		t2		t1		t1		t2		t1	
PHQ stress sum score	4.07 (3.14)		4.33 (3.34)		4.55 (3.45)		0.25*** (2.53)		0.23** (2.86)		0.48*** (2.85)		0.08		0.07		0.15		0.77		0.70		0.63	
Concern about health	0.68 (0.64)		0.69 (0.66)		0.74 (0.67)		0.01 (0.66)		0.05** (0.72)		0.05*** (0.72)		0.01		0.07		0.08		0.41		0.49		0.39	
Concern about weight and looks	0.58 (0.66)		0.68 (0.69)		0.63 (0.67)		0.10*** (0.61)		-0.05*** (0.63)		0.05*** (0.65)		0.15		0.07		0.08		0.45		0.58		0.53	
Low or no sexual desire or pleasure during intercourse	0.49 (0.64)		0.54 (0.67)		0.54 (0.67)		0.05*** (0.63)		0.01 (0.68)		0.06*** (0.68)		0.07		0.01		0.08		0.45		0.54		0.46	
Problems with spouse or (life) partner	0.38 (0.59)		0.41 (0.61)		0.44 (0.63)		0.03* (0.57)		0.03* (0.62)		0.06*** (0.62)		0.05		0.05		0.10		0.48		0.55		0.49	
Burden of caring for children, parents or other family members	0.46 (0.66)		0.44 (0.66)		0.44 (0.65)		-0.02 (0.61)		-0.01 (0.67)		-0.02 (0.67)		0.02		0.01		0.03		0.39		0.58		0.47	
Stress at work or in school	0.62 (0.73)		0.60 (0.72)		0.64 (0.74)		-0.02 (0.61)		0.04*** (0.68)		0.02 (0.70)		0.03		0.06		0.03		0.35		0.65		0.57	
Financial issues or concerns	0.21 (0.47)		0.20 (0.46)		0.30 (0.55)		-0.01 (0.41)		0.10*** (0.51)		0.09*** (0.53)		0.02		0.19		0.18		0.35		0.60		0.50	
Having no one to talk to about issues	0.21 (0.45)		0.33 (0.58)		0.30 (0.55)		0.12*** (0.54)		-0.03*** (0.56)		0.09*** (0.53)		0.24		0.06		0.18		0.48		0.48		0.50	
Something bad that happened recently	0.24 (0.55)		0.25 (0.57)		0.31 (0.60)		0.01 (0.66)		0.06*** (0.73)		0.07*** (0.72)		0.02		0.10		0.12		0.27		0.31		0.22	
Thoughts or dreams about bad events ^a	0.20 (0.47)		0.18 (0.44)		0.21 (0.48)		-0.02* (0.47)		0.04*** (0.52)		0.01 (0.52)		0.05		0.08		0.03		0.35		0.47		0.37	

Note. T1 = October 2020 to April 2021; T2 = March 2021 to June 2021; T3 = May 2022 to November 2022; Δ = difference between the time points; ES = effect size of difference of the scores between the time points measured by Cohen's d ; r_{tt} = discriminatory power calculated by corrected item-total correlations; r_{tt} = Pearson correlation coefficient for test-retest testing; ω = McDonald's Omega; ***, **, * = significance of the mean difference test between the time points (*** = p -value < 0.001, ** = p -value < 0.01, * = p -value < 0.05).

^a "Thoughts or dreams about bad events from the past, e.g., the destruction of one's own home, physical violence or a sexual act under duress".

significantly higher than the pre-pandemic value. Concerning test-retest reliability, the correlation coefficient ranged from $r_{tt} = 0.56$ and $r_{tt} = 0.61$.

4. Discussion

In order to assess the presence and the degree of specific stressors and the overall level of stress during major stressful events on a large scale, short screening methods with high sensitivity to change in the general population are needed. Major health crises such as onset of the COVID-19 pandemic, political repercussions such as a subsequent inflation, and political crises such as the military attack on the Ukraine provided a unique chance to investigate the course of psychosocial stress and its effects on mental and physical health of the German general population. Therefore, in this current study, PHQ stress was implemented several times before and during the COVID-19 pandemic in a large population-based sample. Answering the first objective of this study, we identified a trajectory of a steady and consistent increase of psychosocial stress during the pandemic. Overall, PHQ stress levels significantly rose from $M = 4.07$ ($SD = 3.14$) in late 2020/early 2021 to $M = 4.55$ ($SD = 3.45$) in the second half of 2022. These means were significantly higher than those reported in a 2005 representative survey by [Petrowski et al. \(2024\)](#), who observed a mean score of 2.54 ($SD = 2.81$) in the German population. Psychosocial stress levels have increased since then, most likely due to events such as the global financial crisis in 2008, growing concern about climate change, the COVID-19 pandemic as well as subsequent inflation and Ukraine war, possibly also reflecting sociodemographic changes such as an increase of individuals living alone or of an aging population.

The overall increase of stress in our data, between T1 and T2, was associated with a rise in the items of loneliness and the lack of someone to talk to ($M = 0.12$, $SD = 0.54$, $d = 0.24$), as well as increased concerns about weight and appearance ($M = 0.10$, $SD = 0.61$, $d = 0.18$). Meanwhile, between T2 and T3, the increase was found to be mainly driven by financial concerns, highlighted by the effect sizes of the related items ($d = 0.19$ and $d = 0.18$ for T1 and T2 compared to T3). These findings would be in line with results from other studies that have found a significant and global increase of loneliness and weight gain in the first phase of the pandemic ([Ernst et al., 2022](#); [Heinberg and Steffen, 2021](#); [Hettich et al., 2022](#)). Loneliness declined and contributed less to overall stress as lockdown restrictions were loosened. In its place, financial worries became more relevant between the second and the third time point: Following the onset of the COVID-19 pandemic, there was a global surge in inflation from 2021 until 2022. The third pandemic measurement time point was surveyed in the second half of 2022 – when the effects of inflation were particularly noticeable for consumers, especially with rising energy prices following the onset of the Ukraine war. In fact, a recent study on the compounded effects of multiple crises on mental health found that the rising costs of living, in particular, were associated with increases in physical complaints, mental distress, sleep problems, as well as a decreased life satisfaction ([Richter et al., 2024](#)). Furthermore, financial stress was found to affect mental health in general with longitudinal research showing an increase in depressiveness and generalized anxiety during the pandemic that was attributed to financial stressors such as job loss, a reduction in working hours, as well as general financial strain ([Dragano et al., 2022](#)). That being said, it must be mentioned that we observed only small effect sizes for the changes between time points, suggesting that while increases in loneliness and financial stress – consistent with previous research – may partly explain these shifts, the magnitude of these changes should not be overstated.

To investigate these developments further, we tested the PHQ stress trajectories on a subsample with pre-pandemic data. Here, we found an overall a slight, though not a significant decrease in psychosocial stress compared to pre-pandemic assessment. This decrease might have been due to changes in personal life: work related stress, for example, might have declined since remote working conditions have made commuting

Table 3

Confirmatory factor analysis and longitudinal measurement invariance in the full sample.

	CFI	TLI	RMSEA	SRMR	Δ CFI	Δ TLI	Δ RMSEA	Δ SRMR
Confirmatory factor analysis								
T1	0.947	0.932	0.048	0.047				
T2	0.956	0.943	0.048	0.047				
T3	0.950	0.936	0.053	0.052				
Longitudinal measurement invariance (full sample)								
Configural	0.951	0.937	0.051	0.045				
Metric	0.947	0.942	0.049	0.047	−0.004	0.005	−0.002	0.002
Scalar	0.939	0.941	0.049	0.050	−0.008	−0.001	0.000	0.003
Strict	0.933	0.944	0.048	0.056	−0.006	0.003	−0.001	0.006

Note. CFI = Robust Comparative Fit Index; TLI = Robust Tucker-Lewis Index; RMSEA = Robust Root Mean Square of Approximation; SRMR = Standardized Root Mean Square Residual; Δ = difference in model fits between sequential models.

Table 4

Pearson correlation of the PHQ stress sum score at all time points with relevant factors.

	T1	T2	T3
PHQ-9	0.63***	0.66***	0.66***
GAD2	0.56***	0.54***	0.56***
SSS-8	0.54***	0.54***	0.56***
Age (continuous)	−0.19***	−0.16***	−0.16***
Female sex	0.15***	0.16***	0.14***
At-risk-of-poverty	0.04*	0.05**	0.03*
SES	−0.03*	−0.03	−0.04*
Subjective physical health	0.27***	0.40***	n/a
Subjective mental health	0.45***	0.45***	n/a

Note. T1 = October 2020 to April 2021; T2 = March 2021 to June 2021; T3 = May 2022 to November 2022; n/a = not available for this measurement time point. ***, **, * = p -value (*** = p -value < 0.001, ** = p -value < 0.01, * = p -value < 0.05).

Table 5Sensitivity analysis between pre-pandemic and pandemic time points ($N = 2118$).

	T0	T1	T2	T3
Mean (SD)	4.03 (2.99)	3.86 (3.86)	4.11 (3.15)	4.44 (3.40)
Mean difference test (compared to T0)		0.060	0.453	< 0.001
Test-retest (r_{tt})		0.61	0.56	0.57

Note. T0 = 2017 to 2020; T1 = October 2020 to April 2021; T2 = March 2021 to June 2021; T3 = May 2022 to November 2022; r_{tt} = Pearson correlation coefficient to test for test-retest reliability.

to work no longer necessary and a better work-life balance might have been more easily accomplished (see for example (Shimura et al., 2021)).

The second aim of this study concerned the psychometric evaluation of the PHQ stress scale. The scale showed an acceptable internal consistency with McDonald's omega ranging from 0.77 to 0.79. These findings were consistent with previous results (Petrowski et al., 2024). These relatively low levels may be due to the work- and trauma-related items that do not apply to the whole sample as, for example, almost a quarter of this study sample is unemployed or retired. The test-retest reliability of the measurement instruments ranged from $r = 0.63$ to $r = 0.70$ in the full sample, which can be considered good. A CFA on the full sample confirmed a one factor structure of the PHQ stress scale, though it has to be noted that the TLI did not meet the criteria of good fit (> 0.950), which was also the case in a validation study on a representative sample (Petrowski et al., 2024). Nevertheless, full measurement invariance across time was able to be established as indicated by the Δ CFI that never surpassed the ≥ 0.010 threshold, as proposed by Chen (2007). Heterogeneity in the exposure to specific stressors made testing for measurement invariance between certain subgroups

unfeasible. As mentioned above, stressors such as (care) work, financial strain, trauma, and relationship issues may not affect all groups equally. For instance, as shown in Supplementary Table 5, respondents aged 60 and older reported significantly lower mean scores for concerns about appearance, problems with spouse or partner, burden of child care, work-related stress, and financial worries. These differences are largely age-related, as respondents over the age of 60 may no longer be working (and thus are not burdened by work-related stress), may not have children in their care (and therefore are not exposed to child care-related stressors), and may no longer have a significant partner (and thus are not affected by partner-related issues). Similarly, women reported higher stress levels for nearly all ten measured stressors, indicating that they may be uniquely affected by psychosocial stress. For instance, women may experience more care work-related stress while men may be affected differently due to their gendered responsibilities (Hettich-Damm et al., 2023). Furthermore, respondents at-risk-of-poverty are more likely to suffer from financial stress than more affluent respondents (Petersen et al., 2022). As a result, substantial variance between groups may hinder the establishment of measurement invariance between these subgroups and may impact the overall model fits.

Lastly, we further investigated the associations with other relevant psychological instruments as well as sociodemographic factors, as the third aim of this study. Depressiveness levels were found to be higher in our sample than reported in a standardization paper (Kocalevent et al., 2013): We found a mean of 4.33 ($SD = 3.77$) compared to the mean reported by Kocalevent et al. (2013), which was at 3.1 ($SD = 3.5$) for women and 2.7 ($SD = 3.5$) for men. Similarly, the SSS-8 levels were also significantly higher than those shown by Gierk et al. (2014): 5.64 ($SD = 0.78$) vs. 3.23 ($SD = 3.96$). Only the GAD-2 levels did not differ from those reported in literature with our sample reporting a mean of 0.82 ($SD = 1.12$) and Löwe et al. (2010) reporting a mean of 0.82 ($SD = 1.10$). These increased levels of depressiveness and somatic symptoms during the pandemic, however, have been repeatedly linked to pandemic effects and are therefore consistent with previous findings (Hettich et al., 2022). Regarding the associations to psychosocial stress, we found that depressiveness, generalized anxiety, somatic symptom burden, female sex, being at-risk-of-poverty, as well as bad subjective physical and mental health were positively associated with psychosocial stress (risk factors). Protective factors were higher age and higher socioeconomic status. These associations are in line with previous research and therefore indicate factor validity (Petrowski et al., 2024).

4.1. Limitations

Although the study benefitted from its unique database spanning over an extended period both before and during the pandemic, the community sample's characteristics present limitations. The sample had a relatively high average age of 56.6 years and was financially better off, with only 3.4% of participants being at-risk-of-poverty compared to the 14.7% reported by the German Federal Statistical Office (2023).

Consequently, the representativeness of our findings is limited, and the mean scores should be interpreted with caution. Moreover, although 10250 individuals took part in the GCS baseline examination (T1), only $N = 4412$ provided valid data across all three measurements. This raises the question of dropout patterns and panel attrition effects in longitudinal studies that may further reduce the representativeness of our findings. It is possible that individuals in poor health may become unable to participate in such examinations. Additionally, a portion of the study was conducted during the vaccination rollout, which might have influenced psychosocial stress due to varying vaccine access and the easing of lockdown measures. Lastly, we were not able to establish full measurement invariance across sex and age groups (see Supplementary Table 2), indicating that the scale may function somewhat differently between these subgroups. Further research should investigate a more appropriate factorial structure of the scale.

5. Conclusion

The PHQ stress scale proved to be a reliable tool for assessing psychosocial stress in a longitudinal study. It showed adequate temporal stability and was responsive to large-scale external influences, such as the onset and course of a pandemic and subsequent financial and political stress. Furthermore, our findings indicated that psychosocial stress – assessed using the PHQ stress scale – may not affect subgroups equally, which should be kept in mind in future analyses.

Patents and intellectual property

There are no patents to disclose.

Other activities

There are no additional activities to disclose.

CRedit authorship contribution statement

Julia Petersen: Writing – original draft, Methodology, Investigation, Formal analysis. **Christoph Kasinger:** Writing – review & editing. **Nora Hettich-Damm:** Writing – review & editing. **Elmar Brähler:** Writing – review & editing, Conceptualization. **Katja Petrowski:** Writing – review & editing. **Alexander K. Schuster:** Writing – review & editing. **Karl J. Lackner:** Writing – review & editing. **Jochem König:** Writing – review & editing. **Manfred E. Beutel:** Writing – review & editing, Supervision.

Funding

The study was funded by the European Regional Development Fund and the Ministry of Science and Health of the State of Rhineland-Palatinate (EFRE/REACT-EU, Grant No. 84007232 and No. 84009735); by the ReALity Initiative of the Life Sciences of the Johannes Gutenberg University Mainz for the establishment of a cell bank; and by the National University Medicine Research Network on Covid-19 (“NaFoUni-MedCovid19”, Grant No. 01KX2021) B-FAST for the topics “poor living conditions” and “working conditions” and their association with COVID-19 in the population. Biomaterials are stored at the BioBank Mainz of the University Medical Center Mainz.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgement

We are indebted to all study participants of the Gutenberg COVID-19 study and the Gutenberg Health Study for their contribution. We thank all study staff and researchers involved in the planning and conduct of

the study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jad.2025.119831>.

Data availability

Written consent from GCS study participants does not allow public access to the data. Access to the data in the local database is possible at any time upon request according to the ethics vote. This concept was developed with the local data protection officer and the ethics committee (local ethics committee of the Rhineland-Palatinate Medical Association, German). Interested scientists can make their requests to the Gutenberg COVID-19 Study Steering Committee (E-Mail: info-gcs@unimedizin-mainz.de).

References

- Argabright, S.T., Tran, K.T., Visoki, E., DiDomenico, G.E., Moore, T.M., Barzilay, R., 2022. COVID-19-related financial strain and adolescent mental health. *The Lancet Regional Health - Americas* 16, 100391. <https://doi.org/10.1016/j.lana.2022.100391>.
- Autenrieth, L.K., Asselmann, E., Melzig, C.A., Benke, C., 2024. Fear of COVID-19 predicts increases in anxiety, depressive symptoms, health anxiety, psychosocial distress, and loneliness: findings from a prospective two-year follow-up study. *J. Psychiatr. Res.* 177, 162–168. <https://doi.org/10.1016/j.jpsychires.2024.07.018>.
- Brooks, S.K., Webster, R.K., Smith, L.E., Woodland, L., Wessely, S., Greenberg, N., Rubin, G.J., 2020. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *Lancet* 395 (10227), 912–920. [https://doi.org/10.1016/S0140-6736\(20\)30460-8](https://doi.org/10.1016/S0140-6736(20)30460-8).
- Chen, F.F., 2007. Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Model. Multidiscip. J.* 14 (3), 464–504. <https://doi.org/10.1080/10705510701301834>.
- Cohen, S., Kamarck, T., Mermelstein, R., 1983. A global measure of perceived stress. *J. Health Soc. Behav.* 24 (4), 385. <https://doi.org/10.2307/2136404>.
- Contrada, R.J., 2025. Stress and cardiovascular disease: the role of affective traits and mental disorders. *Annu. Rev. Clin. Psychol.* <https://doi.org/10.1146/annurev-clinpsy-081423-023833>.
- Dragano, N., Reuter, M., Berger, K., 2022. Increase in Mental Disorders during the COVID-19 Pandemic—The Role of Occupational and Financial Strains. *Deutsches Ärzteblatt International, An analysis of the German National Cohort (NAKO) Study*. <https://doi.org/10.3238/arztebl.m2022.0133>.
- Duby, Z., Bunce, B., Fowler, C., Bergh, K., Jonas, K., Dietrich, J.J., Govindasamy, D., Kuo, C., Mathews, C., 2022. Intersections between COVID-19 and socio-economic mental health stressors in the lives of South African adolescent girls and young women. *Child Adolesc. Psychiatry Ment. Health* 16 (1), 23. <https://doi.org/10.1186/s13034-022-00457-y>.
- Dunachik, A., Gerson, K., Glass, J., Jacobs, J.A., Stritzel, H., 2021. Gender, parenting, and the rise of remote work during the pandemic: implications for domestic inequality in the United States. *Gen. Soc.* 35 (2), 194–205. <https://doi.org/10.1177/08912432211001301>.
- Ernst, M., Niederer, D., Werner, A.M., Czaja, S.J., Mikton, C., Ong, A.D., Rosen, T., Brähler, E., Beutel, M.E., 2022. Loneliness before and during the COVID-19 pandemic: a systematic review with meta-analysis. *Am. Psychol.* 77 (5), 660–677. <https://doi.org/10.1037/amp0001005>.
- Eurostat. (n.d.). EU statistics on income and living conditions (EU-SILC) methodology—Monetary poverty. Retrieved 25 May 2022, from [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_\(EU-SILC\)_methodology_-_monetary_poverty#Description](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_monetary_poverty#Description).
- Federal Statistical Office, 2023, May 16. Gut ein Fünftel der Bevölkerung Deutschlands von Armut oder sozialer Ausgrenzung bedroht. https://www.destatis.de/DE/Presse/Pressemittelungen/2023/05/PD23_190_63.html.
- Gierk, B., Kohlmann, S., Kroenke, K., Spangenberg, L., Zenger, M., Brähler, E., Löwe, B., 2014. The somatic symptom scale-8 (SSS-8): a brief measure of somatic symptom burden. *JAMA Intern. Med.* 174 (3), 399. <https://doi.org/10.1001/jamainternmed.2013.12179>.
- Gräfe, K., Zipfel, S., Herzog, W., Löwe, B., 2004. Screening psychischer Störungen mit dem “Gesundheitsfragebogen für Patienten (PHQ-D)”. *Diagnostica* 50 (4), 171–181. <https://doi.org/10.1026/0012-1924.50.4.171>.
- Hammen, C., Kim, E.Y., Eberhart, N.K., Brennan, P.A., 2009. Chronic and acute stress and the prediction of major depression in women. *Depress. Anxiety* 26 (8), 718–723. <https://doi.org/10.1002/da.20571>.
- Heinberg, L.J., Steffen, K., 2021. Social isolation and loneliness during the COVID-19 pandemic: impact on weight. *Curr. Obes. Rep.* 10 (3), 365–370. <https://doi.org/10.1007/s13679-021-00447-9>.
- Hettich, N., Entringer, T.M., Kroeger, H., Schmidt, P., Tibubos, A.N., Braehler, E., Beutel, M.E., 2022. Impact of the COVID-19 pandemic on depression, anxiety, loneliness, and satisfaction in the German general population: a longitudinal

- analysis. *Soc. Psychiatry Psychiatr. Epidemiol.* 57 (12), 2481–2490. <https://doi.org/10.1007/s00127-022-02311-0>.
- Hettich-Damm, N., Petersen, J., Zahn, D., Baumkoetter, R., Wild, P.S., Muenzel, T., Schuster, A.K., Koenig, J., Lackner, K., Pfeiffer, N., Beutel, M.E., Engwicht, E., 2023. Gender differences and the impact of partnership and children on quality of life during the COVID-19 pandemic. *Int. J. Public Health* 68, 1605826. <https://doi.org/10.3389/ijph.2023.1605826>.
- Hu, L., Bentler, P.M., 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct. Equ. Model. Multidiscip. J.* 6 (1), 1–55. <https://doi.org/10.1080/10705519909540118>.
- Jafri, M.R., Zaheer, A., Fatima, S., Saleem, T., Sohail, A., 2022. Mental health status of COVID-19 survivors: a cross sectional study. *Virol. J.* 19 (1), 3. <https://doi.org/10.1186/s12985-021-01729-3>.
- Jorgensen, T.D., Pornprasertmanit, S., Schoemann, A.S., Rosseel, Y., 2022. *semTools: useful tools for structural equation modeling*. In: R Package Version 0.5-6. <https://CRAN.R-project.org/package=semTools>.
- Kalmbach, D.A., Anderson, J.R., Drake, C.L., 2018. The impact of stress on sleep: pathogenic sleep reactivity as a vulnerability to insomnia and circadian disorders. *J. Sleep Res.* 27 (6), e12710. <https://doi.org/10.1111/jsr.12710>.
- Klein, E.M., Brähler, E., Dreier, M., Reinecke, L., Müller, K.W., Schmutz, G., Wölfling, K., Beutel, M.E., 2016. The German version of the Perceived Stress Scale – psychometric characteristics in a representative German community sample. *BMC Psychiatry* 16 (1), 159. <https://doi.org/10.1186/s12888-016-0875-9>.
- Kliem, S., Krieg, Y., Beller, J., Brähler, E., Baier, D., 2021. Psychometric properties of the Somatic Symptom Scale 8 (SSS-8) in a representative sample of German adolescents. *J. Psychosom. Res.* 149, 110593. <https://doi.org/10.1016/j.jpsychores.2021.110593>.
- Kliem, S., Sachser, C., Lohmann, A., Baier, D., Brähler, E., Gündel, H., Fegert, J.M., 2024. Psychometric evaluation and community norms of the PHQ-9, based on a representative German sample. *Front Psychiatry* 15, 1483782. <https://doi.org/10.3389/fpsy.2024.1483782>.
- Kocalevent, R.-D., Hinz, A., Brähler, E., 2013. Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. *Gen. Hosp. Psychiatry* 35 (5), 551–555. <https://doi.org/10.1016/j.genhosppsych.2013.04.006>.
- Kroenke, K., Spitzer, R.L., Williams, J.B.W., Monahan, P.O., Löwe, B., 2007. Anxiety disorders in primary care: prevalence, impairment, comorbidity, and detection. *Ann. Intern. Med.* 146 (5), 317. <https://doi.org/10.7326/0003-4819-146-5-200703060-00004>.
- Lahav, Y., 2020. Psychological distress related to COVID-19 – the contribution of continuous traumatic stress. *J. Affect. Disord.* 277, 129–137. <https://doi.org/10.1016/j.jad.2020.07.141>.
- Lampert, T., Kroll, L.E., 2009. Die Messung des sozioökonomischen Status in sozialepidemiologischen Studien. In: Richter, M., Hurrelmann, K. (Eds.), *Gesundheitliche Ungleichheit*. VS Verlag für Sozialwissenschaften, Wiesbaden, Germany. https://doi.org/10.1007/978-3-531-91643-9_18.
- Löwe, B., Ünlützer, J., Callahan, C.M., Perkins, A.J., Kroenke, K., 2004. Monitoring depression treatment outcomes with the Patient Health Questionnaire-9. *Med. Care* 42 (12), 1194–1201. <https://doi.org/10.1097/00005650-200412000-00006>.
- Löwe, B., Wahl, I., Rose, M., Spitzer, C., Glaesmer, H., Wingenfeld, K., Schneider, A., Brähler, E., 2010. 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.* 122 (1–2), 86–95. <https://doi.org/10.1016/j.jad.2009.06.019>.
- Mercader Rubio, I., Sánchez-López, P., Ángel, N.G., Ruiz, N.F.O., 2022. Psychological consequences of fear of COVID-19: symptom analysis of triggered anxiety and depression disorders in adolescents and young adults. *Int. J. Environ. Res. Public Health* 19 (21), 14171. <https://doi.org/10.3390/ijerph192114171>.
- Meredith, W., Teresi, J.A., 2006. An essay on measurement and factorial invariance. *Med. Care* 44 (Suppl. 3), S69–S77. <https://doi.org/10.1097/01.mlr.0000245438.73837.89>.
- Movsisyan, A., Wendel, F., Bethel, A., Coenen, M., Krajewska, J., Littlecott, H., Stöckl, H., Voss, S., Wollmershäuser, T., Rehfuess, E., 2024. Inflation and health: a global scoping review. *Lancet Glob. Health* 12 (6), e1038–e1048. [https://doi.org/10.1016/S2214-109X\(24\)00133-5](https://doi.org/10.1016/S2214-109X(24)00133-5).
- Navarro, D., 2015. *Learning Statistics with R: A Tutorial for Psychology Students and Other Beginners. (Version 0.6). R package version 0.5.1*. University of New South Wales, Sydney, Australia. <https://learningstatisticswithr.com>.
- Petersen, J., Hettich, N., Baumkötter, R., Wild, P.S., Pfeiffer, N., Münzel, T., König, J., Lackner, K.J., Beutel, M.E., 2022. The burdens of poverty during the COVID-19 pandemic. *Front. Sociol.* 7, 995318. <https://doi.org/10.3389/fsoc.2022.995318>.
- Petrowski, K., Schmalbach, B., Tibubos, A., Brähler, E., Löwe, B., 2024. Psychometric evaluation of the patient health questionnaire stress scale. *J. Affect. Disord.* 357, 37–41. <https://doi.org/10.1016/j.jad.2024.04.089>.
- Revelle, W., 2023. *Psych: Procedures for Psychological, Psychometric, and Personality Research. R package version 2.3.9*. Northwestern University, Evanston, Illinois. <http://CRAN.R-project.org/package=psych>.
- Richter, E.P., Brähler, E., Zenger, M., Stöbel-Richter, Y., Emmerich, F., Junghans, J., Krause, J., Irmscher, L., Berth, H., 2024. Compounded effects of multiple global crises on mental health: a longitudinal study of east German adults. *J. Clin. Med.* 13 (16), 4754. <https://doi.org/10.3390/jcm13164754>.
- Rosseel, Y., 2012. *lavaan: an R package for structural equation modeling*. *J. Stat. Softw.* 48 (2). <https://doi.org/10.18637/jss.v048.i02>.
- Shimura, A., Yokoi, K., Ishibashi, Y., Akatsuka, Y., Inoue, T., 2021. Remote work decreases psychological and physical stress responses, but full-remote work increases presenteeism. *Front. Psychol.* 12, 730969. <https://doi.org/10.3389/fpsyg.2021.730969>.
- Spitzer, R.L., Kroenke, K., Williams, J.B.W., Löwe, B., 2006. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch. Intern. Med.* 166 (10), 1092. <https://doi.org/10.1001/archinte.166.10.1092>.
- Thébaud, S., Hoppen, C., David, J., Boris, E., 2024. Understanding gender disparities in caregiving, stress, and perceptions of institutional support among faculty during the COVID-19 pandemic. *Soc. Sci.* 13 (4), 181. <https://doi.org/10.3390/socsci13040181>.
- 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>.
- Wicke, F.S., Krakau, L., Löwe, B., Beutel, M.E., Brähler, E., 2022. Update of the standardization of the Patient Health Questionnaire-4 (PHQ-4) in the general population. *J. Affect. Disord.* 312, 310–314. <https://doi.org/10.1016/j.jad.2022.06.054>.
- Wild, P.S., Zeller, T., Beutel, M., Blettner, M., Dugi, K.A., Lackner, K.J., Pfeiffer, N., Münzel, T., Blankenberg, S., 2012. Die Gutenberg Gesundheitsstudie. *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 55 (6–7), 824–830. <https://doi.org/10.1007/s00103-012-1502-7>.
- Wilson, J.M., Lee, J., Fitzgerald, H.N., Oosterhoff, B., Sevi, B., Shook, N.J., 2020. Job insecurity and financial concern during the COVID-19 pandemic are associated with worse mental health. *J. Occup. Environ. Med.* 62 (9), 686–691. <https://doi.org/10.1097/JOM.0000000000001962>.