cpe.psychopen.eu | ISSN 2625-3410 | February 2025 | Volume 7 | Issue 1 |

CLINICAL PSYCHOLOGY IN EUROPE

The Official Academic Journal of the European Association of Clinical Psychology and Psychological Treatment

PsychOpen GOLD

CLINICAL PSYCHOLOGY IN EUROPE

The Official Academic Journal of the
European Association of Clinical Psychology and Psychological Treatment

PsychOpen GOLD

CLINICAL PSYCHOLOGY IN EUROPE

February 2025 | Volume 7 | Issue 1

Contents

PsychOpen GOLD is a publishing service by
Leibniz Institute for Psychology Information (ZPID), Germany.
www.leibniz-psychology.org

PsychOpen GOLD

CLINICAL PSYCHOLOGY IN EUROPE

February 2025 | Volume 7 | Issue 1

Letter to the Editor, Commentary

Clinical Psychology in Europe

PsychOpen GOLD

Editorial

Check for updates

Youth, the New Adolescence: A Challenge and a Window of Opportunity for Early Mental Health Interventions

Simone Munsch[1,2], Tina In-Albon[3], Nadine Messerli-Bürgy[4]

[1] *Department of Psychology, Clinical Psychology and Psychotherapy, University of Fribourg, Fribourg, Switzerland.* **[2]** *Food Research and Innovation Center, FRIC, Cluster Food and Mental Health / Psychology, University of Fribourg, Fribourg, Switzerland.* **[3]** *Clinical Child and Adolescent Psychology and Psychotherapy, University of Mannheim, Mannheim, Germany.* **[4]** *Family and Development Research Center, Institute of Psychology, University of Lausanne, Lausanne, Switzerland.*

Clinical Psychology in Europe, 2025, Vol. 7(1), Article e16951, https://doi.org/10.32872/cpe.16951

Published (VoR): 2025-02-28

Corresponding Author: Simone Munsch, Department of Psychology, Clinical Psychology and Psychotherapy, University of Fribourg, Rue P.-A.-de-Faucigny 2, 1700 Fribourg, Switzerland. Phone: +41 26 300 7657. E-mail: simone.munsch@unifr.ch

At the beginning of 2025, we invite you to focus on some of the mental health challenges that lie ahead. Our attention is focused on youth, particularly those between the ages of 12 and 25 (Insel & Fenton, 2005), a period often referred to as the "new adolescence" (Sawyer et al., 2018). While there is evidence on an earlier onset of puberty and on continued growth well into the 20s, there is also knowledge on a delayed timing of role transitions, including the completion of education and parenthood which let to the new conceptualization of this period as the new adolescence (Sawyer et al., 2018). It is also essential to recognize that despite advances in psychotherapy, in neuroscience and psychopharmacological approaches, there are no quick solutions to stop or slow down the global increase of mental health problems in youth. In recent years, this increase has been particularly evident among children and adolescents, driven by factors such as uncertainty, crises, and conflicts in our world.

However, we all wish childhood and adolescence to be a period of positive development that is, at times, maybe feeling shaky but happy, with most young people overcoming perceived challenges and growing out of any problems. While this is still the case for the majority of children, adolescents and young adults, we cannot ignore the fact that the peak age of onset of mental disorders in the general population is around 5.5 and 15.5 years for anxiety disorders, 14.5 years for obsessive-compulsive disorder, 15.5 years for eating disorders, 19.5 years for substance use disorders, 20.5 years for schizophrenia/psychotic disorders and personality disorders, and 20.5 years for mood

disorders (Solmi et al., 2022). In other words, by the age of 24 years, 75% of all mental disorders an individual might experience in his life had already occurred for the first time (Kessler et al., 2001).

Why should the period of adolescence be expanded to 'new adolescence'? This idea is not a TikTok move or of any other social media, but an evidence-based strategy to address the challenges of today's world of children, adolescents and young adults and besides this, not to let windows of opportunity go unused that allow early intervention and prevention improving mental health. Therefore, youth is not only a critical period but a period of meaningful development that allows long-term changes in an individual's functioning. Does that get you excited? We fully understand that for you—the active members of this community —enthusiasm alone is not enough, but independently of your specialization in child, adolescent, or adult mental health, you probably agree that reliable data is needed to show that this period is not only critical for emotional, social, cognitive and occupational development (Thompson et al., 2023), but that it can be influenced in a meaningful way.

Let's start with the term "window of opportunity" which is a sensitive period in the child' and adolescent's development of neurobiological systems such as systems related to threat, reward, social cognition and stress and generally the formation of multiple complex neural systems enabling an individual to live through, and therefore perceive and process a given situation in the context of multiple, at best, coping-oriented experiences and neurobiological systems.

Research shows that in addition to data on the importance of youth for the development of mental disorders (Solmi et al., 2022), we also have data on the sensitivity of the underlying biological systems up to the age of 25 years (Uhlhaas et al., 2023). Although this is good news, it is not enough for us as psychologists because we all know that youth is a time of more complex interplay of different factors. Besides the impact of changes in parent-child and peer interactions, the adolescent's living context (degree of urbanization, economic background, access to education) but also the levels of stigmatization of mental health problems within adolescent's environment and limited access to mental health support all impact on an adolescent's mental wellbeing (Orben et al., 2020; Speyer et al., 2022; van der Wal et al., 2021). It's up to us to highlight the importance of this developmental window and mobilize ourselves, our colleagues, and policymakers to support and improve mental health in this critical period of youth.

While some aspects of today's 12-25-year-olds' daily lives are new to us, we must acknowledge that a person's integration into society has long been recognized as a developmental period (Erikson, 1959). Let's face it. The idea that adolescence doesn't neatly end at the age of 18 has been around for decades and is sometimes consistent with our own "lived experience". Yet, despite this, mental health systems, diagnostics and research efforts worldwide have been slow to fully reflect this understanding and consider it consistently in diagnostics, treatment and prevention.

We do not want to waste precious time in the early year of 2025, and we therefore call on you to contribute to the idea of making the most of these windows of opportunity in your daily research and clinical work, which means:

- **Broaden the age range of your research**: Include experts who are skilled in adapting questionnaires, clinical interviews, and interventions for this age group. Even if "youth" ranges from 12 to 25, it's important to recognize that instruments need to be adapted because cognitive and emotional development limit the assessment of youth.
- **Broaden your treatment approach**: When treating children, adolescents and young adults with or at risk for mental health problems, use your expertise in developmental psychopathology to anticipate future needs. Ask yourself: What will this child's mental health look like at the age of 16? Where should we intervene to support their well-being at the age of 22? How might the family system be able to adapt to the child or young person's needs and promote positive change?
- **Educate parents**: Help parents to better understand their evolving role in the modern world and provide parental training to help them model resilience, cope with daily challenges, and promote mental well-being in their children.
- **Advocate for mental health policies**: Advocate for policies that view youth mental health as a critical window of opportunity, which should be recognized and promoted wherever and whenever possible.
- **Include youth in your research**: Youth with lived experience or youth in the same age range as your research samples can provide valuable feedback and help to advance the understanding of mental health problems in youth.

We thank you for your patience if you find our enthusiasm is too strong at the beginning of this year. However, we have long been committed to contributing to the youth mental health paradigm and would like to do so more this year in our new role as editor-in-chief and editors of CPE. We look forward to your contributions and to the exchange with you on this topic.

We thank the publishers and the reviewers for their excellent contribution to open science and look forward to the upcoming CPE volumes with the same enthusiasm. We thank you for your critical commitment and for including a developmental psychology perspective in your research and clinical work.

Funding: The authors have no funding to report.

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: Simone Munsch and Tina In-Albon are Subject Editors and Nadine Messerli-Bürgy is Editor-in-Chief of Clinical Psychology in Europe.

References

Erikson, E. H. (1959). *Identity and the life cycle*. International Universities Press.

Insel, T. R., & Fenton, W. S. (2005). Psychiatric epidemiology: It's not just about counting anymore. *Archives of General Psychiatry, 62*(6), 590–592. https://doi.org/10.1001/archpsyc.62.6.590

Kessler, R. C., Avenevoli, S., & Ries Merikangas, K. (2001). Mood disorders in children and adolescents: An epidemiologic perspective. *Biological Psychiatry, 49*(12), 1002–1014. https://doi.org/10.1016/S0006-3223(01)01129-5

Orben, A., Tomova, L., & Blakemore, S. J. (2020). The effects of social deprivation on adolescent development and mental health. *The Lancet Child & Adolescent Health, 4*(8), 634–640. https://doi.org/10.1016/S2352-4642(20)30186-3

Sawyer, S. M., Azzopardi, P. S., Wickremarathne, D., & Patton, G. C. (2018). The age of adolescence. *The Lancet Child & Adolescent Health, 2*(3), 223–228. https://doi.org/10.1016/S2352-4642(18)30022-1

Solmi, M., Radua, J., Olivola, M., Croce, E., Soardo, L., Salazar de Pablo, G., Il Shin, J., Kirkbride, J. B., Jones, P., Kim, J. H., Kim, J. Y., Carvalho, A. F., Seeman, M. V., Correll, C. U., & Fusar-Poli, P. (2022). Age at onset of mental disorders worldwide: Large-scale meta-analysis of 192 epidemiological studies. *Molecular Psychiatry, 27*(1), 281–295. https://doi.org/10.1038/s41380-021-01161-7

Speyer, L. G., Ushakova, A., Hall, H. A., Luciano, M., Auyeung, B., & Murray, A. L. (2022). Analyzing dynamic change in children's socioemotional development using the Strengths and Difficulties Questionnaire in a large United Kingdom longitudinal study. *Journal of Psychopathology and Clinical Science, 131*(2), 162–171. https://doi.org/10.1037/abn0000714

Thompson, E. J., Richards, M., Ploubidis, G. B., Fonagy, P., & Patalay, P. (2023). Changes in the adult consequences of adolescent mental ill-health: Findings from the 1958 and 1970 British birth cohorts. *Psychological Medicine, 53*(3), 1074–1083. https://doi.org/10.1017/S0033291721002506

Uhlhaas, P. J., Davey, C. G., Mehta, U. M., Shah, J., Torous, J., Allen, N. B., Avenevoli, S., Bella-Awusah, T., Chanen, A., Chen, E. Y. H., Correll, C. U., Do, K. Q., Fisher, H. L., Frangou, S., Hickie, I. B., Keshavan, M. S., Konrad, K., Lee, F. S., Liu, C. H., . . . Wood, S. J. (2023). Towards a youth mental health paradigm: A perspective and roadmap. *Molecular Psychiatry, 28*(8), 3171–3181. https://doi.org/10.1038/s41380-023-02202-z

van der Wal, J. M., van Borkulo, C. D., Deserno, M. K., Breedvelt, J. J. F., Lees, M., Lokman, J. C., Borsboom, D., Denys, D., van Holst, R. J., Smidt, M. P., Stronks, K., Lucassen, P. J., van Weert, J. C. M., Sloot, P. M. A., Bockting, C. L., & Wiers, R. W. (2021). Advancing urban mental health research: From complexity science to actionable targets for intervention. *The Lancet Psychiatry, 8*(11), 991–1000. https://doi.org/10.1016/S2215-0366(21)00047-X

EACLIPT

Clinical Psychology in Europe (CPE) is the official journal of the European Association of Clinical Psychology and Psychological Treatment (EACLIPT).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.

Clinical Psychology in Europe

PsychOpen GOLD

Systematic Reviews and Meta-analyses

Allegiance and Treatment Quality as Moderators of the Comparative Effectiveness of Psychotherapy? A Systematic Review and Meta-Analysis of Studies Comparing Humanistic Psychotherapy to Other Psychotherapy Approaches

Olivia Schünemann[1], Alessa Jansen[2], Ulrike Willutzki[3], Nina Heinrichs[4]

[1] *Institute of Psychology, TU Braunschweig, Braunschweig, Germany.* **[2]** *Bundespsychotherapeutenkammer, Berlin, Germany.* **[3]** *Department for Psychology and Psychotherapy, Witten/Herdecke University, Witten/Herdecke, Germany.* **[4]** *Department of Psychology, Bielefeld University, Bielefeld, Germany.*

Clinical Psychology in Europe, 2025, Vol. 7(1), Article e9709, https://doi.org/10.32872/cpe.9709

Received: 2022-06-20 • **Accepted:** 2024-10-15 • **Published (VoR):** 2025-02-28

Handling Editor: Winfried Rief, Philipps-University of Marburg, Marburg, Germany

Corresponding Author: Olivia Schünemann, Institute of Psychology, TU Braunschweig, Humboldtstr. 33, 38106 Braunschweig, Germany. Phone: +49 531 391 2852. E-mail: o.schuenemann@tu-bs.de

Supplementary Materials: Materials, Preregistration [see Index of Supplementary Materials]

Abstract

Background: Achieving positive outcomes in comparative RCTs examining psychotherapy interventions may be moderated by other factors than treatments alone, namely allegiance and treatment quality (bona fide, adherence). Using the study sample of a recent comprehensive review on humanistic interventions by the German Scientific Board of Psychotherapy, we assumed that higher allegiance towards non-humanistic approaches and lower treatment quality in the humanistic intervention arm would result in worse outcomes for the humanistic groups.

Method: We included studies in which a humanistic psychotherapy (sub-)approach was compared to another type of psychotherapy. Data was extracted independently by the authors. A priori defined meta-regression analyses were performed with allegiance and treatment quality as main moderators and study quality (risk of bias), type of active control, humanistic psychotherapy and target population (children/adolescents; adults) as exploratory.

Results: The majority of studies showed non-allegiance towards humanistic intervention arms; only about half of the humanistic interventions were bona fide treatments demonstrating high percentages of potential biases in these comparative intervention studies. However, allegiance and

bona fide were significant moderators only for two (allegiance) resp. one (bona fide) of five outcome comparison. Type of active control (cognitive behavioural therapy) and disorder group (anxiety disorders) emerged as further moderators.

Conclusion: We found no clear evidence for allegiance or treatment quality impacting upon treatment outcome in this re-examination. Allegiance and treatment quality were not as relevant for outcomes in this meta-analysis of RCTs as expected.

Keywords

RCT, allegiance, bona fide, study quality, adherence, humanistic psychotherapy

Highlights

- Analysed studies displayed substantial heterogeneity concerning allegiance and treatment quality.
- Only 10% of the studies showed allegiance in favour of HPT.
- Allegiance and bona fide were significant moderators in 2 (alleg.)/1 (bona fide) of 5 comparisons.
- Allegiance and treatment quality were not as relevant for outcomes as expected in our examination.

Humanistic psychotherapy (HPT) has been characterized as a psychotherapeutic approach focusing on "the conditions or stances by which people can come to intimately know themselves and, to the extent possible, to fulfill their aspirations" (Schneider & Leitner, 2002, p. 949). International literature lists HPT as one of the major psychotherapy approaches next to cognitive-behavioural, psychodynamic and systemic psychotherapy (e.g., Lambert, 2013). Like other psychotherapy approaches HPT is a broad and diverse psychotherapeutic approach embracing various subapproaches, for example client / person-centered psychotherapy (Rogers, 1961), constructivist psychotherapy (Neimeyer, 1995), emotion-focused psychotherapy (Greenberg et al., 1998), existential meaning making psychotherapy (Schneider & Krug, 2010), focusing-oriented psychotherapy (Gendlin, 1981), gestalt psychotherapy (Perls et al., 1994) and transpersonal psychotherapy (Wilber et al., 1986; selection suggested by Angus et al., 2015). Schneider and Längle (2012, p. 428) summarized the common assumptions of HPT as follows: "Humanism is concerned with such existential themes as meaning, mortality, freedom, limitation, values, creativity, and spirituality as these arise in personal, interpersonal, social, and cultural contexts. In psychotherapy, humanism places special emphasis on the personal, interpersonal, and contextual dimensions of therapy and on clients' reflections on their relationship with self, others, and the larger psychosocial world". Accordingly, all HPT subapproaches share the following principles: Endorsement of the centrality of an empathic and prizing therapeutic relationship and a focus on the promotion of client experiencing in the therapy process (Elliott et al., 2013). Nevertheless, significant differences between these

subapproaches exist, e.g., in their preferred methods for building a therapeutic alliance, treatment indications, case conceptualizations or concepts for individual treatment planning (e.g., Elliott et al., 2021; GSBP, 2018).

A significant attempt to aggregate empirical results in the context of HPT is reflected in the meta-analysis by Elliott et al. (2013) who analysed 191 HPT-studies (through 2008) involving person-centered, supportive or nondirective, and other HPT subapproaches. Regarding pre-post effects of HPT (n = 191 studies with n = 14,235 clients), Elliott et al. (2013, p. 10) reported a weighted pre-post effect size of d = 0.93 with stable pre-follow up effects up to d = 1.11 when pooling effects of all subapproaches of HPT. Their comparison of HPT to other therapy approaches indicated no differences in pre-post results (weighted effect size d = 0.01; 95% CI [-0.05, 0.07]. However, the comparison group "other therapies" in this meta-analysis is not well described. Comparing HPT to cognitive-behavioural therapy (CBT) only, HPT was described as slightly inferior to CBT (d = -0.13, 95% CI [-0.21, -0.06]; Elliott et al., 2013, p. 10).

There are currently a number of psychotherapy approaches available for practitioners to deliver clinical services around Europe. In Germany, however, HPT is not part of the list accredited psychotherapists can choose from and be reimbursed for their services. To become part of this list, psychotherapy approaches have to undergo two independent evaluations in Germany: (1) a scientific evaluation conducted by the German Scientific Board of Psychotherapy (GSBP [Wissenschaftlicher Beirat Psychotherapie]); and – if they are evaluated positively by the GSBP – (2) subsequent assessment concerning cost-utility aspects (conducted by another, independent board). A recent evaluation of HPT by the GSBP on request of the Work Group Humanistic Psychotherapy (Arbeitsgemeinschaft Humanistische Psychotherapie, 2012) resulted in the conclusion that the approach is not evidence-based according to the Board's criteria.

Potential Reasons for These Discrepant Results: Allegiance and Treatment Quality

The discrepancy between the review result of the GSBP and other international reviews on HPT (e.g, by Elliott et al., 2013; Elliott et al., 2021) can be explained through the different methodological approaches, e.g. using meta-analyses and pooling effects for a certain disorder (Elliott et al., 2013; Elliott et al., 2021) vs. defining a required minimum of evidence (at least three RCTs with sufficient internal and external validity) for a certain number of diagnostic groups (currently: 4 out of 12) in the GSBP approach. Moreover, the discrepant definition of what constitutes HPT as defined by the German Work Group Humanistic Psychotherapy from the international definition, e.g. as defined in the reviews by Elliot is likely to be relevant here. The assignment of subapproaches to HPT is different across author groups and also within type of publications. Elliott et al. (2021) use the term HPT to embrace 10 subapproaches (person-centered therapy, emotion-focused therapy, motivational interviewing, gestalt, existential, psychodrama,

focusing-oriented, expressive and body-oriented as well as "supportive / nondirective"), while in Elliott et al. (2013) the respective analyses included 9 subapproaches, omitting motivational interviewing. On the basis of current discussions in psychotherapy research (e.g., Wampold, 2015), we hypothesized in addition that two key reasons may have an impact on the effectiveness on HPT subapproaches (Schünemann et al., 2019):

1) Researchers' allegiance that may be associated with a decrease of the (possible) effects of the non-preferred treatment condition (Munder et al., 2011), in case of the review of the GSBP a decrease of the effects of subapproaches of HPT. Leykin and DeRubeis (2009, p. 55) define "Allegiance, in the context of treatment outcome research, …[as] a belief in the superiority of a treatment. It usually also entails a belief in the superior validity of the theory of change that is associated with the treatment". Often, this belief is associated with therapy outcomes and may reflect a risk of bias. A meta-analysis on the allegiance bias hypothesis demonstrated a moderate association between allegiance and treatment outcome moderated by methodological quality (Munder et al., 2011), indicating that allegiance as well as methodological quality may need to be taken into account for treatment comparisons (Munder et al., 2011). In the context of the HPT evaluation of the GSBP, researchers' allegiance may have had an effect because many studies in the pool used a study design in which a humanistic subapproach was designed as control group in comparison to other psychotherapeutic approaches (e.g., CBT).

2) We further assumed that treatment quality of HPT studies in the GSBP evaluation may have an impact on outcomes. In this context, the two aspects of bona fide psychotherapy and adherence can be differentiated. Whereas bona fide represents *conceptual quality* of a treatment, adherence is rather concerned with *process quality*. For the present study, bona fide psychotherapy will be defined as follows: mentioning or describing an established psychological approach, psychological treatment principles, a treatment manual or active treatment ingredients (Benish et al., 2008) as well as the requirement that the intervention is implemented by a trained therapist (Wampold et al., 1997). Further, we use a definition of adherence in accordance with Munder et al.'s (2013) definition of treatment integrity: "[…] conceptualized broadly including adherence to specific treatment procedures (e.g., the importance of exposure in psychotherapy for post-traumatic stress disorder), common factors (e.g., therapeutic alliance), and therapist effects (i.e., differences in the effects due to individual therapists)" (Munder et al., 2013, p. 8). Concerning treatment quality, Wampold (2015) summarizes that trainers' competence has only a small (d = 0.14) and adherence to protocol an almost negligible effect (d = 0.04) on treatment outcome. The (non-) effect of adherence to protocol may be explained with associated decreased patient-therapist alliance, increased likelihood of resistance to treatment and a lack of flexibility by the therapist. We hypothesized that the GSBP-evaluation included studies implementing HPT subapproaches (often as control group) in potentially fuzzy and poor quality so that significant HPT mechanisms

could not work properly resulting in potentially decreased effects of the respective HPT subapproach.

Objectives

The aim of this meta-analysis is the re-examination of studies included in the GSBP evaluation of HPT (GSBP, 2018) (based on the application of the Work Group Humanistic Psychotherapy and the additional searches conducted by the GSBP) to examine potential moderating effects of allegiance and treatment quality (bona fide psychotherapy and adherence) on the comparative effectiveness of HPT (including both efficacy and effectiveness studies). We examine whether in the HPT subapproaches vs. others psychotherapeutic interventions included in the GSBP evaluation 1) allegiance to a particular psychotherapeutic intervention and 2) treatment quality (bona fide/adherence) significantly moderate effect sizes. We expected higher allegiance towards other psychotherapy approaches and lower treatment quality to be significant predictors of lower effect sizes in the HPT condition.

Method

Design, Study Search and Procedure

This study was registered in the PROSPERO International prospective register of systematic reviews (CRD42019128983). The present full report relates to the second objective in this protocol (comparison of humanistic subapproaches versus other psychotherapeutic interventions).

The study pool which builds the basis of the present analysis is taken over from the GSBP (2018). For their final sample, the GSBP screened abstracts of n = 481 studies from which N = 114 went into the final pool of evaluated studies. The abstract screening procedure was based on 1) a list of studies submitted by the Work Group Humanistic Psychotherapy (n = 313), and 2) results from an independent literature search conducted by an independent institution. According to pre-specified PICOS criteria, all study abstracts were screened; the ones included went into full-text screening review by two independent reviewers (members of the GSBP) to decide about their further inclusion.

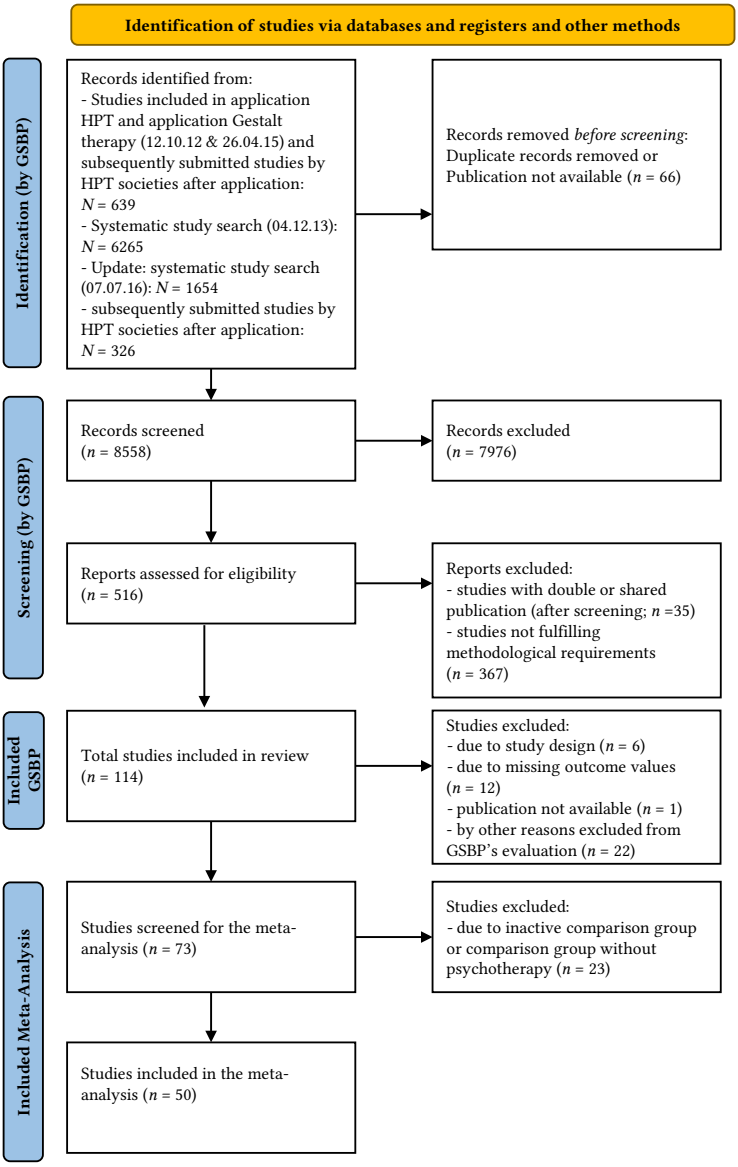
The present systematic review used the identical study pool of the GSBP (N = 114) in order to compare the results of this systematic review directly with the conclusions of the GSBP and by this to be able to provide recommendations for future analyses on psychotherapeutic approaches or methods by the GSBP.

It was necessary to conduct a secondary study screening for eligibility because the present systematic review needs specific study data beyond the data relevant for the purpose of the GSBP. As result, we used N = 50 studies which compared HPT subapproaches

to alternative evidence-based psychotherapeutic interventions (see flowchart Figure 1). The remaining studies were either excluded or relate to inactive control conditions.

Figure 1

PRISMA Flow Diagram Showing the Process for Search and Selection of Studies



Study characteristics for the present meta-analyses were extracted by one author (either OS or AJ). Study results, study quality, allegiance and treatment quality were rated by two of the authors independently (OS & NH or AJ & UW or NH & UW). Inter-rater reliability (Cohen`s Kappa) for adherence (only available for one of the two rater teams: NH, OS) across all intervention groups (IG and CG) was moderate (Altman, 1999): $\kappa = .507$ ($p < .001$); agreement = 76.7%. Inter-rater reliability (intra-class correlation, ICC) for allegiance (MARS index) was .91 (95% CI [.82, .95]; two-way mixed-effects ICC, absolute agreement, average for two coders, same rater team).

Discrepancies were first discussed within dyads and if discrepancies could not be clarified, were discussed among all four authors until a final decision was reached.

Inclusion and Exclusion Criteria

Inclusion criteria were based on those of the GSBP and complemented by the availability of data we needed for the meta-analysis. Inclusion criteria were: 1) RCTs or non-randomized controlled trials with active control group and 2) report of pre- and post-assessments regardless of follow-up assessments. These criteria were complemented with the following exclusion criteria: 1) an active control group was missing (only two or more HPT subapproach-groups with no additional active control group precluding the examination of allegiance); 2) a metric outcome measure or post mean were lacking or 3) an indication of data manipulation could be found. For more details see Schünemann et al. (2019) or Supplementary Materials (*Additional inclusion and exclusion criteria*).

Data Collection

We extracted information on participant characteristics, study characteristics, intervention characteristics, primary and secondary outcomes, risk of bias, allegiance, treatment quality (bona fide and adherence). Guidelines of the Cochrane Collaboration were used to estimate and substitute missing data for outcomes, e.g., calculating standard errors from exactly reported *t*-values. The primary effectiveness outcome was symptom severity at the end of treatment measured on a metric symptom specific scale. Outcomes on self-rating scales (e.g., BDI) were given priority over observer-rated scales (e.g., HDRS). As symptom reduction is not necessarily the primary target of change across different psychotherapy approaches, we extracted data for different outcome domains to analyse moderator effects specific to different outcome domains. Secondary outcomes were interpersonal outcomes (e.g., DAS), general assessment of functioning (e.g., GAF) and quality of life (e.g., WHO QOL). The primary outcome was extracted for short-term (end of intervention) and follow-up if available (6 months after end of intervention or the one closest to 6 months). As primary negative outcome drop-out until end of intervention was extracted.

Assessment of Main and Exploratory Moderators: Study Quality, Allegiance and Treatment Quality

Allegiance was assessed according to the multilevel allegiance rating-scale (MARS) provided by Steinert et al. (2017). This instrument combines information about 1) researchers' allegiance either respective treatment development or contribution to an etiological understanding of the treated disorder; 2) therapists' allegiance; 3) trainers' allegiance and 4) supervisors' allegiance to a total score (0-4; Steinert et al., 2017).

Treatment quality with bona fide psychotherapy and adherence were rated using the definition by Wampold et al. (1997) and Benish et al. (2008) according to the following items: using/citing an established psychotherapy manual; used intervention is based on psychological principles; author mentions HPT (subapproach)/CG on own his/her website; author has other relevant HPT (subapproach)/CG publications; intervention was carried out by trained therapists. To meet the criterion of bona fide two of the items needed to be fulfilled. In the moderator analyses "bona fide" was compared to "non or unclear bona fide". To fulfill the criterion of treatment adherence, treatment conditions and therapeutic procedure needed to be described in detail and adherence must have been proved by external raters. In the moderator analyses "non-adherence" was compared to "adherence". Study quality of the included studies was rated according to the second version of Cochrane's risk of bias tool (RoB 2.0; Higgins et al., 2018) considering the adaptions by Munder and Barth (2018) for its use in psychotherapy outcome research. Thus, the methodological quality was assessed via: 1) bias arising from the randomization process; 2) bias due to missing outcome data; 3) bias in outcome measurement; and 4) bias in selection of the reported result (Cochrane's risk of bias tool) as well as 5) effect of adhering to intervention (see Higgins et al., 2018; Munder & Barth, 2018). For non-randomized controlled trials, the ROBIN-I tool (Sterne, Hernán, et al., 2016) was used to rate risk of bias. For moderator analyses a "low risk of bias" in each category was compared to an "unclear or high risk of bias".

Statistical Analyses

Standardized mean difference for metric measures and odds ratios for rare outcomes between the intervention groups at end of intervention and follow-up were calculated using the intention-to-treat sample, if available. For all analyses, a random effects model with inverse variance weights was applied (DerSimonian & Laird, 2015). Cochran's Q-test was used and quantified using the I^2-statistic (Higgins et al., 2003) to test statistical heterogeneity between study results for significance. Visual examination of funnel plots and Egger's test (Sterne, Egger, & Smith, 2016) were applied to examine possible publication bias.

A priori defined subgroup (in case of categorical predictors) or meta-regression (in case of metric predictors) analyses (univariate) were conducted concerning study

quality, allegiance, treatment quality, type of non-active control (waitlist vs. all others including TAU), type of HPT subapproach (client centered vs. all others) and population (children/adolescents vs. adults). Differences between subgroups were tested formally (Bucher et al., 1997; Deeks et al., 2001; Song et al., 2003). A posteriori (explorative) meta-regression analyses (univariate) were performed in case of considerable heterogeneity between studies for number of sessions, length of intervention, percentage of women, affective disorder, anxiety disorder, post-traumatic stress disorder (PTSD), F 54, study design (no RCT) and comparator (CBT). All analyses were conducted using the metafor package in R (Viechtbauer, 2010).

Results

Descriptive and Additional Results

Because of space limitations additional information and results are documented in the Supplementary Materials. Individual characteristics of all studies are shown in the Supplementary Materials (Supplementary Table 1). All ratings for Cochrane's risk of bias and ROBIN-I tool (study quality) as well as for allegiance and treatment quality for each study can be found in the Supplementary Materials (Supplementary Table 2). Further, the overall main effect sizes are presented in Supplementary Table 3. In order to avoid misunderstandings when interpreting the results of the overall effects a more detailed explanation of the main effects can also be found in Appendix C – Supplementary Information (see Preliminary Results: Overall main effect sizes of HPT). The forest plot for the primary outcome is displayed in Figure 2.

Main Results: Univariate Moderator Effects – The Role of Allegiance and Treatment Quality

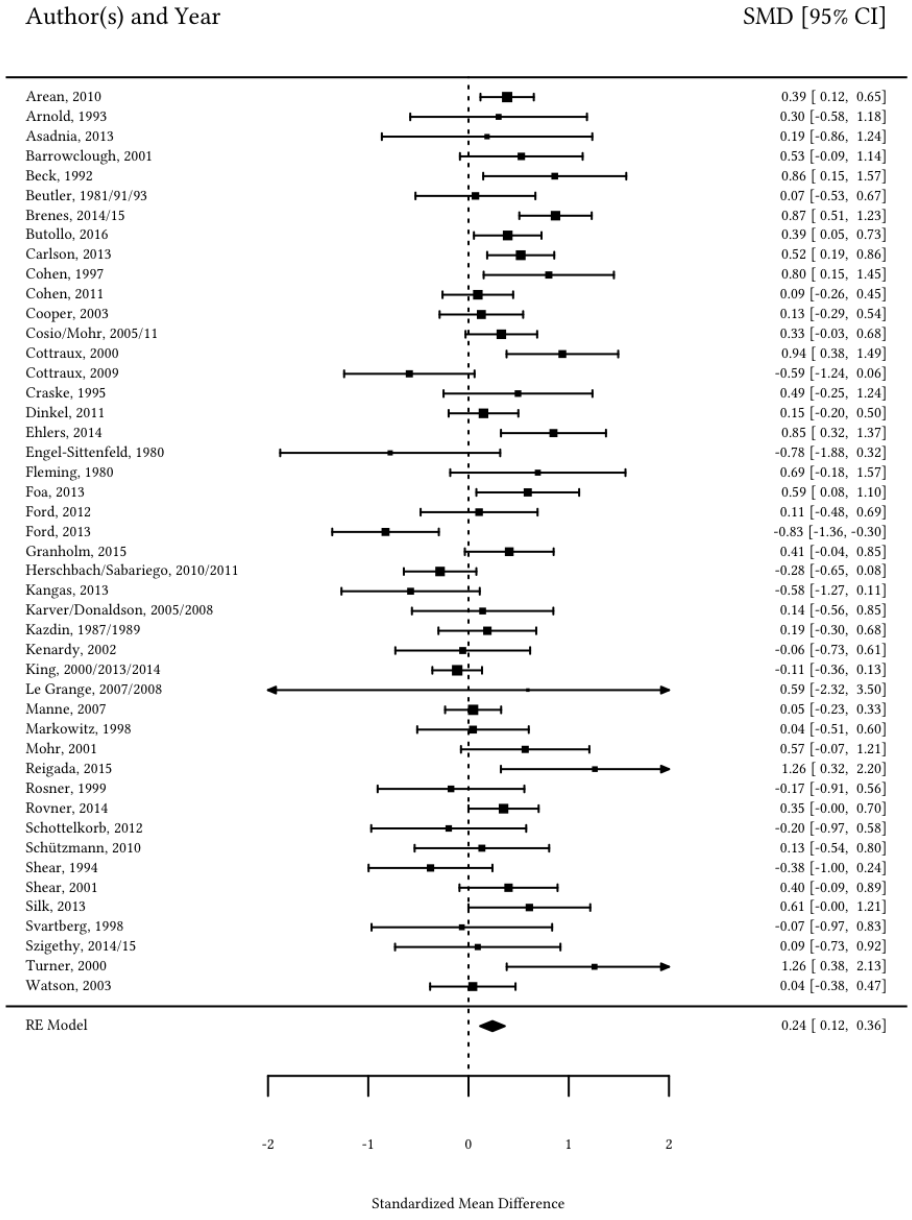
We will first report on moderation effects (pre-specified moderators followed by exploratory moderators) for post-intervention outcomes and then move to reporting the results for the same moderators at follow up.

The results for the moderator analyses for the symptom severity (primary outcome) are presented in Table 1A. No significant moderating effects on the primary comparative effectiveness outcome *symptom severity at end of intervention* could be shown for the pre-specified moderators study quality, allegiance or treatment quality.

Exploratory subgroup analyses indicated that the difference between HPT and other psychotherapies regarding symptom severity increased significantly favouring other psychotherapies when CBT was used as comparator and in studies examining anxiety disorders. Sensitivity analysis for this comparison showed that results were not influenced by the study design.

Figure 2

Exemplary Forest Plot of the Comparison HPT Approaches vs. Other Psychotherapy Only for the Primary Outcome Symptom Severity at End of Intervention



Clinical Psychology in Europe
2025, Vol. 7(1), Article e9709
https://doi.org/10.32872/cpe.9709

Table 1A

Moderators Analysis for Primary Outcome: Symptom Severity at End of Intervention and at Follow-up

	Symptom severity end of intervention (N = 47[b])					Symptom severity at follow-up (N = 30[c])				
Moderators[a]	**Intercept**	**β**	***SE***	***p***	**R^2 in %**	**Intercept**	**β**	***SE***	***p***	**R^2 in %**
no or unclear allegiance against HPT	0.315	-.183	0.125	.144	6.84	**0.365**	**-.208**	**0.103**	**0.044**	**70.62**
non or unclear bona fide HPT	0.185	.116	0.127	.361	< 0.01	0.191	.159	0.118	.178	17.41
non or unclear bona fide CG	0.243	-.042	0.230	.854	< 0.01	0.300	-.241	0.235	.305	0.91
non adherence HPT	0.260	-.054	0.132	.679	< 0.01	0.294	-.018	0.125	.887	< 0.01
non adherence CG	0.226	.070	0.160	.673	< 0.01	0.240	.190	0.140	.175	8.27
low RoB RP	0.175	.111	0.128	.385	< 0.01	0.284	.001	0.126	.996	< 0.01
low RoB AI	0.210	.067	0.129	.602	< 0.01	0.369	-.208	0.124	.094	< 0.01
low RoB MI	0.194	.080	0.129	.537	< 0.01	0.228	.101	0.122	.407	< 0.01
low RoB MO	0.193	.065	0.140	.642	< 0.01	0.324	-.062	0.130	.637	< 0.01
low RoB SR	0.292	-.217	0.148	.142	< 0.01	0.286	-.008	0.165	.960	< 0.01
client centered psychotherapy	0.287	-.060	0.156	.698	< 0.01	0.302	-.022	0.156	.887	< 0.01
population: children	0.213	.128	0.158	.419	< 0.01	0.223	.280	0.143	.050	21.79
age*[d]	0.239	< .001	0.003	.932	< 0.01	0.493	-.006	0.004	.127	19.60
% of women*[d]	0.257	< .001	0.003	.867	< 0.01	0.558	-.004	0.002	.074	44.16
affective disorder*	0.256	-.088	0.160	.584	< 0.01	0.286	-.006	0.150	.969	< 0.01
anxiety disorder*	**0.161**	**.400**	**0.147**	**.006**	**25.84**	0.283	.006	0.198	.975	< 0.01
PTSD*	0.240	-.005	0.163	.977	< 0.01	0.238	.273	0.159	.087	16.12
F 54*	0.283	-.181	0.142	.202	0.40	0.346	-.197	0.124	.114	16.34
design: no RCT*	0.249	-1.030	0.642	.108	2.80	0.285	-.128	0.572	.823	< 0.01
length of intervention HPT*[d]	0.291	< .001	0.006	.966	< 0.01	0.261	.002	0.007	.808	< 0.01
number of sessions HPT*[d]	0.197	.003	0.005	.604	< 0.01	0.300	-.001	0.005	.854	< 0.01
Comparator: CBT*	**-0.357**	**.638**	**0.236**	**.007**	**21.03**	-0.029	.321	0.370	.386	< 0.01
number of sessions CG*[d]	0.154	.006	0.006	.348	< 0.01	0.201	.006	0.007	.436	< 0.01

Note. Grey shadows for main (i.e. pre-specified) moderators. Significant results of moderator analyses (p < .05) are shown in bold. RoB = risk of bias; CG = control group; CBT = cognitive behavioural therapy; NA = moderator analysis could not be conducted due to insufficient variance in moderator. Example for interpretation of a categorical predictor: In studies with no or unclear allegiance against HPT the difference between HPT approaches and other forms of psychotherapy in levels of symptom severity at follow-up is significantly reduced (β = -0.208, p = .044). Example for interpretation of a metric predictor: In studies with one percentage more women the difference between HPT approaches and other forms of psychotherapy in levels of symptom severity at end of intervention is not significantly increased (β = < .001, p = .867). [a]N varies for individual moderator analysis as studies with missing data in moderators were removed from metaregression analyses. [b]One study (Dietz et al., 2015) was identified as an outlier and removed from all further meta-analysis. [c]One study (Herschbach et al., 2010) was identified as an outlier and removed from all further meta-analysis. [d]metric predictor (all other predictors are dichotomous). *A posteriori analyses.

For the outcome *symptom severity at follow-up* subgroup analyses indicated that the difference between HPT in comparison to other psychotherapies was reduced significantly in studies with no or unclear allegiance against HPT (β = -0.208, p = .044; for examples for interpretation, please see table note).

Results for secondary outcome domains are shown in Table 1B. For the outcome *interpersonal problems* subgroup analyses indicated that the difference between HPT in comparison to other psychotherapies was reduced significantly in studies with no or unclear allegiance against HPT. Also, the difference between HPT in comparison to other psychotherapies was reduced when bona fide HPT was used, in studies with a low risk of bias due to missing data and in studies not examining anxiety disorders. Yet, somewhat contradictory, the difference increased in favour of other psychotherapies when the control group was not adherent to intervention. No significant moderator effects were detected for the outcomes *general functioning* and *quality of life*.

The results for some of the outcome measures should be interpreted with caution because of the small number of studies (n = 9 for interpersonal problems, n = 12 for general functioning, n = 8 for quality of life) in comparison to the large number of moderator analyses.

To test whether primary moderators were sufficiently independent, we examined their associations post-hoc via Fisher's Exact Test (FET). Results showed a significant association between allegiance and bona fide HPT (p = .01, FET). Moreover, bona fide HPT was significantly associated with adherence HPT (p = .007, FET) as well as with adherence to the control condition (p = .02, FET). There were no other significant associations between primary moderators.

Discussion

The purpose of this study was to investigate the effects of allegiance and treatment quality (bona fide psychotherapy, adherence) for assessing the comparative effectiveness of psychotherapy. We used data comparing the effects of HPT subapproaches (from a study pool used in a recent evaluation of the GSBP) to other evidence-based treatments.

Clinical Psychology in Europe
2025, Vol. 7(1), Article e9709
https://doi.org/10.32872/cpe.9709

Table 1B

Moderators Analysis for Secondary Outcome Domains

	Interpersonal problems (N = 9[b])					General functioning (N = 12)					Quality of life (N = 8)[c]				
Moderators[a]	**Intercept**	**β**	***SE***	***p***	**R^2 in %**	**Intercept**	**β**	***SE***	***p***	**R^2 in %**	**Intercept**	**β**	***SE***	***p***	**R^2 in %**
No or unclear allegiance against HPT	**0.502**	**-.519**	**0.169**	**.002**	**> 0.99**	-0.275	.331	0.224	.140	7.22	-0.378	.280	0.171	.102	87.49
Non or unclear bona fide HPT	**0.027**	**.568**	**0.197**	**.004**	**> 0.99**	0.084	-.335	0.235	.154	2.40	-0.098	-.280	0.171	.102	87.49
Non or unclear bona fide CG	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Non adherence HPT	0.013	.353	0.231	.127	18.77	-0.165	.069	0.239	.772	< 0.01	-0.099	-.234	0.163	.151	> 0.99
Non adherence CG	**0.071**	**.830**	**0.290**	**.004**	**> 0.99**	-0.100	-.170	0.320	.596	< 0.01	-0.090	-.290	0.170	.079	> 0.99
Low RoB RP	**0.028**	**.362**	**0.169**	**.032**	**> 0.99**	0.036	-.275	0.227	.226	8.16	-0.137	-.104	0.221	.639	< 0.01
Low RoB AI	0.075	.288	0.270	.285	< 0.01	-0.026	-.214	0.231	.353	< 0.01	-0.363	.213	0.204	.296	22.38
Low RoB MI	**0.404**	**-.408**	**0.163**	**.012**	**> 0.99**	-0.037	-.166	0.240	.489	< 0.01	-0.214	-.003	0.211	.987	< 0.01
Low RoB MO	0.254	-.120	0.272	.659	< 0.01	-0.047	-.140	0.240	.559	< 0.01	-0.280	.081	0.256	.752	< 0.01
Low RoB SR	0.157	.063	0.278	.819	< 0.01	-0.105	-.086	0.273	.753	< 0.01	-0.249	.108	0.205	.597	< 0.01
Client centered psychotherapy	-0.102	.391	0.229	.088	41.74	0.016	-.176	0.314	.576	< 0.01	-0.165	-.084	0.211	.691	< 0.01
Population: children	0.143	.196	0.333	.555	< 0.01	-0.120	-.048	0.322	.882	< 0.01	NA	NA	NA	NA	NA
Age*[d]	0.298	-.004	0.012	.748	< 0.01	-0.102	-.001	0.012	.947	< 0.01	-0.384	.004	0.011	.729	< 0.01
% of women*[d]	0.721	-.008	0.007	.238	14.74	0.003	-.002	0.005	.646	< 0.01	-0.558	.005	0.005	.321	< 0.01
Affective disorder*	0.273	-.290	0.263	.270	< 0.01	-0.185	.291	0.293	.322	< 0.01	NA	NA	NA	NA	NA
Anxiety disorder*	**0.071**	**.829**	**0.291**	**.004**	**> 0.99**	-0.081	-.317	0.318	.318	0.05	-0.180	-.290	0.308	.346	8.89
PTSD*	0.229	-.191	0.292	.514	< 0.01	-0.020	-.395	0.233	.091	24.76	-0.215	-.005	0.222	.981	< 0.01
F 54*	NA	NA	NA	NA	NA	-0.133	.063	0.446	.888	< 0.01	-0.264	.093	0.202	.645	< 0.01
Design: no RCT*	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Length of intervention HPT*[d]	0.237	.003	0.016	.855	< 0.01	-0.063	-.002	0.010	.801	< 0.01	-0.434	.005	0.008	.552	< 0.01
Number of sessions HPT*[d]	0.341	-.008	0.007	.274	< 0.01	-0.417	.013	0.007	.057	26.08	-0.278	.005	0.011	.665	< 0.01
Comparator: CBT*	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Number of sessions CG*[d]	0.243	-.001	0.016	.938	< 0.01	-0.335	.008	0.013	.536	< 0.01	-0.261	.004	0.009	.691	< 0.01

Note. Grey shadows for main (i.e. pre-specified) moderators. Significant results of moderator analyses (p < .05) are shown in bold. RoB = risk of bias; CG = control group; CBT = cognitive behavioural therapy; NA = moderator analysis could not be conducted due to insufficient variance in moderator. Example for interpretation: In studies with no or unclear allegiance against HPT the difference between HPT approaches and other forms of psychotherapy in levels of symptom severity at follow-up is significantly reduced (β = -0.208, p = .044). [a]N varies for individual moderator analysis as studies with missing data in moderators were removed from metaregression analyses. [b]One study (Dietz et al., 2015) was identified as an outlier and removed from all further meta-analysis. [c]Quality of Life: higher score for better QoL. [d]metric predictor (all other predictors are dichotomous). *A posteriori analyses.

Moderator analyses did not indicate consistent effects of allegiance, treatment quality or study quality on the comparison between HPT and other forms of psychotherapy across all outcomes and assessment points. At the end of intervention, we found a moderating effect in one *(interpersonal problems*) out of four outcomes (*symptom severity, general level of functioning, quality of life*) for allegiance (no or unclear against HPT subapproach) and treatment quality (non bona fide HPT subapproach, non-adherence to control condition) and – to a somewhat lesser extent – also for study quality. At follow-up (with only one outcome domain still available: symptom severity), allegiance significantly moderated comparative treatment effects: in studies with no allegiance against HPT the difference between HPT and other forms of psychotherapy was significantly reduced for *symptom severity at follow-up*. The beta coefficient for symptom severity at end of intervention is pointing in the same direction (ß = -0.183) as at follow-up (ß = -0.208) but is not as pronounced in size.

Exploratory subgroup analyses showed that the difference between subapproaches of HPT in comparison to other psychotherapies was reduced in studies *not* examining anxiety disorders and in studies *not* using CBT as a comparator.

Allegiance

The high proportion of *allegiance* against HPT as well as the high number of non-bona fide treatments in the HPT-treatment arm in the present meta-analysis has also been reported in other studies (Cuijpers et al., 2012; Elliott et al., 2013; Elliott et al., 2021). Cuijpers et al. (2012) concluded that non-directive supportive therapy (as a subapproach of HPT) for depression is equally effective as other psychological treatments after researcher allegiance was controlled. Similarly, Elliott et al. (2013) report a drop in (weighted) effects when comparing supportive therapies to other psychotherapy approaches. In addition, researcher allegiance has been demonstrated to have an impact on outcomes in psychotherapy studies not focusing on HPT alone (Munder et al., 2013).

When focusing on the impact of allegiance, it is important to keep in mind that it is not useful to interpret it as a purposeful attempt to skew results but rather take it into account as a human tendency to believe in one's own ideas and practices in a way that objectivity is compromised (Lomangino, 2016; Yoder et al., 2019). This perspective is supported by research that shows that awareness and acceptance of its potential impact may reduce its effects (Munder et al., 2013). Another option to avoid a bias through allegiance are allegiance-controlled trials where the interventions in the different conditions are planned and supervised by proponents of the respective approaches (Barkham et al., 2021; Leichsenring et al., 2009). Similarly, more recently, bona fide is also explicitly considered in trial designs: In their recent RCT comparing person-centered experiential therapy (PCET as a HPT subapproach) and CBT, Barkham et al. (2021) carefully implemented bona fide treatment arms with a similar level of professional training with the help of trainers qualified in the respective interventions.

Adherence

We did not find moderation effects of adherence on comparative effectiveness. A recent systematic review and meta-analysis of the role of adherence for outcome in child and adolescent psychotherapy found a small (statistically significant) (interventional, non-comparative) effect size for adherence of $d = 0.096$ (95% CI [.058, .124]. Considering the very small effect size, the authors conclude - despite the statistical significance - that other factors than adherence are much more relevant to consider for outcome. They included primary studies if "external or independent observer rated measures of adherence or competence, measures across multiple time points (…)., and interrater agreement established in the study or use of coders trained to this level (…) (ICC) > -0,60 or Kappa > -0,61 or percent agreement > - 90% or the score is based on agreement by multiple rates" (Collyer et al., 2020, p. 419). Our criteria to evaluate adherence were less demanding than in Collyer et al. (2020). When adherence as a construct is significantly linked to outcome, quite likely its actual impact upon explaining differences in outcomes remains very small with the presented mean effect size estimation of $r < 0.10$ (i.e. accounting for less than 1% of the variance in therapy outcome).

Study Quality

With respect to *study quality* (risk of bias, assessed via RoB), we found no clear evidence for moderation effects. These results are in line with the recent studies by Cuijpers, Quero, et al. (2021) as well as Hoppen and Morina (2020). In this context it has to be kept in mind that study quality is quite likely not fully independent from allegiance: Munder et al. (2011) found that allegiance is more strongly related to outcome in studies with lower methodological quality. The inclusion criteria of the GSBP define relatively high methodological standards. These standards may have prevented that allegiance effects unfold as strongly in our study as maybe in other meta-analyses where inclusion criteria are less methodologically strict.

Strengths and Limitations

The present study compasses a considerable number of trials, examined important predefined moderators and demonstrates the challenges in evaluating the significance of allegiance and treatment quality (bona fide and adherence). However, in conducting this meta-analysis, we also were faced with a number of problems:

Limitations: Allegiance

Allegiance was defined via an aggregate score taking into account researchers' allegiance, therapists' allegiance, trainers' allegiance and supervisors' allegiance was used (Steinert et al., 2017). This aggregation forecloses to identify which type of allegiance may be more (or less) relevant for treatment outcome.

Limitations: Treatment Quality

Following Benish et al. (2008) and Wampold et al. (1997) we used a sophisticated bona fide rating scheme with five criteria: use of a manual; application of an intervention based on psychological principles; author(s) indicate(s) affiliation with the examined intervention (sub)approach on own website; author has other publications concerning the examined intervention (sub)approach; therapists received a training (Benish et al., 2008; Wampold et al., 1997). We decided to use a liberal categorization with only two of the above aspects necessary in order to categorize the intervention as an overall bona fide treatment. Despite these seemingly clear criteria, the assessment of bona fide turned out to be challenging because of often insufficient information.

Further difficulties arose concerning the manual criterion. Many studies reported the use of a manual, also in the HPT condition. However, referenced manuals differed considerably in terms of accessibility, comprehensiveness as well as in terms of content (*common* manuals vs. review or overview articles). This constitutes a general problem in psychotherapy research and practice, affecting not only HPT (Willutzki & Andermann, 2019). Therefore, extensive discussions about the manual criterion were necessary, and decisions on the categorization are sometimes not clear cut. For rating adherence, all following criteria had to be fulfilled: treatment conditions needed to be described precisely; the exact (sub)approach must be presented in detail; adherence had to be rated by external observers. Related to the latter, studies sometimes indicated that adherence was rated by external observers but the actual result was missing. In these cases, we decided to rate the criterion of adherence liberally, even if authors did not report their adherence results.

Further Limitations

The most common diagnostic groups of the systematic review were psychological and behavioural factors associated with disorders or diseases classified elsewhere (F54). The GSBP used this category for a wide scope of applications. Thus, there were also studies included examining subjects with (*only*) a diagnosis of a somatic illness (e.g., patients with age-related macular degeneration, Rovner et al., 2014).

While we based our analysis on the 114 studies taken into account by the GSBP it is important to note that our study pool for the moderator analysis is not identical to the study pool underlying the differentiated GSBP-evaluation of HPT (GSBP, 2018). Discrepancies are due to differences between the method paper (GSBP, 2010) and our study protocol. As an example, the GSBP excluded some of the 114 studies from further analysis due to insufficient methodological quality. However, studies with low methodological quality were included in the present meta-analysis as study quality was one of the moderators by using Cochrane's risk of bias or ROBIN-I tool. On the other hand, we had to exclude 19 studies for which either the study design was not appropriate for the present research question (e.g., HPT subapproach vs. other HPT subapproach or HPT

subapproach mixed with other psychotherapy approach), or no mean values were given needed to conduct statistical analyses going beyond the evaluation of the GSBP.

As in many meta-analyses trying to explore heterogeneity among studies via subgroup analyses we face power problems, particularly because studies are not evenly distributed between subgroups. It is still useful and necessary to run the respective subgroup analyses. On the other hand, one could also argue that experiment wise error rates may assist in arguing that the significant results may have occurred by chance alone in the absence of any true effects. Further, it has to be kept in mind that no evidence of the moderators' impact does not mean that this is evidence of no difference (Cuijpers, Griffin, & Furukawa, 2021). Finally, most of the Humanistic studies were studies based on Rogerian psychotherapy and thus results are particularly relevant for this kind of interventions.

Conclusion

The present results could not demonstrate effects of the examined moderators on treatment outcome. However, we mainly examined studies with high study quality. This is crucial to consider for generalizing our results and it is necessary to prove whether results can be replicated for studies with low study quality.

Funding: There are no funders to report for this submission.

Acknowledgments: The authors would like to thank all members of the German Scientific Board of Psychotherapy in the period from 2014 to 2018 for their openness to have UW and NH follow up on these research objectives which is based on the joint work done in the GSBP.

Competing Interests: AJ is an employee of the Federal Chamber of Psychotherapists. NH and UW are members of the German Scientific Board of Psychotherapy. All authors are trained in cognitive behavioral therapy; UW is also trained in systemic therapy.

Reporting Guidelines: The study followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA; Page et al., 2021).

Data Availability: The set of extracted data is available upon request.

Supplementary Materials

The Supplementary Materials contain the following items:

- **Study protocol:** Schünemann et al. (2019)
- **Online appendices:** Schünemann et al. (2025S):
  - ***Appendix A – Supplementary Tables***
    - *Supplementary Table 1.* Study characteristics ($N = 50$)
    - *Supplementary Table 2.* Study quality (Cochrane's risk of bias tool for RCT or ROBIN-I for non-RCT); Allegiance, Bona fide and Adherence rating ($N = 50$)
    - *Supplementary Table 3.* Different outcome domains for the comparison humanistic psychotherapy approaches vs. other psychotherapies
  - **Appendix B – Supplementary Figure**
    - *Supplementary Figure 1.* Exemplary funnel plot of the comparison HPT approaches vs. other psychotherapy only for the primary outcome symptom severity at end of intervention
  - **Appendix C – Supplementary Information**
    - *The evaluation procedure of the GSBP*
    - *Additional inclusion and exclusion criteria*
    - *Preliminary Results: Overall main effect sizes of HPT*

Index of Supplementary Materials

Schünemann, O., Jansen, A., Willutzki, U., & Heinrichs, N. (2025S). *Supplementary materials to "Allegiance and treatment quality as moderators of the comparative effectiveness of psychotherapy? A systematic review and meta-analysis of studies comparing humanistic psychotherapy to other psychotherapy approaches"* [Online appendices]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.15954

References

Altman, D. G. (1999). *Practical statistics for medical research.* Chapman & Hall/CRC. https://doi.org/10.1201/9780429258589

Angus, L., Watson, J. C., Elliott, R., Schneider, K. J., & Timulak, L. (2015). Humanistic psychotherapy research 1990-2015: From methodological innovation to evidence-supported treatment outcomes and beyond. *Psychotherapy Research, 25*(3), 330–347. https://doi.org/10.1080/10503307.2014.989290

Arbeitsgemeinschaft Humanistische Psychotherapie [Work Group Humanistic Psychotherapy]. (2012). *Antrag an den Wissenschaftlichen Beirat Psychotherapie nach § 11 PsychThG auf „Beurteilung der wissenschaftlichen Anerkennung des Verfahrens Humanistische Psychotherapie"* [Humanistic Psychotherapy Application (HPT-A)]. http://www.aghpt.de/texte/aghpt-antrag an den wbp_(1)_antragstext_2012-10-12.pdf

Barkham, M., Saxon, D., Hardy, G. E., Bradburn, M., Galloway, D., Wickramasekera, N., Keetharuth, A. D., Bower, P., King, M., Elliott, R., Gabriel, L., Kellett, S., Shaw, S., Wilkinson, T., Connell, J.,

Harrison, P., Ardern, K., Bishop-Edwards, L., Ashley, K., . . . Brazier, J. E. (2021). Person-centred experiential therapy versus cognitive behavioural therapy delivered in the English Improving Access to Psychological Therapies service for the treatment of moderate or severe depression (PRaCTICED): A pragmatic, randomised, non-inferiority trial. *The Lancet Psychiatry, 8*(6), 487–499. https://doi.org/10.1016/S2215-0366(21)00083-3

Benish, S. G., Imel, Z. E., & Wampold, B. E. (2008). The relative efficacy of bona fide psychotherapies for treating post-traumatic stress disorder: A meta-analysis of direct comparisons. *Clinical Psychology Review, 28*(5), 746–758. https://doi.org/10.1016/j.cpr.2007.10.005

Bucher, H. C., Guyatt, G. H., Griffith, L. E., & Walter, S. D. (1997). The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. *Journal of Clinical Epidemiology, 50*(6), 683–691. https://doi.org/10.1016/S0895-4356(97)00049-8

Collyer, H., Eisler, I., & Woolgar, M. (2020). Systematic literature review and meta-analysis of the relationship between adherence, competence and outcome in psychotherapy for children and adolescents. *European Child & Adolescent Psychiatry, 29*(4), 417–431. https://doi.org/10.1007/s00787-018-1265-2

Cuijpers, P., Driessen, E., Hollon, S. D., van Oppen, P., Barth, J., & Andersson, G. (2012). The efficacy of non-directive supportive therapy for adult depression: A meta-analysis. *Clinical Psychology Review, 32*(4), 280–291. https://doi.org/10.1016/j.cpr.2012.01.003

Cuijpers, P., Griffin, J. W., & Furukawa, T. A. (2021). The lack of statistical power of subgroup analyses in meta-analyses: A cautionary note. *Epidemiology and Psychiatric Sciences, 30*, Article e78. https://doi.org/10.1017/S2045796021000664

Cuijpers, P., Quero, S., Noma, H., Ciharova, M., Miguel, C., Karyotaki, E., Cipriani, A., Cristea, I.-A., & Furukawa, T. A. (2021). Psychotherapies for depression: A network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. *World Psychiatry, 20*(2), 283–293. https://doi.org/10.1002/wps.20860

Deeks, J. J., Altman, D. G., & Bradburn, M. J. (2001). Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. In M. Egger, G. D. Smith, & D. G. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (pp. 285–312). BMJ Publishing Group. https://doi.org/10.1002/9780470693926.ch15

DerSimonian, R., & Laird, N. (2015). Meta-analysis in clinical trials revisited. *Contemporary Clinical Trials, 45*(Pt A), 139–145. https://doi.org/10.1016/j.cct.2015.09.002

Dietz, L. J., Weinberg, R. J., Brent, D. A., & Mufson, L. (2015). Family-based interpersonal psychotherapy for depressed preadolescents: Examining efficacy and potential treatment mechanisms. *Journal of the American Academy of Child and Adolescent Psychiatry, 54*(3), 191–199. https://doi.org/10.1016/j.jaac.2014.12.011

Elliott, R., Watson, J., Greenberg, L. S., Timulak, L., & Freire, E. (2013). Research on humanistic-experiential psychotherapies. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 495–538). Wiley.

Elliott, R., Watson, J., Timulak, L., & Sharbanee, J. (2021). Research on humanistic-experiential psychotherapies: Updated review. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (pp. 421–468). John Wiley & Sons.

Gendlin, E. T. (1981). *Focusing*. Bantam Books.

Greenberg, L. S., Rice, L. N., & Elliott, R. (1998). *Facilitating emotional change: The moment-by-moment process*. Guilford.

Herschbach, P., Book, K., Dinkel, A., Berg, P., Waadt, S., Duran, G., Engst-Hastreiter, U., & Henrich, G. (2010). Evaluation of two group therapies to reduce fear of progression in cancer patients. *Supportive Care in Cancer, 18*(4), 471–479. https://doi.org/10.1007/s00520-009-0696-1

Higgins, J. P. T., Savović, J., Page, M. J., & Sterne, J. A. C. (2018). *Revised Cochrane risk-of-bias tool for randomized trials (RoB 2)*. https://sites.google.com/site/riskofbiastool/welcome/rob-2-0-tool/current-version-of-rob-2

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ (Clinical Research Ed.), 327*(7414), 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hoppen, T. H., & Morina, N. (2020). Is high-quality of trials associated with lower treatment efficacy? A meta-analysis on the association between study quality and effect sizes of psychological interventions for pediatric PTSD. *Clinical Psychology Review, 78*, Article 101855. https://doi.org/10.1016/j.cpr.2020.101855

Lambert, M. J. (2013). Introduction and historical overview. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 3–20). Wiley.

Leichsenring, F., Hoyer, J., Beutel, M., Herpertz, S., Hiller, W., Irle, E., Joraschky, P., Konig, H. H., de Liz, T., Nolting, B., Pohlmann, K., Salzer, S., Schauenburg, H., Stangier, U., Strauss, B., Subic-Wrana, C., Vormfelde, S., Weniger, G., Willutzki, U., . . . Leibing, E. (2009). The social phobia psychotherapy research network: The first multicenter randomized controlled trial of psychotherapy for social phobia: Rationale, methods and patient characteristics. *Psychotherapy and Psychosomatics, 78*(1), 35–41. https://doi.org/10.1159/000162299

Leykin, Y., & DeRubeis, R. J. (2009). Allegiance in psychotherapy outcome research: Separating association from bias. *Clinical Psychology: Science and Practice, 16*(1), 54–65. https://doi.org/10.1111/j.1468-2850.2009.01143.x

Lomangino, K. M. (2016). Countering cognitive bias: Tips for recognizing the impact of potential bias on research. *Journal of the Academy of Nutrition and Dietetics, 116*(2), 204–207. https://doi.org/10.1016/j.jand.2015.07.014

Munder, T., & Barth, J. (2018). Cochrane's risk of bias tool in the context of psychotherapy outcome research. *Psychotherapy Research, 28*(3), 347–355. https://doi.org/10.1080/10503307.2017.1411628

Munder, T., Brütsch, O., Leonhart, R., Gerger, H., & Barth, J. (2013). Researcher allegiance in psychotherapy outcome research: An overview of reviews. *Clinical Psychology Review, 33*(4), 501–511. https://doi.org/10.1016/j.cpr.2013.02.002

Munder, T., Gerger, H., Trelle, S., & Barth, J. (2011). Testing the allegiance bias hypothesis: A meta-analysis. *Psychotherapy Research, 21*(6), 670–684. https://doi.org/10.1080/10503307.2011.602752

Neimeyer, R. A. (1995). Constructivist psychotherapies: Features, foundations, and future directions. In R. A. Neimeyer & M. J. Mahoney (Eds.), *Constructivism in psychotherapy* (pp. 11–38). American Psychological Association. https://doi.org/10.1037/10170-001

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.), 372*, Article 71. https://doi.org/10.1136/bmj.n71

Perls, F. S., Hefferline, R., & Goodman, P. (1994). *Gestalt therapy: Excitement and growth in the human personality*. Gestalt Journal Press.

Rogers, C. R. (1961). *On becoming a person: A therapist's view of psychotherapy*. Constable.

Rovner, B. W., Casten, R. J., Hegel, M. T., Massof, R. W., Leiby, B. E., Ho, A. C., & Tasman, W. S. (2014). Low vision depression prevention trial in age-related macular degeneration: A randomized clinical trial. *Ophthalmology, 121*(11), 2204–2211. https://doi.org/10.1016/j.ophtha.2014.05.002

Schneider, K. J., & Krug, O. T. (2010). *Existential-humanistic therapy* [Theories of Psychotherapy Series]. https://doi.org/10.1037/12050-000

Schneider, K. J., & Längle, A. (2012). The renewal of humanism in psychotherapy: A roundtable discussion. *Psychotherapy, 49*(4), 427–429. https://doi.org/10.1037/a0027111

Schneider, K. J., & Leitner, L. M. (2002). Humanistic psychotherapy. In M. Hersen & W. Sledge (Eds.), *Encyclopedia of psychotherapy* (pp. 949–957). Elsevier. https://doi.org/10.1016/B0-12-343010-0/00116-1

Schünemann, O., Jansen, A., Willutzki, U., & Heinrichs, N. (2019). Allegiance bias and treatment quality as moderators of the effectiveness of humanistic psychotherapy: Protocol for a systematic review and meta-analysis. *JMIR Research Protocols, 8*(11), Article e15140. https://doi.org/10.2196/15140

Song, F., Altman, D. G., Glenny, A.-M., & Deeks, J. J. (2003). Validity of indirect comparison for estimating efficacy of competing interventions: Empirical evidence from published meta-analyses. *BMJ (Clinical Research Ed.), 326*(7387), Article 472. https://doi.org/10.1136/bmj.326.7387.472

Steinert, C., Munder, T., Rabung, S., Hoyer, J., & Leichsenring, F. (2017). Psychodynamic therapy: As efficacious as other empirically supported treatments? A meta-analysis testing equivalence of outcomes. *The American Journal of Psychiatry, 174*(10), 943–953. https://doi.org/10.1176/appi.ajp.2017.17010057

Sterne, J. A. C., Egger, M., & Smith, G. D. (2016). Investigating and dealing with publication and other biases. In M. Egger, G. D. Smith, & D. G. Altman (Eds.), *Systematic reviews in health care: Meta-analysis in context* (pp. 189–208). BMJ Publishing Group.

Sterne, J. A. C., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A.-W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., . . . Higgins, J. P. T. (2016). Robins-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ (Clinical Research Ed.), 355*, Article i4919. https://doi.org/10.1136/bmj.i4919

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor Package. *Journal of Statistical Software, 36*(3). https://doi.org/10.18637/jss.v036.i03

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry, 14*(3), 270–277. https://doi.org/10.1002/wps.20238

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "all must have prizes.". *Psychological Bulletin, 122*(3), 203–215. https://doi.org/10.1037/0033-2909.122.3.203

Wilber, K., Engler, J., & Brown, D. P. (1986). *Transformations of consciousness: Conventional and contemplative perspectives on development* [New science library]. Shambhala.

Willutzki, U., & Andermann, L. (2019). Manuale in der Psychotherapie – ein schillerndes Thema [Manuals in psychotherapy]. *PiD – Psychotherapie im Dialog, 20*(4), 74–77. https://doi.org/10.1055/a-0771-7771

Wissenschaftlicher Beirat Psychotherapie [German Scientific Board of Psychotherapy]. (2010). *Verfahrensregeln zur Beurteilung der wissenschaftlichen Anerkennung von Methoden und Verfahren der Psychotherapie. Version 2.8* [Procedural rules for assessing the scientific recognition of methods and procedures of psychotherapy].

Wissenschaftlicher Beirat Psychotherapie [German Scientific Board of Psychotherapy]. (2018). Gutachten zur wissenschaftlichen Anerkennung der Humanistischen Psychotherapie [Report on the scientific recognition of humanistic psychotherapy]. *Deutsches Arzteblatt, 3*, 1–14. https://doi.org/10.3238/arztebl.2018.gut_hpt01

Yoder, W. R., Karyotaki, E., Cristea, I.-A., van Duin, D., & Cuijpers, P. (2019). Researcher allegiance in research on psychosocial interventions: Meta-research study protocol and pilot study. *BMJ Open, 9*(2), Article e024622. https://doi.org/10.1136/bmjopen-2018-024622

Clinical Psychology in Europe

PsychOpen GOLD

Systematic Reviews and Meta-analyses

ICD-11 Prolonged Grief Disorder, Physical Health, and Somatic Problems: A Systematic Review

James Cunningham[1], Mark Shevlin[1], Catalina Cerda[1], Eoin McElroy[1]

[1] *School of Psychology, Ulster University, Coleraine, United Kingdom.*

Clinical Psychology in Europe, 2025, Vol. 7(1), Article e14351, https://doi.org/10.32872/cpe.14351

Received: 2024-04-09 • **Accepted:** 2024-08-26 • **Published (VoR):** 2025-02-28

Handling Editor: Winfried Rief, Philipps-University of Marburg, Marburg, Germany

Corresponding Author: Mark Shevlin, Ulster University (Psychology), Room H256, Cromore Road, Coleraine, Northern Ireland, BT52 1SA. Phone: +442870123141. E-mail: m.shevlin@ulster.ac.uk

Supplementary Materials: Materials, Preregistration [see Index of Supplementary Materials]

Abstract

Background: Since Prolonged Grief Disorder's (PGD) inclusion as a mental health disorder in the ICD-11 in 2018, much of the peer-reviewed research has focused on its prevalence, assessment, and co-occurrence with other mental health disorders. There is also emerging research literature on the association between PGD and physical and somatic health outcomes. In light of this, the objective of this review was to identify and summarise the extant research on the association between PGD, and outcomes related to physical health and somatic complaints among bereaved individuals.

Method: A systematic review utilized electronic databases (Web of Science, MEDLINE, Cochrane Library, PsycINFO) up to October 10, 2023. Included were cohort and cross-sectional studies since 2018 exploring links between ICD-11 PGD and physical/somatic health outcomes. Two researchers independently identified eligible studies meeting inclusion/exclusion criteria, employing quality assessment instruments to evaluate methodological rigor.

Results: From the 418 articles that were initially screened, 18 met the inclusion criteria. The studies reported significant associations between PGD and physical health, somatic symptom distress, insomnia severity, blood pressure, bodily distress syndrome, chronic physical diseases, and poor- caregiver health profiles.

Conclusion: Out of the 18 studies eligible for analysis, 13 (72%) established a significantly strong or moderate association between PGD and physical or somatic illness, highlighting the intricate nature of this connection. Further research is required to assess the breadth of physical and somatic health problems associated with PGD and to understand the psychological and biological mechanisms that underpin these observed relationships.

Keywords

prolonged-grief, bereavement, PGD, physical, somatic, somatization, illness

Highlights

- Most studies found a strong to moderate link between PGD and physical/somatic illness.
- PGD impacts caregiver health, somatic distress, insomnia, and comorbid chronic diseases.
- Findings align with PTSD, with clinically relevant psychological and medical effects.

Physical health is defined by the Centers for Disease Control and Prevention as the condition of one's body, with the ability to carry out daily activities without experiencing pain, discomfort, or limitation (Elgaddal et al., 2022). Somatic problems, on the other hand, are physical symptoms that are not caused by an identifiable medical condition (Kolappa et al., 2013). The association between mental health disorders and physical health or somatic symptoms has been consistently reported in the research literature. For example, depression has been identified as a risk factor for long-term physical conditions such as diabetes (Cosgrove et al., 2008; Gonzalez et al., 2008), cancer (Massetti et al., 2017; Massie, 2004), and cardiac disease (Berg et al., 2018; Chaddha et al., 2016; Dhar & Barton, 2016).

Studies conducted by Haug et al. (2004) and Carlehed et al. (2017) have also explored how depression relates to physical symptoms in large community samples, revealing a strong and significant relationship between depression and experiencing functional somatic symptoms. Moreover, Gili et al. (2010) reported a higher prevalence of depression among primary care patients with chronic somatic diseases compared to their physically healthy counterparts.

One disorder that has consistently been found to be associated with physical and somatic problems is post-traumatic stress disorder (PTSD). The ICD-11 outlines PTSD as a mental health condition that can emerge after experiencing a threatening or horrifying event or a sequence of such events (Barbano et al., 2019) and the associated allostatic load has been argued to cause physical morbidity (McFarlane, 2010). There has been a plethora of studies examining the association between PTSD, trauma exposure, physical illness, and somatization, and various systematic reviews have analyzed and described the extant research evidence. An early systematic review by Qureshi et al. (2009) found evidence for a consistent association between PTSD and arthritis, however, mixed results were observed for conditions such as diabetes, coronary heart disease, and stroke. In a more comprehensive systematic review of 62 studies, Pacella et al. (2013) reported a significant association between PTSD and overall poorer physical health outcomes. This encompassed general health symptoms, medical conditions, and health-related quality of life. Gupta's (2013) review further emphasized the link between PTSD and diverse

medical conditions by highlighting the severity of PTSD symptoms to be significantly associated with an increased risk of physical conditions such as hypertension and coronary heart disease. Sleep disturbances, such as sleep paralysis, were also prevalent in PTSD patients, suggesting a multifaceted impact on physical health. Afari et al.'s (2014) systematic review of 71 studies indicated that individuals with reported exposure to trauma were more likely to have functional somatic syndromes, with PTSD also identified as a contributor to cardiovascular and immune-mediated disorders. Lastly, Ryder et al.'s (2018) meta-analysis underscored a robust association between PTSD and increased risks of cardiovascular, metabolic, and musculoskeletal disorders. Collectively, these studies emphasize the intricate connection between PTSD and various physical health outcomes.

Poorer physical and somatic health status also appear to be associated with stressful life experiences such as bereavement. Parkes (1964) was among the first to show a significant correlation between bereavement and physical health in older adults by reporting a 65% increase in medical consultation rates among a sample of widows following bereavement. Large-sample cross-sectional research from Thimm et al. (2020) also demonstrated that severe grief reactions in elderly individuals were significantly associated with self-reported physical health problems as well as an increased use of health services. Additionally, Sillis et al. (2022) and Toblin et al. (2012) have shown that this association was also present in samples of younger people by reporting significant associations between grief and somatic complaints among bereaved university students and infantry soldiers. Moreover, a systematic review by Ennis and Majid (2021) found a significant, positive relationship between bereavement and adverse physical and physiological health outcomes, including inflammation, cardiovascular risk, chronic pain, and mortality.

A significant issue in the field of bereavement has been the lack of acknowledgment of enduring, distressing grief reactions as specific conditions related to grief. There has been a warranted reluctance to pathologize any form of grief, leading to inconsistencies in its definitions and measurement. As a result, depression was often diagnosed instead. However, the inclusion of Prolonged Grief Disorder (PGD) in the 11th Revision of the International Classification of Diseases (ICD-11: WHO, 2019) and the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5-TR) (American Psychiatric Association, 2022) has facilitated a more standardized approach to the study of grief.

On the other hand, the published peer-reviewed literature exploring PGD and physical health problems has also not been systematically examined since PGD was officially classified as a mental health disorder. To address this, we conducted a systematic review of the scientific literature to investigate the association between ICD-11 PGD, and outcomes related to physical health and somatic complaints among bereaved individuals. By synthesizing existing research, this review aims to provide a clearer understanding of the impact of PGD on physical and somatic health, which could inform clinical practices, guide future research, and ultimately contribute to improved care for bereaved

individuals. This review represents the first comprehensive assessment of the evidence for associations between ICD-11 PGD, and physical and somatic health outcomes since PGD's inclusion in the ICD-11.

Method

The protocol for this systematic review was preregistered at the PROSPERO repository (CRD42023471080) on 10/10/2023 (for access, see Cunningham et al., 2023S). To ensure transparency and completeness in the processing and reporting of the results, the PRISMA 2020 guidelines (Page et al., 2021) were adhered to.

Inclusion and Exclusion Criteria

This systematic review incorporated any form of quantitative studies that met the following inclusion criteria:

1. The study reported original, empirical research published in peer-reviewed journals, that utilized quantitative and validated measures of Prolonged Grief Disorder (PGD) and physical or somatic illness.
2. Investigated the association between PGD symptoms from standardized assessment tools and physical and somatic health symptoms.
3. Included a report of quantitative measures of association or group difference such as correlations, odds ratio, *t*-test, etc.

The exclusion criteria were:

1. Non-peer reviewed published research studies.
2. Research that did not employ a quantitative methodology.
3. Single-item quantitative scale measurement of PGD or physical or somatic illness.
4. Non-English language.
5. Studies prior to 2018.

Search Strategy

Four electronic databases Web of Science, MEDLINE, Cochrane Library, and PsycINFO up to the 10th of October 2023 were searched using full-text terms to identify studies reporting an association between PGD, and physical and somatic health symptoms. The search was limited to research studies published in the English language since 2018 that underwent peer review. Searches were conducted using Boolean operators of the following search terms:

> "prolonged grief disorder" OR "prolonged grief" OR "traumatic grief" AND "somatic symptoms" OR "physical illness".

In addition, reference lists of selected studies were screened for any other relevant study.

Reporting Guidelines

This article was prepared in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Page et al., 2021). Adherence to PRISMA standards ensures that the research was reported with transparency and rigor, providing a clear, comprehensive, and reproducible account of the systematic review process. Following these guidelines enhanced the quality and integrity of our research findings.

Data Collection, Extraction and Quality Assessment

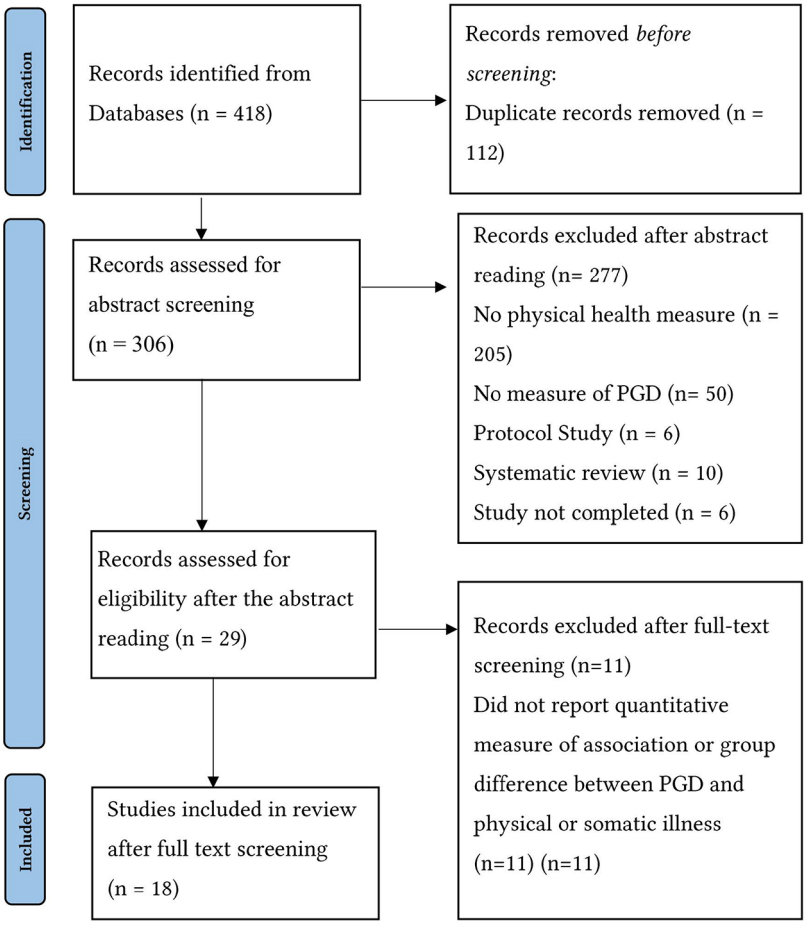
After identifying studies that met the inclusion/exclusion criteria, the researchers retrieved the full-text articles. Two independent reviewers (J.C and C.C) assessed the articles for eligibility, and any disagreements were resolved by consensus. The reviewers were not blinded to the journals or authors of the studies. The researchers created a standardized data extraction sheet to gather information on publication details, study location, methodological features (such as sample size and study design), exposure and outcome measures, PGD type, and the scales used for physical and somatic health outcomes (Supplementary Table 3). The evaluation then focused on the appropriateness of quality assessment tools to measure the level of bias in each study. The resultant tool was a modification of the two most relevant instruments. The Joanna Briggs Institute critical appraisal checklist for analytical cross-sectional studies (JBI) (Joanna Briggs Institute, 2017a) (Supplementary Table 1) was applied to cross-sectional studies, while the JBI critical appraisal checklist for cohort studies (Joanna Briggs Institute, 2017b) (Supplementary Table 2) was employed for longitudinal studies. The description of effect sizes were based on Cohen (1988) descriptions of mean difference (small $d = 0.20$, medium $d = 0.50$, and large $d \geq 0.80$) and correlations (small $r = .10$, medium $r = 0.30$, and large $r \geq .50$).

Results

Details of the search and selection of studies is presented in Figure 1. Out of the initial screening based on title and abstract, 418 articles were identified, 112 of which were duplicates, and once removed, 306 articles remained. There was a high degree of agreement between the two reviewers (24 and 25 articles) in selecting articles that met the inclusion criteria (kappa = .62, $t = 10.90$, $p < .001$). After full-text screening and discussion, a final set of 18 articles were selected to take forward to full review.

Figure 1

PRISMA Flow Diagram Showing the Process for Search and Selection of Studies



Results of the reviewed studies are summarized in Table 1, covering information on associations between PGD, and physical and somatic health outcomes, mode of bereavement, sample characteristics, study design, measures, main findings, and risk of bias.

Clinical Psychology in Europe
2025, Vol. 7(1), Article e14351
https://doi.org/10.32872/cpe.14351

Table 1

Summary of Associations Between PGD and Physical, Somatic Health Outcomes

Study	Mode of bereavement (natural, sudden/ unexpected/ specific illness) Time since bereavement	Sample Size and Characteristics	Study Design (Cross-sectional/longitudinal) Single group or case-control/ comparison	Measures of PGD and Physical and Somatic Health	Main findings	Risk of Bias
Lundorff et al. (2020) Denmark	Loss of a spouse (Natural) 2, 6, and 11 months post-loss	*N* = 857 Female: 69.8% Male: 30.2% Mean age: 70.30	Prospective Longitudinal Single group 11-months post loss	The 13-item Prolonged Grief-scale (PG-13; Prigerson et al., 2009); Revised ICG-R (Prigerson & Jacobs, 2001); The Short-Form Health Survey (Ware et al., 1996)	Physical health significantly predicted the moderate-stable class, EST = -0.041, *SE* = 0.016, *p* = .008 which also included substantial proportions of probable PGD cases, and approached significantly as a predictor of the prolonged grief class EST = -0.041, *SE* = 0.016 *p* = .052	Low
Killikelly et al. (2020) Cross-National study China, Switzerland and the United States	Loss of a loved one (Natural) 6 to 36 months	*N* = 539 Chinese Speaking: 325 German speaking: 214 Female: 72.4% Male: 27.6% Mean age total: 35.39 Chinese sample: 33.14 German sample: 38.71	Cross-sectional Comparison group	International ICD-11 Prolonged Grief Disorder Scale (Killikelly & Maercker, 2017) The Somatic Symptom scale (Gierk et al., 2014)	Correlation coefficients between PGD (IPGDS) and somatic symptoms (SSS-8) showed moderate-level relationships for each of the three IPGDS scales for both samples. Chinese Speaking sample: IPGDS 32 items & SSS-8 = .538, IPGDS 13 items & SSS-8 = .480, and IPGDS standard with cultural supplement & SSS-8 = .540. German Speaking sample: IPGDS 32 items & SSS-8 = .508, IPGDS 13 items & SSS-8 = .458 and IPGDS standard with cultural supplement & SSS-8 = .514	Low
Vogel et al. (2021) Germany	Loss of a loved one Natural at least 6 months previously	*N* = 20 Female: 80% Male: 20% Mean age: 56	Prospective Longitudinal Single group 3 month	PG-13 (Prigerson et al., 2009) The Screening for Somatoform Disorders (SOMS-7D; Rief & Hiller, 2003)	There were no significant differences between the PGD group before and after the person-centered therapy intervention in regard to somatoform symptoms (SOMS-7D) t0-t1 *d* = 0.07, t0-t2 *d* = 0.29 *p* = .665	Moderate
Miller et al. (2020) United States	Loss of a loved one Illness: Cancer 6 to 15 months post-loss	*N* = 198 Female: 61% Male: 39% Mean age: 64.40	Prospective Longitudinal Single group study: However latent class mixture modeling is used to characterize caregiver health by identifying distinct profiles 15 months post loss	PG-13 (Prigerson et al., 2009) Overall health was assessed with 3 separate measures: a single self-report item, The health subscale of Caregiver Reaction Assessment (Given et al., 1992), and The Meeting Physical Demands subscale of the Perceived Self-Care and Daily Living Competencies Scale (Caserta et al., 2004; Utz et al., 2012).	Two distinct health profiles were identified in the total sample. Poorer Health profile group (*n* = 49; 25%) had significantly greater health impact from caregiving *d* = 0.85 (*p* < .0001), more self-reported health problems *d* = 0.53 (*p* = .002), and greater difficulty meeting the physical demands of daily life *d* = 1.16 (*p* < .0001) than the distinct profile (*n* = 149). Regression models showed that having a poorer caregiver health profile was a significant predictor of higher levels of grief symptoms *d* = 4.62 (*p* < .001) in the subsample of participants who were eligible for the bereavement analyses (*N* = 81).	Moderate

Study	Mode of bereavement (natural, sudden/ unexpected/ specific illness) Time since bereavement	Sample Size and Characteristics	Study Design (Cross-sectional/longitudinal) Single group or case-control/ comparison	Measures of PGD and Physical and Somatic Health	Main findings	Risk of Bias
Marcussen et al. (2021) Cross-National study Denmark, Australia and Norway	Loss of a parent Cancer, sudden unexpected, suicide and chronic disease does not provide data on the time since bereavement	*N* =190 Female: 91% Male: 9% Mean age: 17.90	Cross-sectional Comparison group	The PG-13 (Prigerson et al., 2009) The CMDQ 36 (Lu et al., 2008; Tebeka et al., 2016). (Bodily distress syndrome subscale)	Prolonged grief and bodily distress syndrome showed a weak correlation at .24. There was a significant difference between the divorced parental death group n = 52 compared to the non-divorced parental death group n = 130 on bodily distress syndrome d = 0.375 p = .04. The risk of bodily distress syndrome was found to be significantly associated with parental divorce before parental death B = 3.53, p = .009.	Moderate
Lenger et al. (2020) Denmark	Loss of a patient caregivers who experienced the death of patients 6 months after bereavement	*N* = 2,125 Female: 70% Male: 30% Mean age: 62.00	Prospective Longitudinal Single-group However, sample split into with PGD and without PGD 6 months post-loss	The PG-13 (Prigerson et al., 2009) Short Form Health Survey-36 (SF-36) (Ware, 1999) Subscales: Physical functioning, role-physical bodily pain, and general health	Poor physical health status during caregiving predicted prolonged grief disorder: odds ratio 1.05 (95% CI [1.04, 1.07]). The physical subscales of physical functioning odds ratio 1.02, (95% CI [1.02, 1.03]). Role physical odds ratio, 1.02, (95% CI [1.01, 1.02]). Bodily pain odds ratio 1.03, (95% CI [1.02, 1.03]). and general health odds ratio 1.04, (95% CI [1.03, 1.04]). all predicted prolonged grief disorder.	Moderate
Zhou et al. (2020) China	Loss of an only child Violent and non-violent 6 months after bereavement	*N* = 1,030 Female: 62% Male: 38% Mean age: 59.91	Cross-sectional Single-group	The PG-13 (Prigerson et al., 2009) The presence of chronic physical diseases was assessed through a series of binary questions. (Yin et al., 2018). Cumulative Illness Rating Scale (Linn et al., 1968) The number of chronic physical diseases was calculated and coded into a score ranging from zero to six.	More comorbid chronic physical diseases were significantly related to the increased risk of Prolonged grief disorder t = 10.25, β = .33 (95% CI [1.03, 1.51])	Low
Zhang et al. (2020) China	Loss of an only child Disease and accident meantime post-loss 7.6 years	*N* = 149 Female: 60% Male: 40% Mean age: 62.25	Cross-sectional Comparison study	The PG-13 (Prigerson et al., 2009) Information about whether the participants had underlying chronic diseases was recorded. Number of outpatient visits for physical health or other reasons in the past year. It was ranked in 4 levels.	The overall morbidity of osteoarthrosis in the PGD-positive group was significantly higher than that in the PGD-negative group ($\chi 2$ = 7.18, p < .007). There was no significant difference in the number of hospital visits between the two groups.	Low
Pohlkamp et al. (2019) Sweden	Loss of a child Cancer 1 to 5 years after loss	*N* = 225 Female: 59% Male: 41% Mean age: 46.00	Cross-sectional Single-group study	The PG-13 (Prigerson et al., 2009) The Insomnia Severity Index (Morin, 1993)	In symptoms of insomnia, there was no significant effect of years since loss, F 4 = 1.12, .35 and no difference between genders, F 1 = 1.92, .17. There was no significant interaction between years since loss and gender on insomnia F 4 = 1.16, .33.	Moderate

Clinical Psychology in Europe
2025, Vol. 7(1), Article e14351
https://doi.org/10.32872/cpe.14351

Study	Mode of bereavement (natural, sudden/ unexpected/ specific illness) Time since bereavement	Sample Size and Characteristics	Study Design (Cross-sectional/longitudinal) Single group or case-control/ comparison	Measures of PGD and Physical and Somatic Health	Main findings	Risk of Bias
Sveen et al. (2020) Sweden	Loss of a significant other Traumatic event, in the past 5 years	*N* = 123 Female: 81% Male: 19% Mean age: 37.85	Prospective Longitudinal Subsamples: comparison group Ongoing longitudinal study (TRACES study)	The PG-13 (Prigerson et al., 2009) The Symptom Checklist 27 (SCL-27) (Hardt et al., 2004)	PG-13 correlations with the somatization subscale were stronger in the bereavement group 0.57 (p < .001) compared to the comparison group 0.29: Z value = 1.77. There were no significant differences between the bereavement group n = 72 and the comparison group n = 51 on Somatization, D = -0.145.	Moderate
de Lang et al. (2023) Netherlands	Loss of a loved one, Natural accident and suicide 1 month to more than 5 years	*N* = 343 Female: 88% Male: 12% Mean age: 54.00	Prospective Longitudinal Single-group 1 year post-loss	Traumatic Grief Inventory Self-Report Plus (TGI-SR; Lenferink et al., 2022). Based on the TGI-SR (Boelen & Smid, 2017) The Insomnia Severity Index (Morin, 1993)	Correlations between prolonged grief and insomnia symptoms all showed a moderate relationship between the two variables across the three time points at 6 month intervals. PGS at time 1 displayed a weaker correlation over time against insomnia symptoms over time (T1 .39), (T2 .37) (T3 .35) PGS at time 2 ebbed and flowed as a correlation over time against insomnia symptoms overtime (T1 .35), (T2 .47) (T3 .42) PGS at time 3 displayed a stronger correlation over time against insomnia symptoms over time (T1 .36), (T2 .44) (T3 .49). All correlations are significant, p < .001 Participants with higher traits of prolonged grief symptoms also reported higher traits of insomnia symptoms b = .022 (p < .001). For insomnia symptoms, there was a significant autoregressive path (p = .011) and a cross-lagged effect from insomnia to prolonged grief symptoms b = .023 (p = .028).	Moderate
Hennemann et al. (2023) Cross-National study Germany, Switzerland and Ireland	Loss of a close loved one natural, accident, suicide, substance abuse, homicide, and natural disaster. No specific duration since the loss	*N* = 1,337 Female: 76% Male: 24% Mean age: 23.74	Cross-sectional Single-group	International ICD-11 Prolonged Grief Disorder Scale (Killikelly & Maercker, 2017) Somatic Symptom Scale (Gierk et al., 2014)	The direct effect of PGD on somatic symptom distress remained significant when including mediators (c = 0.03, p = .003), indicating a partial mediation of somatic symptom distress. 23% of the variance in explaining somatic symptom distress was explained by prolonged grief disorder b = 0.48 p = < .001. Two-thirds of individuals with possible PGD reported high or very high levels of somatic symptom distress in the SSS-8, which is remarkably higher than prevalences in the general population Non-PGD M = 6.88 PGD M = 12.91 p < .001, d = 0.90.	Low
Carlsson et al. (2023) Sweden	Loss of family member cardiac arrest six months after loss	*N* = 108 Female: 69% Male: 31% Mean age: 61.50	Cross-sectional Single-group: However, subsamples of Spouses and non-spouses were conducted	11 items of the 13-item PG-13 (Prigerson et al., 2009) The RAND-36 measured health-related quality of life (Hays & Morales, 2001)	Spouses reported more problems with symptoms of prolonged grief and self-reported health than non-spouses (p < .001). No significant differences were found between spouses and non-spouses in terms of symptoms of prolonged grief and self-reported	High

Clinical Psychology in Europe
2025, Vol. 7(1), Article e14351
https://doi.org/10.32872/cpe.14351

Study	Mode of bereavement (natural, sudden/ unexpected/ specific illness) Time since bereavement	Sample Size and Characteristics	Study Design (Cross-sectional/longitudinal) Single group or case-control/ comparison	Measures of PGD and Physical and Somatic Health	Main findings	Risk of Bias
					health. 25% of family members in the present study reported their general health as fair or poor and their health to be worse compared to a year ago	
Palitsky et al. (2023) United States	Loss of close relative Natural within the past year	*N* = 59 Female: 69% Male: 31% Mean age: 66	Cross-sectional Single-group	The PG-13 (Prigerson et al., 2009) GE Dinamap Pro 100 BP Monitors: Provided measures of SBP Systolic blood pressure and Diastolic blood pressure DBP	Increases were observed in SBP from baseline (mean [standard error], or *M* [*SE*] = 124.32 [15.01] mm Hg) to immediately post-GR (mean [standard deviation], *M* [*SD*] = 145.43 [25.17], $p < .001$, 95% CI [16.68, 25.52]). DBP also increased from baseline (*M* [*SD*] = 69.05 [8.47]) to immediately post-GR (*M* [*SD*] = 77.15 [10.67], $p < .001$, 95% CI [5.87, 10.34]). Prolonged grief disorder also significantly predicted SBP (B = 0.447, *SE* = 0.215, $p = .042$, 95% CI [0.024, 0.871]).	Low
Kaiser et al. (2022) Germany	Loss of a loved one hematological cancer time since loss not outlined	*N* = 87 Female: 83% Male: 17% Mean age: 47.32 Intervention G = 47.80 WCG = 46.84	Prospective Longitudinal A randomized controlled trial with a waitlist control group 1 year post-loss	The German version of the ICG (Prigerson et al., 1995) 12-item Short-Form Health Survey (Bullinger, 1995)	No significant group interaction was found for prolonged grief and physical health, sleep quality, or somatization. A significant within-group effect of time was found in the IG and the WCG for prolonged grief and somatization at $p = .03$ (WCG) and $p < .001$ (IG) *D* = -0.01	Low
Comtesse et al. (2020) Germany	Loss of a child, partner, parent and other. natural and un-natural 6 months post-loss	*N* = 113 Female: 81% Male: 19% Mean age: 51.68	Prospective Longitudinal Comparison group	The PG-13 (Prigerson et al., 2009) The Screening for Somatoform Disorders (SOMS7D; Rief & Hiller, 2003)	There was no significant difference between the PGD/PGD emp groups and non-PGD/ PGD emp groups in somatization $t = 2.16$ D = 0.643 and $t = 1.62$ D = 0.39. The persistent complex bereavement disorder group showed an approaching significance difference result $p = .055$ compared to the non-PCBD group D = 0.136	High
Maccallum and Bryant (2020) Australia	Loss of a partner, child, parent Sibling or other medical, accident, suicide and Homicide, 6 months after loss	*N* = 215 Female: 82% Male: 18% Mean age: 49.24	Cross-sectional Single-group	The PG-13 (Prigerson et al., 2009) The WHOQOL-BRIEF (Power et al., 1999)	Regularized partial correlation network analysis showed a significant negative association between prolonged grief disorder and physical health EL = -0.02	Moderate
Yıldırım (2023) Turkey	Loss of first-degree relative COVID-19, natural and unnatural deaths 12-24 months after loss	*N* = 68 Female: 85% Male: 15% Mean age: 45.35 PGD group = 41.90 Non PGD group = 48.80	Cross-sectional Single-group: however, subgroups of PGD and NO PGD were used	The PG-13 (Prigerson et al., 2009) Insomnia Severity Index (Morin, 1993)	Positive correlations between PGD severity and insomnia ($r = 0.501$; $p < 0.01$) There was a significant difference between the PGD group *n* = 30 and the Non PGD group *N* = 38 in severity of insomnia $t = 2.63$ $p = .011$	Low

Clinical Psychology in Europe 2025, Vol. 7(1), Article e14351 https://doi.org/10.32872/cpe.14351

Across the 18 studies eligible for examination, 13 (72%) demonstrated a significantly strong or moderate association between PGD and physical or somatic illness. This was displayed across divergent research designs, types of loss and different somatic and physical health problems. Cross-sectional and longitudinal designs were used in all studies. Using a cross-sectional design Killikelly et al. (2020) reported a moderate correlation between PGD (three IPGDS sub-scales) and somatic symptoms, and this is consistent with Hennemann et al. (2023) who reported that a significant proportion of variance (R^2 = 23%) in somatic related distress was attributed to PGD. In contrast, Maccallum and Bryant (2020) identified a negative association between prolonged grief and physical health. In prospective longitudinal studies, Comtesse et al. (2020) found no significant differences in somatization between individuals with PGD and those without it. In contrast, Sveen et al. (2020) highlighted stronger correlations between prolonged grief and somatization in bereavement. Vogel et al. (2021) showed no significant differences in somatoform symptoms pre and post-person-centered therapy, while Kaiser et al. (2022) found no significant group interaction but observed within-group effects over time. However, both intervention studies featured small sample sizes and a 4:1 ratio of females to males, impacting statistical power and generalizability.

The types of loss reported in the studies included in this review were mostly losing a child, spouse/partner, or parent as well as the losses of patients and family members. Zhou et al. (2020) and Zhang et al. (2020) examined Chinese parents who had lost an only child and they reported associations between an increased risk of PGD and chronic physical diseases. Studies on spousal loss by Lundorff et al. (2020) and Carlsson et al. (2023) identified spousal grief symptoms as a predictor of physical health problems. Marcussen et al. (2021) found a strong correlation between prolonged grief and bodily distress syndrome in a sample who had experienced parental loss. Similarly, Lenger et al. (2020) and Miller et al. (2020) showed a significant association between prolonged grief symptoms and poorer physical health in a sample of bereaved caregivers. These findings suggest that the type of relationship with the deceased may influence the nature and severity of health outcomes associated with prolonged grief, with different relationships potentially leading to specific patterns of physical and somatic symptoms.

There were specific health outcomes that were found to be associated with grief. Yıldırım (2023) and de Lang et al. (2023) both reported a significant association between grief severity and insomnia, although this was not replicated in the Pohlkamp et al. (2019) study. The diversity of outcome types that have been investigated is reflected in the study by Palitsky et al. (2023) who found a significant association between PGD and systolic and diastolic pressure.

Supplementary Tables 1 and 2 show the comprehensive evaluation of bias risk for each study, conducted through the Joanna Briggs Institute critical appraisal checklist for analytical cross-sectional and cohort studies. This showed that 40% of cross-sectional studies exhibited moderate to high levels of bias, in contrast to the higher rate of 75% for

the longitudinal studies. Significant heterogeneity was also noted. The primary bias in cross-sectional studies stemmed from the lack of control over confounding variables. Few studies controlled for participant's previous physical health status, which is significant since individuals experiencing loss tend to be older, and older individuals tend to have more physical health complaints (James et al., 2018; Wu et al., 2022). In contrast, incomplete follow-up in cohort studies contributed to the most common element of potential bias.

Discussion

This is the first systematic review of peer-reviewed published studies assessing the association between ICD-11 prolonged grief disorder (PGD) and outcomes related to physical and somatic health among bereaved individuals since PGD was included in the ICD-11. Among the 18 eligible studies, 13 (72%) reported moderate (Carlsson et al., 2023; de Lang et al., 2023; Killikelly et al., 2020; Lenger et al., 2020; Lundorff et al., 2020; Sveen et al., 2020) to strong associations (Hennemann et al., 2023; Marcussen et al., 2021; Miller et al., 2020; Palitsky et al., 2023; Yıldırım, 2023; Zhang et al., 2020; Zhou et al., 2020) between PGD and physical or somatic illness. There were a number of studies that reported non-significant associations, or failed to report *p*-values and were unclear in describing effect sizes (Comtesse et al., 2020; Kaiser et al., 2022; Maccallum & Bryant, 2020; Pohlkamp et al., 2019; Vogel et al., 2021). It appears that there is reliable scientific evidence, with a relatively low risk of bias, that the experience of prolonged grief is associated with poorer physical health and a higher risk of somatization. This prompts inquiry into the underlying mechanisms connecting these phenomena. Several theoretical frameworks, including attachment theory, the stress response syndrome, and the dual-process model, offer potential explanations for these observed associations.

First, attachment theory (Bowlby, 1958, 2018) provides a conceptual basis for understanding the substantial and moderate associations observed between PGD and intimate types of loss. The loss of a child, especially for mothers, has been shown to produce higher rates of PGD compared to other close loved ones (Buur et al., 2024; Goldstein et al., 2019). Attachment theory helps explain these findings due to the intense emotional bonds between parents and children, making such losses especially devastating. Moreover, the significant associations in our review studies by (Zhang et al., 2020; Zhou et al., 2020), indicate an increased risk of PGD and chronic physical diseases among Chinese parents who had lost an only child. Attachment anxiety has also been shown to predict membership into PGD groups over depression and low-symptom groups, demonstrating incremental predictive ability for both prolonged grief and somatic symptoms (Field et al., 2005; King & Werner, 2012; Maccallum & Bryant, 2018). In light of this, future research could employ quantitative measurement scales and tools for assessing attachment styles based on attachment theory. This could explore associations between

attachment types and physical and somatic health outcomes for individuals meeting ICD-11 PGD criteria. If an association between attachment styles and physical/somatic health outcomes were found to be consistent, such findings may help shape practices and policies, such as identifying profiles of attachment types that pose a high risk of physical/somatic health outcomes.

Second, Horowitz's (1986) stress response syndrome (SRS) offers a robust framework for understanding the significant associations between PGD, insomnia, and excessive blood pressure. The SRS delineates between psychological and physiological responses that individuals may undergo following traumatic or highly stressful events. This persistent state of hypervigilance has the potential to magnify the grieving process and contribute to mental health challenges, adversely impacting somatic and physical well-being (Joiner et al., 1999; Riemann et al., 2010). Regarding insomnia, the SRS would explain heightened emotional distress during nighttime, exacerbating the challenges of coping with complicated grief in solitude (Baker et al., 2016; Germain et al., 2005, 2006; Lancel et al., 2020). Furthermore, elevated blood pressure in prolonged grief sufferers may stem from persistent emotional distress and difficulties in adapting to loss, triggering complex stress responses (Mason & Duffy, 2019). In consideration of this evidence, future research could utilize biological markers and neuroimaging techniques to study hyperarousal in PGD, insomnia, and elevated blood pressure. Objective sleep monitoring (polysomnography or actigraphy) could quantify disruptions in sleep architecture. Results may show correlations between hyperarousal markers and specific sleep parameters, supporting interventions such as cognitive-behavioral therapy for insomnia (CBT-I). Advocating for CBT-I inclusion in treatment plans and workplace policies accommodating insomnia due to prolonged grief could be significant. Identifying factors moderating prolonged grief and elevated blood pressure may also inform tailored interventions and prevention strategies.

A broader perspective on the association between PGD and physical health may also be gained by examining how chronic stress and inflammation, which elucidate similar relationships in other mental disorders such as Post-Traumatic Stress Disorder (PTSD) and Major Depressive Disorder (MDE), apply to PGD. Both PTSD and MDE are linked to prolonged activation of the stress response, leading to increased inflammation (Ehlert et al., 2001; Slavich & Irwin, 2014; Wichmann et al., 2017). This inflammatory process contributes to various physical health issues, including cardiovascular disease and metabolic disorders (Black & Garbutt, 2002; Liu et al., 2017). Given that PGD involves sustained emotional distress, analogous stress-induced inflammatory pathways may also underlie the physical health problems observed in PGD. Moreover, PTSD and MDE are associated with somatic complaints such as chronic pain and gastrointestinal issues (Gupta, 2013; Thom et al., 2019), which may similarly manifest in PGD as physical symptoms due to intense grief and emotional turmoil. By exploring these parallels, researchers may

gain a deeper understanding of the mechanisms through which PGD impacts physical well-being, thus guiding future research and clinical practice.

Lastly, the dual process model of coping with bereavement (Stroebe & Schut, 1999) offers a bidirectional insight into the significant associations between PGD and physical/somatic illness following unnatural loss through loss-oriented and restoration-oriented stressors (Tur et al., 2022). Unnatural or traumatic loss poses unique challenges to the grieving process, triggering intense emotions such as shock, disbelief, and intrusive thoughts (Layne et al., 2018; Lobb et al., 2010; Walsh, 2007). These emotions fall under loss-oriented stressors, as they prompt individuals to face the reality of their abnormal loss. Simultaneously, coping with the aftermath of unnatural loss involves practical challenges, such as legal processes, funeral arrangements, and dealing with the societal aftermath. It is highly plausible that individuals experiencing unnatural loss may oscillate between addressing their emotional pain and engaging in such constructive tasks. For instance, someone grieving the sudden abnormal loss of a loved one in an accident may alternate between processing the emotional trauma and dealing with the administrative aspects, such as legal procedures or insurance matters. This consistent fluctuation may create cognitive dissonance (Festinger, 1957) in those experiencing unnatural loss which may elucidate the substantially significant associations observed between PGD and physical and somatic illness through abnormal loss circumstances. For example, Dickerson and Kemeny (2004) have shown how stressors that involve social-evaluative threats a key component of cognitive dissonance lead to significant increases in cortisol levels, which has been shown to suppress the immune system, making individuals more vulnerable to illness. In the context of PGD, the ongoing internal conflict and chronic stress may result in a sustained physiological response, thereby weakening immunity and increasing susceptibility to physical ailments. Additionally, cognitive dissonance has also been linked to an increased risk of cardiovascular disease. Linden et al. (2007) found that stress arising from conflicting emotions or behaviours which are key elements of cognitive dissonance significantly heightens the risk of hypertension and other cardiovascular problems. For individuals with PGD, the persistent cognitive dissonance they experience may intensify their stress, thereby increasing the likelihood of developing cardiovascular issues. Keeping this in consideration, future research could refine and adapt existing prolonged grief and coping scales to better align with the nuances of the dual process model. A dual-process model questionnaire could focus on addressing specific components of prolonged grief that contribute to cognitive dissonance and potential physical and somatic health complications. Policymakers could integrate such screening tools into routine health assessments, while employers and community organizations could offer more targeted support programs.

The studies under review exhibited both strengths and limitations. They notably demonstrated consistency in measuring PGD, alongside showcasing geographical and cultural diversity, which enriched external and ecological validity. However, this cultural

diversity may explain the assorted findings found across the reviewed studies regarding the strength of the association between PGD and physical and somatic health outcomes. Future research could investigate this by examining how cultural factors influence this relationship, potentially through incorporating culturally sensitive measures in assessments. The majority of studies also presented substantial sample sizes, often supported by reported power analyses. On the other hand, the bias risk evaluation revealed differences in bias levels. Cross-sectional studies tended to have lower bias than longitudinal studies. Moderate bias was noted in cross-sectional studies, while higher bias was observed in cohort studies. Both designs exhibited relatively low levels of high bias. Future studies, especially in cohort designs, can benefit from proactive strategies and experimental designs to mitigate bias and enhance generalizability. However, given the intrinsic difficulty in manipulating grief as an emotional state in experimental settings, researchers must approach this challenge with caution and creativity. Methodologically, it's noteworthy that the majority of studies relied on self-reported measures that lacked control for confounding variables, while only 44% utilized longitudinal methodology, potentially impacting internal validity and result interpretability. Substantial heterogeneity was observed among the studies analyzed, with four distinct scales employed to evaluate physical health and five to measure somatic health outcomes. Moreover, three studies adopted alternative quantification methods, including the use of monitors, chronic disease assessments, and outpatient visits. This disparity in measurement complicates direct result comparisons, as it's unclear if differences stem from variable characteristics or scale usage. Developing universal physical and somatic health scales could address this, offering standardized measures across cultures. This would aid cross-cultural comparisons and deepen our understanding of physical and somatic health outcomes. The included studies also exhibited a fairly high mean age of 50 which may not capture the unique prolonged grief experiences of younger individuals who may have different coping mechanisms, support structures, and life contexts compared to older adults.

In conclusion, this pioneering review on PGD's association with physical and somatic illness exhibited numerous strengths such as the consistent measurement of PGD, substantial sample sizes, and a high level of regional diversity. However, limitations included disparities in bias levels between transverse and cohort studies, heterogeneity in attaining the measurement of physical and somatic illness and the use of self-reported measures that lacked control for confounding variables. The reviewed results revealed a hierarchy of associations. Most studies demonstrated a significantly strong or moderate association between PGD and physical or somatic illness. Notable findings include PGD's impact on caregiver health decline, somatic symptom distress, insomnia severity, and comorbid chronic diseases such as osteoarthrosis and elevated blood pressure. These results are consistent with PTSD findings and highlight the clinically relevant effect sizes both psychologically and medically. These findings may assist in the differential diagnosis of PGD by emphasizing the unique combination of psychological and physio-

logical symptoms, which can help distinguish PGD from other disorders such as PTSD. Given the significant impact of PGD on physical health, it is important to consider these physiological symptoms more prominently in the diagnostic process to ensure comprehensive assessment and appropriate treatment.

An important additional consideration is the impact of behavioural changes associated with PGD on overall health. PGD has been shown to cause behavioural changes that contribute to poor physical and mental health. For instance, individuals with complicated grief may engage in behaviours such as binge drinking, smoking, and a lack of physical activity (Stroebe et al., 2007). These behaviours can exacerbate chronic illness, which in turn impacts an individual's mental health and affects their ability to participate effectively in therapy (Lando, 2006). Understanding these interactions is crucial, as social withdrawal known as a common response in PGD (Szuhany et al., 2021) can lead to further physical and mental health problems. A comprehensive approach to PGD treatment must consider these behavioural changes and their impact on overall health to enhance therapeutic outcomes and support holistic recovery.

Future research avenues include integrating quantitative tools based on attachment theory for intimate losses in routine PGD screenings or employing biological markers and neuroimaging techniques to study hyperarousal in PGD, insomnia, and elevated blood pressure. Additionally, the dual process model of coping with bereavement could be utilized through a standardized questionnaire tailored to measure the framework, potentially predicting physical or somatic health issues among prolonged grief sufferers. However, future studies must prioritize methodological rigor, diverse participant samples, and ethical standards to ensure valid and applicable findings in clinical practice.

Funding: The authors have no funding to report.

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: The authors have declared that no competing interests exist.

Ethics Statement: This article includes studies that involved human participants, all of which received appropriate ethical approval from their respective ethics committees. Ethical approval was not required for this review, as it involved the analysis of publicly available data.

Social Media Accounts: @mark_shevlin_

Preregistration: The protocol for this systematic review was preregistered at the PROSPERO repository (CRD42023471080), and is available at https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=471080

Reporting Guidelines: This article was prepared following the PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Page et al., 2021).

Data Availability: All materials are freely available from the corresponding author on request.

Supplementary Materials

The Supplementary Materials contain the following items:

- **Preregistered PROSPERO Protocol** (Cunningham et al., 2023S)
- **Online Appendices** (Cunningham et al., 2025S):
  - **Appendix A**: JBI Critical Appraisal Checklist for Analytical Cross-Sectional Studies.
  - **Appendix B**: JBI Critical Appraisal Checklist for Cohort Studies.
  - **Appendix C:** Standardized Data Extraction Sheet.
  - **Appendix D:** Descriptions of the included studies.
  - **Appendix E:** References from Systematic Review.

Index of Supplementary Materials

Cunningham, J., Shevlin, M., Cerda, C., & McElroy, E. (2023S). *ICD-11 prolonged grief disorder, physical health and somatic problems: A systematic review* [Preregistration]. PROSPERO. https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=471080

Cunningham, J., Shevlin, M., Cerda, C., & McElroy, E. (2025S). *Supplementary materials to "ICD-11 prolonged grief disorder, physical health, and somatic problems: A systematic review"* [Online appendices]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.16037

References

Note. References marked with an asterisk (*) are the articles included in the systematic review.

Afari, N., Ahumada, S. M., Wright, L. S., Mostoufi, S., Golnari, G., Reis, V., & Cuneo, J. G. (2014). Psychological trauma and functional somatic syndromes. *Psychosomatic Medicine, 76*(1), 2–11. https://doi.org/10.1097/PSY.0000000000000010

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text revision). American Psychiatric Association Publishing.

Baker, A. W., Keshaviah, A., Horenstein, A., Goetter, E. M., Mauro, C., Reynolds, C. F., Zisook, S., Shear, M. K., & Simon, N. M. (2016). The role of avoidance in complicated grief: A detailed examination of the Grief-Related Avoidance Questionnaire (GRAQ) in a large sample of individuals with complicated grief. *Journal of Loss and Trauma, 21*(6), 533–547. https://doi.org/10.1080/15325024.2016.1157412

Barbano, A. C., Van Der Mei, W. F., Bryant, R. A., Delahanty, D. L., deRoon-Cassini, T., Matsuoka, Y., Olff, M., Qi, W., Ratanatharathorn, A., Schnyder, U., Seedat, S., Kessler, R. C., Koenen, K. C., & Shalev, A. Y. (2019). Clinical implications of the proposed ICD-11 PTSD diagnostic criteria. *Psychological Medicine, 49*(3), 483–490. https://doi.org/10.1017/S0033291718001101

Boelen, P. A., & Smid, G. E. (2017). The Traumatic Grief Inventory Self-Report Version (TGI-SR): Introduction and preliminary psychometric evaluation. *Journal of Loss and Trauma, 22*(3), 196–212. https://doi.org/10.1080/15325024.2017.1284488

Bowlby, J. (1958). The nature of the child's tie to his mother. *The International Journal of Psycho-Analysis, 39*, 350–373.

Bowlby, J. (2018). The nature of the child's tie to his mother. In A. C. Furman & S. T. Levy (Eds.), *Influential papers from the 1950s* (pp. 222–273). Routledge.

Berg, S. K., Rasmussen, T. B., Thrysoee, L., Thorup, C. B., Borregaard, B., Christensen, A. V., Mols, R. E., Juel, K., & Ekholm, O. (2018). Mental health is a risk factor for poor outcomes in cardiac patients: Findings from the National Denheart survey. *Journal of Psychosomatic Research, 112*, 66–72. https://doi.org/10.1016/j.jpsychores.2018.07.002

Black, P. H., & Garbutt, L. D. (2002). Stress, inflammation and cardiovascular disease. *Journal of Psychosomatic Research, 52*(1), 1–23. https://doi.org/10.1016/S0022-3999(01)00302-6

Bullinger, M. (1995). German translation and psychometric testing of the SF-36 health survey: Preliminary results from the IQOLA project. *Social Science & Medicine, 41*(10), 1359–1366. https://doi.org/10.1016/0277-9536(95)00115-N

Buur, C., Zachariae, R., Komischke-Konnerup, K., Marello, M. M., Schierff, L. H., & O'Connor, M. (2024). Risk factors for prolonged grief symptoms: A systematic review and meta-analysis. *Clinical Psychology Review, 107*, Article 102375. https://doi.org/10.1016/j.cpr.2023.102375

Carlehed, G., Katz, J., & Nordin, S. (2017). Somatic symptoms of anxiety and depression: A population-based study. *Mental Health & Prevention, 6*, 57–62. https://doi.org/10.1016/j.mhp.2017.03.005

*Carlsson, N., Alvariza, A., Bremer, A., Axelsson, L., & Årestedt, K. (2023). Symptoms of prolonged grief and self-reported health among bereaved family members of persons who died from sudden cardiac arrest. *OMEGA – Journal of Death and Dying, 87*(1), 66–86. https://doi.org/10.1177/00302228211018115

Caserta, M. S., Lund, D. A., & Obray, S. J. (2004). Promoting self-care and daily living skills among older widows and widowers: Evidence from the Pathfinders Demonstration Project. *OMEGA – Journal of Death and Dying, 49*(3), 217–236. https://doi.org/10.2190/9BH0-N565-Y40G-QDN9

Chaddha, A., Robinson, E. A., Kline–Rogers, E., Alexandris-Souphis, T., & Rubenfire, M. (2016). Mental health and cardiovascular disease. *The American Journal of Medicine, 129*(11), 1145–1148. https://doi.org/10.1016/j.amjmed.2016.05.018

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

*Comtesse, H., Vogel, A., Kersting, A., Rief, W., Steil, R., & Rosner, R. (2020). When does grief become pathological? Evaluation of the ICD-11 diagnostic proposal for prolonged grief in a treatment-seeking sample. *European Journal of Psychotraumatology, 11*(1), Article 1694348. https://doi.org/10.1080/20008198.2019.1694348

Cosgrove, M. P., Sargeant, L. A., & Griffin, S. J. (2008). Does depression increase the risk of developing type 2 diabetes? *Occupational Medicine, 58*(1), 7–14. https://doi.org/10.1093/occmed/kqm105

*de Lang, T. A., Buyukcan-Tetik, A., de Jong, P. J., Lancel, M., & Eisma, M. C. (2023). Cross-lagged analyses of prolonged grief and depression symptoms with insomnia symptoms. *Behavior Therapy, 54*(3), 510–523. https://doi.org/10.1016/j.beth.2022.12.004

Dhar, A. K., & Barton, D. A. (2016). Depression and the link with cardiovascular disease. *Frontiers in Psychiatry, 7*, Article 33. https://doi.org/10.3389/fpsyt.2016.00033

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychological Bulletin, 130*(3), 355–391. https://doi.org/10.1037/0033-2909.130.3.355

Ehlert, U., Gaab, J., & Heinrichs, M. (2001). Psychoneuroendocrinological contributions to the etiology of depression, posttraumatic stress disorder, and stress-related bodily disorders: The role of the hypothalamus–pituitary–adrenal axis. *Biological Psychology, 57*(1–3), 141–152. https://doi.org/10.1016/S0301-0511(01)00092-8

Elgaddal, N., Kramarow, E. A., & Reuben, C. A. (2022). *Physical activity among adults aged 18 and over: United States, 2020.* US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Ennis, J., & Majid, U. (2021). "Death from a broken heart": A systematic review of the relationship between spousal bereavement and physical and physiological health outcomes. *Death Studies, 45*(7), 538–551. https://doi.org/10.1080/07481187.2019.1661884

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford University Press.

Field, N. P., Gao, B., & Paderna, L. (2005). Continuing bonds in bereavement: An attachment theory-based perspective. *Death Studies, 29*(4), 277–299. https://doi.org/10.1080/07481180590923689

Germain, A., Caroff, K., Buysse, D. J., & Shear, M. K. (2005). Sleep quality in complicated grief. *Journal of Traumatic Stress, 18*(4), 343–346. https://doi.org/10.1002/jts.20035

Germain, A., Shear, K., Monk, T. H., Houck, P. R., Reynolds, C. F., Frank, E., & Buysse, D. J. (2006). Treating complicated grief: Effects on sleep quality. *Behavioral Sleep Medicine, 4*(3), 152–163. https://doi.org/10.1207/s15402010bsm0403_2

Gierk, B., Kohlmann, S., Kroenke, K., Spangenberg, L., Zenger, M., Brähler, E., & Löwe, B. (2014). The Somatic Symptom Scale–8 (SSS-8). *JAMA Internal Medicine, 174*(3), 399–407. https://doi.org/10.1001/jamainternmed.2013.12179

Gili, M., Comas, A., García-García, M., Monzón, S., Serrano-Blanco, A., & Roca, M. (2010). Comorbidity between common mental disorders and chronic somatic diseases in primary care patients. *General Hospital Psychiatry, 32*(3), 240–245. https://doi.org/10.1016/j.genhosppsych.2010.01.013

Given, C. W., Given, B., Stommel, M., Collins, C. E., King, S., & Franklin, S. (1992). The Caregiver Reaction Assessment (CRA) for caregivers to persons with chronic physical and mental impairments. *Research in Nursing & Health, 15*(4), 271–283. https://doi.org/10.1002/nur.4770150406

Goldstein, R. D., Petty, C. R., Morris, S. E., Human, M., Odendaal, H., Elliott, A., Tobacco, D., Angal, J., Brink, L., Kinney, H. C., & Prigerson, H. G. (2019). Pre-loss personal factors and prolonged grief disorder in bereaved mothers. *Psychological Medicine, 49*(14), 2370–2378. https://doi.org/10.1017/S0033291718003264

Gonzalez, J. S., Safren, S. A., Delahanty, L. M., Cagliero, E., Wexler, D. J., Meigs, J. B., & Grant, R. W. (2008). Symptoms of depression prospectively predict poorer self-care in patients with Type 2 diabetes. *Diabetic Medicine, 25*(9), 1102–1107. https://doi.org/10.1111/j.1464-5491.2008.02535.x

Gupta, M. A. (2013). Review of somatic symptoms in post-traumatic stress disorder. *International Review of Psychiatry, 25*(1), 86–99. https://doi.org/10.3109/09540261.2012.736367

Hardt, J., Egle, U. T., Kappis, B., Hessel, A., & Brähler, E. (2004). Die Symptom-Checkliste SCL-27 [The Symptom Checklist SCL-27]. *Psychotherapie, Psychosomatik, Medizinische Psychologie, 54*(5), 214–223. https://doi.org/10.1055/s-2003-814786

Haug, T. T., Mykletun, A., & Dahl, A. A. (2004). The association between anxiety, depression, and somatic symptoms in a large population: The HUNT-II study. *Psychosomatic Medicine, 66*(6), 845–851. https://doi.org/10.1097/01.psy.0000145823.85658.0c

Hays, R. D., & Morales, L. S. (2001). The RAND-36 measure of health-related quality of life. *Annals of Medicine, 33*(5), 350–357. https://doi.org/10.3109/07853890109002089

*Hennemann, S., Killikelly, C., Hyland, P., Maercker, A., & Witthöft, M. (2023). Somatic symptom distress and ICD-11 prolonged grief in a large intercultural sample. *European Journal of Psychotraumatology, 14*(2), Article 2254584. https://doi.org/10.1080/20008066.2023.2254584

Horowitz, M. J. (1986). Stress-response syndromes: A review of posttraumatic and adjustment disorders. *Psychiatric Services, 37*(3), 241–249. https://doi.org/10.1176/ps.37.3.241

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, Z., Abera, S. F., Abil, O. Z., Abraha, H. N., Abu-Raddad, L. J., Abu-Rmeileh, N. M. E., Accrombessi, M., . . . Murray, C. J. L. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet, 392*(10159), 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7

Joanna Briggs Institute. (2017a). *Checklist for analytical cross sectional studies.* https://jbi.global/sites/default/files/2019-05/JBI_Critical_Appraisal-Checklist_for_Analytical_Cross_Sectional_Studies2017_0.pdf

Joanna Briggs Institute. (2017b). *Checklist for cohort studies.* https://jbi.global/sites/default/files/2020-08/Checklist_for_Cohort_Studies.pdf

Joiner, T. E., Jr., Steer, R. A., Beck, A. T., Schmidt, N. B., Rudd, M. D., & Catanzaro, S. J. (1999). Physiological hyperarousal: Construct validity of a central aspect of the tripartite model of depression and anxiety. *Journal of Abnormal Psychology, 108*(2), 290–298. https://doi.org/10.1037/0021-843X.108.2.290

*Kaiser, J., Nagl, M., Hoffmann, R., Linde, K., & Kersting, A. (2022). Therapist-assisted web-based intervention for prolonged grief disorder after cancer bereavement: Randomized controlled trial. *JMIR Mental Health, 9*(2), Article e27642. https://doi.org/10.2196/27642

Killikelly, C., & Maercker, A. (2017). Prolonged grief disorder for ICD-11: The primacy of clinical utility and international applicability. *European Journal of Psychotraumatology, 8*(Suppl 6), Article 1476441. https://doi.org/10.1080/20008198.2018.1476441

*Killikelly, C., Zhou, N., Merzhvynska, M., Stelzer, E.-M., Dotschung, T., Rohner, S., Sun, L. H., & Maercker, A. (2020). Development of the International Prolonged Grief Disorder Scale for the ICD-11: Measurement of core symptoms and culture items adapted for Chinese and German-speaking samples. *Journal of Affective Disorders, 277*, 568–576. https://doi.org/10.1016/j.jad.2020.08.057

King, L. C., & Werner, P. D. (2012). Attachment, social support, and responses following the death of a companion animal. *OMEGA – Journal of Death and Dying, 64*(2), 119–141. https://doi.org/10.2190/OM.64.2.b

Kolappa, K., Henderson, D. C., & Kishore, S. P. (2013). No physical health without mental health: Lessons unlearned? *Bulletin of the World Health Organization, 91*(1), 3–3A. https://doi.org/10.2471/BLT.12.115063

Lancel, M., Stroebe, M., & Eisma, M. C. (2020). Sleep disturbances in bereavement: A systematic review. *Sleep Medicine Reviews, 53*, Article 101331. https://doi.org/10.1016/j.smrv.2020.101331

Lando, H. A. (2006). Reflections on 30+ years of smoking cessation research: From the individual to the world. *Drug and Alcohol Review, 25*(1), 5–14. https://doi.org/10.1080/09595230500459461

Layne, C. M., Kaplow, J. B., Oosterhoff, B., Hill, R. M., & Pynoos, R. S. (2018). The interplay between posttraumatic stress and grief reactions in traumatically bereaved adolescents: When trauma, bereavement, and adolescence converge. *Adolescent Psychiatry, 7*(4), 266–285. https://doi.org/10.2174/2210676608666180306162544

*Lenger, M. K., Neergaard, M. A., Guldin, M.-B., & Nielsen, M. K. (2020). Poor physical and mental health predicts prolonged grief disorder: A prospective, population-based cohort study on caregivers of patients at the end of life. *Palliative Medicine, 34*(10), 1416–1424. https://doi.org/10.1177/0269216320948007

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

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

Linden, W., Phillips, M. J., & Leclerc, J. (2007). Psychological treatment of cardiac patients: A meta-analysis. *European Heart Journal, 28*(24), 2972–2984. https://doi.org/10.1093/eurheartj/ehm504

Linn, B. S., Linn, M. W., & Gurel, L. (1968). Cumulative Illness Rating Scale. *Journal of the American Geriatrics Society, 16*(5), 622–626. https://doi.org/10.1111/j.1532-5415.1968.tb02103.x

Liu, Y.-Z., Wang, Y.-X., & Jiang, C.-L. (2017). Inflammation: The common pathway of stress-related diseases. *Frontiers in Human Neuroscience, 11*, Article 316. https://doi.org/10.3389/fnhum.2017.00316

Lu, W., Mueser, K. T., Rosenberg, S. D., & Jankowski, M. K. (2008). Correlates of adverse childhood experiences among adults with severe mood disorders. *Psychiatric Services, 59*(9), 1018–1026. https://doi.org/10.1176/ps.2008.59.9.1018

*Lundorff, M., Bonanno, G. A., Johannsen, M., & O'Connor, M. (2020). Are there gender differences in prolonged grief trajectories? A registry-sampled cohort study. *Journal of Psychiatric Research, 129*, 168–175. https://doi.org/10.1016/j.jpsychires.2020.06.030

Maccallum, F., & Bryant, R. A. (2018). Prolonged grief and attachment security: A latent class analysis. *Psychiatry Research, 268*, 297–302. https://doi.org/10.1016/j.psychres.2018.07.038

*Maccallum, F., & Bryant, R. A. (2020). A network approach to understanding quality of life impairments in prolonged grief disorder. *Journal of Traumatic Stress, 33*(1), 106–115. https://doi.org/10.1002/jts.22383

*Marcussen, J., Hounsgaard, L., O'Connor, M., Möller, S., Wilson, R., & Thuen, F. (2021). Parental death in young adults with divorced compared to non-divorced parents: The effect on prolonged grief and mental health. *Death Studies, 45*(6), 437–450. https://doi.org/10.1080/07481187.2019.1648337

Mason, T. M., & Duffy, A. (2019). Complicated grief and cortisol response: An integrative review of the literature. *Journal of the American Psychiatric Nurses Association, 25*(3), 181–188. https://doi.org/10.1177/1078390318807966

Massetti, G. M., Thomas, C. C., King, J., Ragan, K., & Lunsford, N. B. (2017). Mental health problems and cancer risk factors among young adults. *American Journal of Preventive Medicine, 53*(3, Suppl. 1), S30–S39. https://doi.org/10.1016/j.amepre.2017.04.023

Massie, M. J. (2004). Prevalence of depression in patients with cancer. *Journal of the National Cancer Institute Monographs, 2004*(32), 57–71. https://doi.org/10.1093/jncimonographs/lgh014

McFarlane, A. C. (2010). The long-term costs of traumatic stress: Intertwined physical and psychological consequences. *World Psychiatry, 9*(1), 3–10. https://doi.org/10.1002/j.2051-5545.2010.tb00254.x

*Miller, L. M., Utz, R. L., Supiano, K., Lund, D. A., & Caserta, M. S. (2020). Health profiles of spouse caregivers: The role of active coping and the risk for developing prolonged grief symptoms. *Social Science & Medicine, 266*, Article 113455. https://doi.org/10.1016/j.socscimed.2020.113455

Morin, C. M. (1993). *Insomnia Severity Index*. PsycTESTS. https://doi.org/10.1037/t07115-000

Pacella, M. L., Hruska, B., & Delahanty, D. L. (2013). The physical health consequences of PTSD and PTSD symptoms: A meta-analytic review. *Journal of Anxiety Disorders, 27*(1), 33–46. https://doi.org/10.1016/j.janxdis.2012.08.004

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T., Mulrow, C. D., Shamseer, L., Tetzlaff, J., Akl, E. A., Brennan, S., Chou, R., Glanville, J., Grimshaw, J., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E., Mayo-Wilson, E., McDonald, S., . . . Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research Ed.), 372*, Article 71. https://doi.org/10.1136/bmj.n71

*Palitsky, R., Wilson, D., Friedman, S. E., Ruiz, J. M., Sullivan, D., & O'Connor, M.-F. (2023). The relationship of prolonged grief disorder symptoms with hemodynamic response to grief recall among bereaved adults. *Psychosomatic Medicine, 85*(6), 545–550. https://doi.org/10.1097/PSY.0000000000001223

Parkes, C. M. (1964). Effects of bereavement on physical and mental health—A study of the medical records of widows. *British Medical Journal, 2*(5404), 274–279. https://doi.org/10.1136/bmj.2.5404.274

*Pohlkamp, L., Kreicbergs, U., & Sveen, J. (2019). Bereaved mothers' and fathers' prolonged grief and psychological health 1 to 5 years after loss—A nationwide study. *Psycho-Oncology, 28*(7), 1530–1536. https://doi.org/10.1002/pon.5112

Power, M., Bullinger, M., & Harper, A. E. (1999). The World Health Organization WHOQOL-100: Tests of the universality of quality of life in 15 different cultural groups worldwide. *Health Psychology, 18*(5), 495–505. https://doi.org/10.1037/0278-6133.18.5.495

Prigerson, H. G., Horowitz, M. J., Jacobs, S., Parkes, C. M., Aslan, M., Goodkin, K., Raphael, B., Marwit, S. J., Wortman, C. B., Neimeyer, R. A., Bonanno, G. A., Block, S. D., Kissane, D. W., Boelen, P. A., Maercker, A., Litz, B. T., Johnson, J. G., First, M. B., & Maciejewski, P. K. (2009). Prolonged grief disorder: Psychometric validation of criteria proposed for DSM-V and ICD-11. *PLoS Medicine, 6*(8), Article e1000121. https://doi.org/10.1371/journal.pmed.1000121

Prigerson, H. O., & Jacobs, S. C. (2001). Traumatic grief as a distinct disorder: A rationale, consensus criteria, and a preliminary empirical test. In M. S. Stroebe, R. O. Hansson, W.

Stroebe, & H. Schut (Eds.), *Handbook of bereavement research: Consequences, coping, and care* (pp. 613–645). American Psychological Association. https://doi.org/10.1037/10436-026

Prigerson, H. G., Maciejewski, P. K., Reynolds, C. F., Bierhals, A. J., Newsom, J. T., Fasiczka, A. L., Frank, E., Doman, J., & Miller, M. W. (1995). Inventory of Complicated Grief: A scale to measure maladaptive symptoms of loss. *Psychiatry Research, 59*(1–2), 65–79. https://doi.org/10.1016/0165-1781(95)02757-2

Qureshi, S. U., Pyne, J. M., Magruder, K. M., Schulz, P. E., & Kunik, M. E. (2009). The link between post-traumatic stress disorder and physical comorbidities: A systematic review. *The Psychiatric Quarterly, 80*(2), 87–97. https://doi.org/10.1007/s11126-009-9096-4

Rief, W., & Hiller, W. (2003). A new approach to the assessment of the treatment effects of somatoform disorders. *Psychosomatics, 44*(6), 492–498. https://doi.org/10.1176/appi.psy.44.6.492

Riemann, D., Spiegelhalder, K., Feige, B., Voderholzer, U., Berger, T., Perlis, M. L., & Nissen, C. (2010). The hyperarousal model of insomnia: A review of the concept and its evidence. *Sleep Medicine Reviews, 14*(1), 19–31. https://doi.org/10.1016/j.smrv.2009.04.002

Ryder, A. L., Azcarate, P. M., & Cohen, B. E. (2018). PTSD and physical health. *Current Psychiatry Reports, 20*(12), Article 116. https://doi.org/10.1007/s11920-018-0977-9

Sillis, L., Claes, L., & Andriessen, K. (2022). Association between grief and somatic complaints in bereaved university and college students. *International Journal of Environmental Research and Public Health, 19*(19), Article 12108. https://doi.org/10.3390/ijerph191912108

Slavich, G. M., & Irwin, M. R. (2014). From stress to inflammation and major depressive disorder: A social signal transduction theory of depression. *Psychological Bulletin, 140*(3), 774–815. https://doi.org/10.1037/a0035302

Stroebe, M., & Schut, H. (1999). The dual process model of coping with bereavement: Rationale and description. *Death Studies, 23*(3), 197–224. https://doi.org/10.1080/074811899201046

Stroebe, M., Schut, H., & Stroebe, W. (2007). Health outcomes of bereavement. *Lancet, 370*(9603), 1960–1973. https://doi.org/10.1016/S0140-6736(07)61816-9

*Sveen, J., Bondjers, K., Heinsoo, J., & Arnberg, F. K. (2020). Psychometric evaluation of the Swedish version of the Prolonged Grief Disorder-13 (PG-13) in a bereaved mixed trauma sample. *Frontiers in Psychiatry, 11*, Article 541789. https://doi.org/10.3389/fpsyt.2020.541789

Szuhany, K. L., Malgaroli, M., Miron, C. D., & Simon, N. M. (2021). Prolonged grief disorder: Course, diagnosis, assessment, and treatment. *Focus – American Psychiatric Publishing, 19*(2), 161–172. https://doi.org/10.1176/appi.focus.20200052

Tebeka, S., Hoertel, N., Dubertret, C., & Le Strat, Y. (2016). Parental divorce or death during childhood and adolescence and its association with mental health. *The Journal of Nervous and Mental Disease, 204*(9), 678–685. https://doi.org/10.1097/NMD.0000000000000549

Thimm, J. C., Kristoffersen, A. E., & Ringberg, U. (2020). The prevalence of severe grief reactions after bereavement and their associations with mental health, physical health, and health service utilization: A population-based study. *European Journal of Psychotraumatology, 11*(1), Article 1844440. https://doi.org/10.1080/20008198.2020.1844440

Thom, R., Silbersweig, D. A., & Boland, R. J. (2019). Major depressive disorder in medical illness: A review of assessment, prevalence, and treatment options. *Psychosomatic Medicine, 81*(3), 246–255. https://doi.org/10.1097/PSY.0000000000000678

Toblin, R. L., Riviere, L. A., Thomas, J. L., Adler, A. B., Kok, B. C., & Hoge, C. W. (2012). Grief and physical health outcomes in U.S. soldiers returning from combat. *Journal of Affective Disorders, 136*(3), 469–475. https://doi.org/10.1016/j.jad.2011.10.048

Tur, C., Campos, D., Suso-Ribera, C., Kazlauskas, E., Castilla, D., Zaragozá, I., García-Palacios, A., & Quero, S. (2022). An internet-delivered cognitive-behavioral therapy (iCBT) for prolonged grief disorder (PGD) in adults: A multiple-baseline single-case experimental design study. *Internet Interventions, 29*, Article 100558. https://doi.org/10.1016/j.invent.2022.100558

Utz, R. L., Caserta, M. S., & Lund, D. A. (2012). Grief, depressive symptoms, and physical health among recently bereaved spouses. *The Gerontologist, 52*(4), 460–471. https://doi.org/10.1093/geront/gnr110

*Vogel, A., Comteße, H., Nocon, A., Kersting, A., Rief, W., Steil, R., & Rosner, R. (2021). Feasibility of present-centered therapy for prolonged grief disorder: Results of a pilot study. *Frontiers in Psychiatry, 12*, Article 534664. https://doi.org/10.3389/fpsyt.2021.534664

Walsh, F. (2007). Traumatic loss and major disasters: Strengthening family and community resilience. *Family Process, 46*(2), 207–227. https://doi.org/10.1111/j.1545-5300.2007.00205.x

Ware, J. E., Jr. (1999). SF-36 health survey. In M. E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment* (pp. 1227–1246). Lawrence Erlbaum Associates.

Ware, J. E., Jr., Kosinski, M., & Keller, S. D. (1996). A 12-Item Short-Form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220–233. https://doi.org/10.1097/00005650-199603000-00003

Wichmann, S., Kirschbaum, C., Böhme, C., & Petrowski, K. (2017). Cortisol stress response in post-traumatic stress disorder, panic disorder, and major depressive disorder patients. *Psychoneuroendocrinology, 83*, 135–141. https://doi.org/10.1016/j.psyneuen.2017.06.005

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/browse11/l-m/en

Wu, Y., Tao, Z., Qiao, Y., Chai, Y., Liu, Q., Lu, Q., Zhou, H., Li, S., Mao, J., Jiang, M., & Pu, J. (2022). Prevalence and characteristics of somatic symptom disorder in the elderly in a community-based population: A large-scale cross-sectional study in China. *BMC Psychiatry, 22*(1), Article 257. https://doi.org/10.1186/s12888-022-03907-1

*Yıldırım, Y. E. (2023). Prevalence of prolonged grief disorder and related clinical factors among Turkish psychiatric outpatients during the COVID-19 pandemic. *OMEGA – Journal of Death and Dying*. Advance online publication. https://doi.org/10.1177/00302228231187296

Yin, Q., Shang, Z., Zhou, N., Wu, L., Liu, G., Yu, X., Zhang, H., Xue, H., & Liu, W. (2018). An investigation of physical and mental health consequences among Chinese parents who lost their only child. *BMC Psychiatry, 18*(1), Article 45. https://doi.org/10.1186/s12888-018-1621-2

*Zhang, H., Shang, Z., Wu, L., Sun, Z., Zhang, F., Sun, L., Zhou, Y., Wang, Y., & Liu, W. (2020). Prolonged grief disorder in Chinese Shidu parents who have lost their only child. *European*

Journal of Psychotraumatology, 11(1), Article 1726071. https://doi.org/10.1080/20008198.2020.1726071

*Zhou, N., Wen, J., Stelzer, E.-M., Killikelly, C., Yu, W., Xu, X., Shi, G., Luo, H., Wang, J., & Maercker, A. (2020). Prevalence and associated factors of prolonged grief disorder in Chinese parents bereaved by losing their only child. *Psychiatry Research, 284*, Article 112766. https://doi.org/10.1016/j.psychres.2020.112766

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Dropout From Trauma-Focused Treatment for PTSD in a Naturalistic Setting

Verena Semmlinger[1], Keisuke Takano[2], Larissa Wolkenstein[1], Antje Krüger-Gottschalk[3], Sascha Kuck[3], Anne Dyer[4], Andre Pittig[5], Georg W. Alpers[6], Thomas Ehring[1,7]

[1] *Department of Psychology, LMU Munich, Munich, Germany.* **[2]** *Human Informatics and Interaction Research Institute (HIIRI), National Institute of Advanced Industrial Science and Technology (AIST), Tsukuba, Japan.* **[3]** *Institute of Psychology, WWU Münster, Münster, Germany.* **[4]** *Central Institute of Mental Health, ZISG Mannheim, Mannheim, Germany.* **[5]** *Translational Psychotherapy, Institute of Psychology, Georg-August-University of Göttingen, Göttingen, Germany.* **[6]** *Department of Psychology, School of Social Science, University of Mannheim, Mannheim, Germany.* **[7]** *German Center for Mental Health (DZPG), Munich, Germany.*

Clinical Psychology in Europe, 2025, Vol. 7(1), Article e14491, https://doi.org/10.32872/cpe.14491

Received: 2024-04-26 • **Accepted:** 2024-09-15 • **Published (VoR):** 2025-02-28

Handling Editor: Winfried Rief, Philipps-University of Marburg, Marburg, Germany

Corresponding Author: Verena Semmlinger, Department of Psychology, LMU Munich, 80802 Munich, Germany. Phone +49 89 2180 5171; Fax: +49 89 2180 5224. E-mail: verena.semmlinger@psy.lmu.de

Supplementary Materials: Materials [see Index of Supplementary Materials]

Abstract

Background: Although evidence-based interventions for posttraumatic stress disorder (PTSD) are highly effective, on average about 20% of patients drop out of treatment. Despite considerable research investigating PTSD treatment dropout in randomized controlled trials (RCTs), findings in naturalistic settings remain sparse.

Objective: Therefore, the present study investigated the frequency and predictors of dropout in trauma-focused interventions for PTSD in routine clinical care.

Method: The sample included n = 195 adults with diagnosed PTSD, receiving trauma-focused, cognitive behavioral therapy in routine clinical care in three outpatient centers. We conducted a multiple logistic regression analysis with the following candidate predictors of dropout: patient variables (e.g., basic sociodemographic status and specific clinical variables) as well as therapist's experience level and gender match between therapist and patient.

Results: Results showed a dropout rate of 15.38%. Age (higher dropout probability in younger patients) and living situation (living with parents predicted lower dropout probability compared to

living alone) were significant predictors of dropout. Dropout was not significantly associated with the therapist's experience level and gender match.

Conclusions: In conclusion, routinely assessed baseline patient variables are associated with dropout. Ultimately, this may help to identify patients who need additional attention to keep them in therapy.

Keywords

treatment dropouts, posttraumatic stress disorder, prediction, psychotherapy, clinical practice, naturalistic setting

Highlights

- About 15% of patients receiving PTSD treatment in routine clinical care dropped out.
- This rate is lower than found in previous studies.
- Age and living situation were the only variables related to dropout.

Evidence-based interventions for posttraumatic stress disorder (PTSD) have been shown to be highly effective (e.g., Mavranezouli et al., 2020). However, about 20% of patients receiving an intervention for PTSD drop out of treatment (e.g., Varker et al., 2021). As treatment dropout can lead to lower treatment effectiveness and reduced probability of improvement (Barrett et al., 2008; Varker et al., 2021), PTSD treatment dropout is an important clinical challenge. On a general level, dropout can be defined as termination of an initiated treatment before the symptoms that had caused the patient to seek treatment have been alleviated (Swift et al., 2009; Swift & Greenberg, 2012). Despite repeated efforts to establish a common standard, there remains a lack of consensus in the literature regarding the operationalization of dropout, resulting in different variants being observed (e.g., Barrett et al., 2008; Imel et al., 2013). One criterion that is common in many different operationalization methods is that dropout is a unilateral decision by the patient without mutual agreement or discussion of the decision with the therapist (Swift et al., 2012). In clinical practice, therapist judgement has been discussed for many years as a preferred operationalization method (Swift & Greenberg, 2012; Wierzbicki & Pekarik, 1993) that can be combined with an objective measure to ensure reliability and comparability (Semmlinger & Ehring, 2022).

Previous research has focused on estimating the prevalence of dropout from psychological treatment in randomized controlled trials (RCTs). Across different mental disorders, a large-scale meta-analysis found a weighted average dropout-rate of 19.7%, 95% CI [18.7, 20.7] (Swift & Greenberg, 2012). The average dropout rate reported from evidence-based treatments for PTSD is comparable to this general dropout rate. In a recent meta-analysis investigating dropout from guideline-recommended psychological treatments for PTSD in RCTs, Varker et al. (2021) reported an average dropout rate of 20.9%, 95% CI [17.2, 24.9]. Similar dropout rates have been estimated by other previous

meta-analyses that focus on a wider range of treatment orientations and settings (e.g., Imel et al., 2013: 18.3%, 95% CI [14.8, 21.8]; Lewis et al., 2020: 16%, 95% CI [14, 18]). While there is a vast body of research investigating dropout in RCTs, less is known about dropout rates from treatment for PTSD in routine clinical care. In a systematic review investigating dropout from outpatient treatment for PTSD in a sample of veterans with combat-related PTSD, Goetter et al. (2015) estimated a dropout rate of 36%, 95% CI [26.2, 43.9]. A recent meta-analysis including both RCTs and non-RCTs reported a weighted average dropout rate of 41.5% from trauma-focused CBT for PTSD (Mitchell et al., 2022). It is worth noting that, due to the focus of their analysis, Mitchell et al. (2022) only reported the average dropout rate across all studies and did not include information on the weighted dropout rates for RCTs and non-RCTs separately. Dropout rates for the included non-RCT studies were 35%, 67.5%, and 72.2% (Mitchell et al., 2022).

For dropout from PTSD treatment a number of predictors have been discussed. First, baseline PTSD symptom severity might influence dropout, evidence however is mixed. While Varker et al. (2021) did not find a significant effect, Mitchell et al. (2022) showed higher clinician-rated baseline PTSD symptom severity scores in patients dropping out of treatment compared to completers (Hedge's g = .50, 95% CI [-.95, -.04], p < .05). It is worth noting that this effect applied only to clinician-rated but not to self-rated PTSD severity. Zandberg et al. (2016) added to these findings by examining the influence of the rate of improvement on dropout as a function of symptom severity. The authors showed that for patients with high baseline severity, high dropout rates were associated with both very fast and very slow PTSD improvement, in contrast to patients with low baseline severity, who showed high dropout rates only with fast improvement. The loss of motivation and reduction in the credibility of treatment caused by slow improvement of PTSD symptoms might result in a higher risk of dropout in patients with high PTSD severity (Zandberg et al., 2016).

Second, comorbidity is often discussed as a possible predictor, especially comorbid depression, generalized anxiety disorder (GAD), alcohol disorder, and borderline personality disorder (BPD) (e.g. Steindl et al., 2003). However, the findings are contradictory and potential mechanisms are still unknown (e.g., Angelakis & Nixon, 2015; Mitchell et al., 2022; Snoek et al., 2021). As possible explanations, different studies have discussed depressed patients' reduced ability for emotional processing (Angelakis & Nixon, 2015) or the possible exacerbation of PTSD symptomatology and the increase of psychosocial impairment as a result of comorbid BPD (Frías & Palma, 2015). Specifically, with regard to dropout, a handful of studies have reported an effect of co-occurring depression (e.g., Zayfert et al., 2005), anxiety (e.g., McDonagh et al., 2005; van Minnen et al., 2002), or comorbid personality disorder (e.g., McDonagh et al., 2005) on dropout. However, recent large-scale meta-analyses did not find a significant relationship between comorbidity and dropout from PTSD treatment (Mitchell et al., 2022; Snoek et al., 2021; Varker et al., 2021).

Third, other pretreatment clinical variables might be associated with dropout in PTSD treatments. However, results to date are inconsistent and findings only rely on few studies. Possible predictors are difficulties in emotional regulation (no effect: Belleau et al., 2017; Shnaider et al., 2022; effect: Bremer-Hoeve et al., 2023; Gilmore et al., 2020), anger (no effect: Hinton et al., 2022; van Minnen et al., 2002; mixed results: Rizvi et al., 2009), impaired social functioning (effect: Zayfert et al., 2005), dissociative symptoms (no effect: Hagenaars et al., 2010), and childhood trauma (effect: Miles & Thompson, 2016; mixed results: Resick et al., 2014; no effect: van Minnen et al., 2002). In addition, the patient's trauma response and maladaptive processing (e.g. avoidance, rumination, overgeneralization) may be associated with dropout (Alpert et al., 2020; Shayani et al., 2023). Alpert et al. (2020) found that more negative emotions and ruminative processing predicted lower dropout, whereas overgeneralization was associated to higher dropout. In contrast, Shayani et al. (2023) did not find an effect of overgeneralization, ruminative processing, and negative emotions, but did find that higher levels of avoidance were associated with higher dropout.

Concerning sociodemographic variables, only for the variable age is there a reasonable indication that younger age might be predictive for dropout in PTSD treatment (Garcia et al., 2011; Goetter et al., 2015; Rizvi et al., 2009). However, in two recent meta-analyses, none of the sociodemographic variables (including age) was found to be a consistent predictor across studies (Lewis et al., 2020; Varker et al., 2021).

The majority of studies investigating dropout in PTSD treatment have used an RCT design. Therefore, much less is known about dropout in naturalistic settings. To our knowledge, there is only one review with a veterans sample (Goetter et al., 2015) and few studies (Garcia et al., 2011; van Minnen et al., 2002) specifically investigating dropout in routine clinical care. Transferring results from efficacy studies (RCTs) to naturalistic therapeutic settings might be problematic (Leichsenring, 2004; Schindler et al., 2011). Despite the well-known strength of RCTs it has been discussed whether randomization in RCTs and the strict use of diagnosis specific treatment manuals impose artificial conditions that do not reflect the complexities of clinical practice. Therefore, naturalistic studies are required (Leichsenring, 2004).

The aim of the present study was to investigate the frequency and predictors of dropout in trauma-focused, guideline-recommended interventions for PTSD in routine clinical care. Due to the lack of research on the prevalence and predictors from PTSD treatment in naturalistic settings, our analyses followed an exploratory approach.

Method

Participants

Data was assessed at three university-based outpatient centers providing treatment for PTSD in Germany, located at LMU Munich (Dataset 1) as well as the University of Münster and the Otto Selz Institute at the University of Mannheim (Dataset 2). The sample consisted of 195 adult patients receiving treatment for PTSD. All data was collected as part of effectiveness studies evaluating trauma-focused cognitive behavioral therapy (TF-CBT) for PTSD in routine clinical care (previous, different analysis only on Dataset 2: Krüger-Gottschalk et al., 2024; Schumm et al., 2022, 2023). At pretreatment, all patients met DSM-5 diagnostic criteria for PTSD assessed via the Clinician-Administered PTSD Scale for DSM-5 (CAPS-5) (Weathers, Blake, et al., 2013), and were between 18 and 65 years old. Only participants who had already terminated their treatment at the respective institution and had attended at least one treatment session were included in the study. Exclusion criteria included current psychotic disorder, current substance dependence, or current suicidal intent (First, Williams, Karg, & Spitzer, 2016). Sociodemographic and clinical characteristics of the sample are presented in Table 1.

Treatment

Treatment in all outpatient centers consisted of trauma-focused cognitive behavioral therapy following the same treatment manual. Due to the naturalistic setting of the study, no randomization took place and there was no control condition. The treatment manual is based on empirically tested therapy concepts (especially Ehlers & Clark's cognitive therapy approach, Ehlers & Wild, 2022, as well as DBT-PTSD principles, Bohus et al., 2020) and follows a modularized phase-based approach (see also Ehring, 2019). It includes three consecutive phases. Phase 1 can be summarized as preparation for trauma-focused therapy, including providing a theoretical rationale, increasing treatment motivation, or reducing risky or self-destructive behavior where needed. Phase 2 consisted of the trauma-focused interventions. Therapists could choose between different trauma-focused interventions, including Prolonged Exposure, cognitive therapy, Imagery Rescripting, trigger analyses and discrimination training, as well as cognitive interventions targeting dysfunctional assumptions. Phase 3 was the final phase of treatment and focused on improving quality of life, resuming activities, and relapse prevention. The treatment plan was intended to take each patient through all three phases, with the number of sessions required for each phase and the selection of modules within each phase varying from patient to patient. Depending on the current symptomatology, deviations from this phase structure had to be made in individual cases.

Treatment sessions were usually provided on a weekly basis, with a regular session duration of 50 minutes. The overall average treatment length was $M = 36.6$ sessions ($SD = 23.4$). The average treatment length for dropout cases was $M = 23.3$ sessions

(*SD* = 17.5) and *M* = 39.3 (*SD* = 23.5) for patient who did not drop out. On average patients underwent *M* = 5.0 (*SD* = 1.3) preparatory sessions. This is higher than typically reported in RCTs for PTSD, whereas 12 – 16 sessions are more frequently used. In the German healthcare system patients are permitted to receive up to 80 treatment sessions. Therefore, the reported number of sessions used in our study is typical of the German healthcare system. Second, PTSD treatment in RCTs is often provided in 90-100 min sessions, which means that the treatment dose received in the current study is not that different to typical RCT settings. The treatments were conducted by either licensed CBT therapists (39.2%) or psychotherapists in training (60.8%) employed at the outpatient centers. Supervision by a CBT therapist with expertise in PTSD treatment was regularly provided, on average at every second session. Given the naturalistic nature of the study, it was not feasible to implement formal fidelity checks. The majority of the therapists were female (86.4%).

Measures

The baseline assessment included sociodemographic data, namely age, gender, marital status, living situation, and education. Clinical variables were assessed using clinical interviews and psychometric questionnaires. In addition, two therapist variables, i.e., experience level and gender match, were coded as potential predictors of dropout. For each patient, we revised the patient files, analyzing the therapeutic session protocols.

Dropout

Dropout was operationalized using the therapist's judgement, and the termination had to be initiated by the patient, without a mutual agreement that termination was the best choice. Therapists routinely documented this information in patient files on a treatment termination form. In exceptional cases, where no information was provided, we used an elaborate file analysis, i.e., analyzing the three last session protocols for each respective patient, to retrieve the information needed. If no or only ambiguous information could be obtained, the patient was excluded from the study.

Clinician-Administered PTSD Scale for DSM-5 (CAPS-5)

The CAPS-5 (Weathers, Blake, et al., 2013; German translation by Schnyder, 2013) is a structured diagnostic interview that assesses posttraumatic stress symptoms in the past month. Symptoms are rated on a five-point Likert scale ranging from 0 = absent to 4 = extreme, with a rating of 2 or higher indicating the presence of a symptom (Weathers et al., 2018). The presence of at least one symptom per cluster "intrusive symptoms" and "avoidance", and at least two symptoms per cluster "changes in mood and cognition" and "hyperarousal" indicated the presence of a PTSD diagnosis. The CAPS-5 is a gold-standard clinical interview with good reliability and validity (Weathers et al., 2018).

Structured Clinical Interview for DSM (SCID)

The SCID (First, Williams, Karg, & Spitzer, 2016; Wittchen et al., 1997) was used to assess the presence of comorbid disorders. The SCID for personality disorders (First, Williams, Smith Benjamin, & Spitzer, 2016; Fydrich et al., 1997) was administered to assess the presence of comorbid personality disorders. The SCID is a gold-standard clinical interview to assess diagnostic criteria according to the DSM. For each disorder, interview questions along the DSM criteria allow the rating of diagnostic symptoms as present or absent.

PTSD-Checklist for DSM-5 (PCL-5)

The PCL-5 (Weathers, Litz, et al., 2013; German version by Krüger-Gottschalk et al., 2017) was used to assess posttraumatic symptom severity. The assessment consists of 20 items, corresponding to the DSM-5 PTSD criteria. Distress caused by each symptom is rated on a five-point Likert scale ranging from 0 = not at all to 4 = extremely. Symptom severity was obtained as a sum score of all 20 items (range 0 to 80). The German PCL-5 has demonstrated high internal consistency (Cronbach's $\alpha = .95$) (Krüger-Gottschalk et al., 2017). In the current study, internal consistency was also high ($\alpha = .87$). Please note that Cronbach's alpha for all analyzed questionnaires was calculated on the non-imputed dataset.

Childhood Trauma Questionnaire (CTQ-28)

Exposure to traumatic childhood experiences was assessed with the CTQ-28 (Bernstein et al., 2003; German version by Klinitzke et al., 2012). The CTQ-28 is a self-report questionnaire consisting of 28 items, rated on a five-point Likert scale ranging from 1 = never true to 5 = very often true. A sum score for all items (range 25 to 128) was calculated. The German CTQ-28 shows overall good psychometric properties. The internal consistency for the four subscales without physical neglect is high ($\alpha \geq .80$), while the physical neglect subscale shows weak internal consistency ($\alpha = .55$) (Klinitzke et al., 2012). In the current study, internal consistency was good ($\alpha = .95$) for the total CTQ score.

Inventory of Interpersonal Problems (IIP-32)

The IIP-32 was used to assess interpersonal problems (Horowitz et al., 2000; German version by Thomas et al., 2011). The self-report questionnaire contains 32 items, assessing interpersonal behavior that the participant either finds difficult or shows in excess. The items are rated on a five-point Likert scale ranging from 0 = not at all to 4 = extremely. In the parent studies, different item versions of the questionnaire were used (IIP-127, IIP-64, IIP-32). For the main analyses, we used the IIP-32 version and narrowed the long versions down to the IIP-32. We calculated the IIP-32 total score as the mean of the eight scale scores (Horowitz et al., 2000). The internal consistency of the German IIP-32 was rated as satisfactory to good; for the individual scales it ranged from $\alpha = .60$ to $\alpha = .83$ (Thomas et

al., 2011). In the current study the internal consistency for the total IIP-32 was high (α = .90).

Dissociative Experience Scale (DES)

Dissociative symptoms were assessed with the Dissociative Experience Scale (DES) (Bernstein & Putnam, 1986; German version by Spitzer et al., 2004, called FDS-20). The DES is a 20-item self-report questionnaire. Items are rated on a scale ranging from 0% (never) to 100% (all the time). We used the total mean score to determine the overall dissociation. The DES showed good psychometric measures and the internal consistency was α = .93 (Spitzer et al., 2004). In the current study internal consistency was high α = .93.

Posttraumatic Cognitions Inventory (PTCI) and Interpretation of Symptoms Inventory (IPSI)

Posttraumatic cognitions were assessed using a combined version of the PTCI (Foa et al., 1999) and the IPSI (Dunmore et al., 1999) (German versions by Ehlers & Boos, 2000). The self-report questionnaire assesses negative cognitions and beliefs in response to a traumatic experience (PTCI) and to posttraumatic symptoms (IPSI). The 48 items are rated on a seven-point Likert scale ranging from 1 = totally disagree to 7 = totally agree. We used the total sum score for PTCI and the IPSI mean (Ehlers, 1999). The German PTCI has demonstrated high internal consistency of α = .95 and good overall psychometric properties (Müller et al., 2010). The internal consistency reported for the IPSI was α = .84 (Dunmore et al., 2001). In the current study internal consistency was high for PTCI (α = .92) and IPSI (α = .92).

Difficulties in Emotion Regulation Scale (DERS)

Emotional dysregulation was assessed with the self-report questionnaire DERS (Gratz & Roemer, 2004; German version by Ehring et al., 2008). The 36 items are rated on a five-point Likert scale ranging from 1 = almost never to 5 = almost always. We used the DERS sum score (range 36 to 180) to determine possible difficulties in emotion regulation. The German version of the DERS has an excellent internal consistency of α = .96 (for the sum score) (Kruse et al., 2024). In the current study, internal consistency was excellent, α = .94.

Procedure

The studies were approved by the local ethics committees at the LMU Munich, University of Münster, and the University of Mannheim. All three outpatient centers are specialized in the treatment of patients with trauma-related disorders. Participants referred to these centers were screened for eligibility. If eligible, participants received detailed information about the respective study, and written informed consent was obtained. Due to the

naturalistic setting, participants were not randomized to different conditions but received standard care (see treatment). After the baseline assessment had taken place, the treatment was initiated at the next possible date.

All candidate predictor variables were assessed at baseline. The baseline assessment session consisted of clinical interviews (CAPS-5; SCID) as well as sociodemographic and clinical questionnaires. As treatment was delivered in a naturalistic setting, a substantial effort was made to prevent premature termination of treatment as part of the standard procedure. In the case of excused absence, a new appointment offer was made; in the case of unexcused absence, patients were called by the therapists to make a new appointment. If no contact could be made after several attempts, a letter was sent asking the patient to get in contact within a defined period of time to guarantee continued access to treatment. If the patient clearly expressed the desire to discontinue treatment, no further attempts to contact them were made.

Statistical Analyses

All statistical analyses were conducted using R (Version 4.2.0). Datasets from two parent studies were merged for the current analyses. The dropout rate was calculated as the proportion of the patients who dropped out to the total number of patients who had started the treatment. There was a notable amount of missing data in some questionnaires (M = 7%, SD = 4%, max = 27%). The missing data was assumed to be missing at random (MAR) (Bhaskaran & Smeeth, 2014), and was imputed using the iterative procedure of conditional multiple imputation technique on an item level, i.e., before calculating the respective sum score. Conditional multiple imputation was realized by the five-step procedure proposed by Rubin (1976) and Kropko et al. (2014), using the R *Multivariate Imputation by Chained Equations (mice)* package (van Buuren & Groothuis-Oudshoorn, 2011). The number of multiple imputations as well as the number of iterations were set to five (m = 5, maxit = 5), and we used predictive mean matching (pmm) as the imputation method for continuous variables and the logistic regression (logreg) as the imputation method for dichotomous variables. We conducted a sensitivity analysis to ensure that the results were not affected by multicollinearity due to highly correlated items in the dataset or by the use of the multiple imputed dataset for our main analysis.

First, we tested the differences in demographics and baseline symptom levels between patients who dropped out and those who did not. Next, zero-order associations were examined between dropout and the predictors of interest using point-biserial correlation on the imputed data. We then conducted a multiple logistic regression analysis (maximum likelihood estimation; imputed data) to investigate the unique effects of the variables on dropout after controlling for the effect of the other variables in the model. The level of significance was set as α = .05. We included the following variables as potential predictors of dropout (all assessed at the beginning of treatment): age, gender, marital status, living situation, education, posttraumatic symptom severity (PCL), exposure to

traumatic childhood experiences (CTQ), interpersonal problems (IIP), overall dissociation (DES), posttraumatic cognitions in response to the traumatic experience (PTCI) and to posttraumatic symptoms (IPSI), emotional dysregulation (DERS), number of previous treatments (outpatient and inpatient), number of comorbid disorders (all comorbid disorder), comorbid personality disorder, therapist's experience level (registered vs. in training), and gender match.

Although our primary focus was on the effects of each predictor on dropout, we were interested in how well the logistic regression model would predict dropout. We evaluated the prediction performance using leave-one-out cross-validation on the imputed datasets. The following three performance measures were computed (as medians across imputed datasets): accuracy (i.e., the number of patients who were correctly identified by the model as dropouts or non-dropouts divided by the total number of patients), sensitivity (i.e., the number of dropouts correctly identified as dropouts by the model divided by the number of dropouts), and specificity (i.e., the number of non-dropouts correctly identified as non-dropouts divided by the number of non-dropouts). In addition, Receiver Operating Characteristic (ROC) analysis was performed to evaluate the discriminatory power of the logistic regression model. The area under the ROC curve (AUC) was calculated to summarize the overall performance of the model, again as median AUC across the multiple imputed datasets. The AUC typically ranges from 0 to 1, with 1 indicating the perfect separation and with 0.5 meaning random separation (or poor prediction performance).

Results

Descriptives and Demographics

The sample consisted of 195 patients, with a mean age of 36.14 years (*SD* = 13.02 years). The majority of patients were female (75.9%). Ninety-six patients (56.8%) had at least one comorbid disorder. The mean baseline PTSD symptom severity (PCL) was *M* = 46.2 (*SD* = 14.5), indicating a high severity of PTSD symptoms. Patients in the sample experienced a variation of traumatic events, including accidental trauma, victimization, or trauma predominantly related to death threat. There was a significant association between dropout and age (see Table 1), but not with respect to the other variables studied. The descriptive statistics for all demographic and clinical measures of the sample are presented in Table 1.

Table 1

Descriptive Statistics of the Sample, of Dropouts, and of No Dropout at Baseline

	Total	Dropout	No Dropout	
Variable	***n* (%) / *M* (*SD*)**	***n* (%) / *M* (*SD*)**	***n* (%) / *M* (*SD*)**	***t* or χ^2 (*p*)**
Gender[a]				0.35 (.56)
Female	148 (75.9%)	21 (70.0%)	127 (77.0%)	
Male	47 (24.1%)	9 (30.0%)	38 (23.0%)	
Age in years[b]	36.1 (13.02)	29.97 (10.11)	37.28 (13.21)	**3.40 (.001)**
Marital status[c]				0.73 (.70)
Single	112 (59,6%)	19 (65.5%)	93 (58.5%)	
Married	58 (30.8%)	7 (24.1%)	51 (32.1%)	
Divorced/widowed	18 (9.6%)	3 (10.4%)	15 (9.4%)	
Living situation[b]				3.90 (.27)
Alone	41 (21.9%)	7 (24.1%)	34 (21.4%)	
With partner	106 (56.7%)	14 (48.3%)	92 (57.9%)	
With parents	23 (12.3%)	2 (10.3%)	21 (13.2%)	
Other	17 (9.1%)	5 (17.2%)	12 (7.5%)	
Highest education level[d]				4.15 (.25)
University degree	35 (18.5%)	3 (10.0%)	32 (20.1%)	
High school[†]	35 (18.5%)	9 (30.0%)	26 (16.4%)	
Secondary school[‡]	102 (54.0%)	16 (53.3%)	86 (54.1%)	
Other	17 (9.0%)	2 (6.7%)	15 (9.4%)	
Previous treatment[e]				0.63 (.43)
yes	106 (58.6%)	14 (50.0%)	92 (60.1%)	
no	75 (41.4%)	14 (50.0%)	61 (39.9%)	
Comorbid PD[f]				< .001 (1.0)
yes	15 (8.6%)	2 (6.9%)	13 (8.9%)	
no	160 (91.4%)	27 (93.1%)	133 (91.1%)	
Number of CD[g]	0.98 (1.1)	0.89 (0.91)	0.99 (1.13)	0.54 (.60)
Gender match[h]				0.02 (.89)
Match	107 (73.3%)	19 (70.4%)	88 (73.9%)	
No match	39 (26.7%)	8 (29.6%)	31 (26.1%)	
Approval therapist[i]				0.02 (.89)
Licensed	56 (39.2%)	11 (42.3%)	45 (38.5%)	
Non-licensed	87 (60.8%)	15 (57.7%)	72 (61.5%)	

Variable	Total	Dropout	No Dropout	
	n (%) / *M* (*SD*)	*n* (%) / *M* (*SD*)	*n* (%) / *M* (*SD*)	*t* or χ^2 (*p*)
Clinical measures[a]				
PCL-5	46.2 (14.5)	47.0 (12.2)	46.1 (15.1)	-0.33 (.74)
CTQ-28	55.2 (22.9)	49.6 (15.9)	56.2 (24.2)	1.46 (.15)
IIP-32	1.6 (0.6)	1.6 (0.5)	1.7 (0.7)	0.54 (.59)
DES	2.0 (1.8)	2.2 (1.5)	2.0 (1.9)	-0.55 (.58)
PTCI	131.7 (36.3)	135.3 (33.2)	131.0 (37.6)	-0.59 (.55)
IPSI	3.5 (1.5)	4.0 (1.2)	3.5 (1.5)	-1.77 (.08)
DERS	103.8 (27.4)	103.3 (23.9)	103.9 (28.1)	0.11 (.91)

[a]*n* = 195, [b]*n* = 187, [c]*n* = 188, [d]*n* = 189, [e]*n* = 181, [f]*n* = 175, [g]*n* = 167, [h]*n* = 146, [i]*n* = 143.
[†]High school: 12-13 years of schooling, according to the German school system; [‡]Secondary school: 9-10 years of schooling, according to the German school system; with partner = with partner and/or child(ren) in own apartment; with parents = with parents/one parent; previous treatment = previous psychological treatment (inpatient and/or outpatient); comorbid PD = comorbid personality disorder; number of CD = number of comorbid disorders; *M, SD,* and *t* values for the clinical measures were calculated on the imputed dataset; significant effects are displayed in bold.

Dropout in Trauma Focused-Treatment for PTSD

A total of 30 out of 195 patients (15.38%) were classified as dropouts according to our criteria.

Analysis of Dropout Prediction

Association Between Dropout and Predictor Variables

Point-biserial correlations were calculated on the imputed dataset to examine the zero-order associations between dropout and the predictor variables. Results revealed a significant positive correlation between dropout and age (r = -.19, p = .02) but not between dropout and any other variable. See Supplementary Materials, Table S.1 for a complete correlation matrix of all variables studied.

Prediction of Dropout

To examine the unique influence of the variables of interest on dropout (0 = no dropout, 1 = dropout), a multiple logistic regression was performed on the imputed data. The results indicated that age (β = - 0.07, p = .04) and living situation (β = -2.16, p = .04) were significant predictors of dropout (see Table 2). Results showed that younger individuals were more likely to drop out of treatment, with an *OR* of 0.94. Patients who lived with their parents were at lower risk of dropout compared to those who lived alone (*OR* = 0.12).

Table 2

Results of the Logistic Regression Analysis

Variable	β	*SE*	*t*	*OR*	*LL*	*UL*	*p*
Intercept	-0.46	2.13	-0.22	0.63	0.01	47.31	.82
Gender (Ref. = female)	0.87	0.65	1.34	2.39	0.67	8.59	.18
Age	-0.07	0.03	-2.12	0.94	0.88	1.00	**.04**
Marital status (Ref. = single)							
Married	-0.33	0.67	-0.49	0.72	0.19	2.73	.63
Divorced/widowed	0.35	0.91	0.39	1.42	0.23	8.79	.70
Living situation (Ref. = alone)							
With partner	-0.11	0.68	-0.17	0.89	0.23	3.42	.87
With parents	-2.16	1.02	-2.11	0.12	0.02	0.88	**.04**
Other	0.05	0.83	0.06	1.05	0.20	5.41	.95
Highest education level (Ref. = uni. degree)							
High school	1.12	0.83	1.35	3.07	0.59	15.90	.18
Secondary school	0.45	0.79	0.57	1.57	0.33	7.43	.57
Other	0.99	1.10	0.90	2.68	0.31	23.41	.37
Previous treatment (Ref. = no)	-0.39	0.54	-0.73	0.68	0.23	1.97	.47
Comorbid PD (Ref. = yes)	0.92	0.93	0.99	2.52	0.39	16.30	.33
Number of CD	0.03	0.31	0.08	1.03	0.52	2.02	.93
Gender match (Ref. = match)	-0.20	0.62	-0.33	0.82	0.24	2.80	.74
Approval therapist (Ref. = licensed)	-0.02	0.51	-0.05	0.98	0.36	2.66	.96
Clinical measures							
PCL-5	-0.01	0.02	-0.34	0.99	0.95	1.04	.73
CTQ-28	-0.01	0.01	-0.81	0.99	0.96	1.02	.41
IIP-32	0.02	0.58	0.04	1.02	0.32	3.26	.97
DES	-0.04	0.19	-0.22	0.96	0.66	1.40	.83
PTCI	0.01	0.01	0.66	1.01	0.99	1.03	.51
IPSI	0.47	0.27	1.72	1.60	0.93	2.76	.09
DERS	-0.02	0.02	-1.12	0.98	0.95	1.01	.26

Note. Ref. = reference category; with partner = with partner and/or child(ren) in own apartment; with parents = with parents/one parent; uni. degree = university degree; previous treatment = previous psychological treatment (inpatient and/or outpatient); comorbid PD = comorbid personality disorder; number of CD = number of comorbid disorders; *OR* = Odds ratio; lower and upper CI refer to the corresponding 95% confidence intervals of the *OR*; significant effects are displayed in bold.

Prediction Performance

Using leave-one-out cross-validation on the imputed datasets, we evaluated the prediction performance of the logistic regression model in distinguishing between people who

dropped out vs. those who did not dropout from the treatment. The model showed an accuracy of 80.5%. This accuracy score should be interpreted carefully as the data was not balanced between dropout (15.38%) and no dropout (84.62%). Indeed, the specificity was excellent (95.2%) although the sensitivity was poor (3.3%), meaning that the model is not good at identifying dropouts. ROC analysis showed an AUC value of 0.58, indicating the marginal discriminatory power of the logistic regression model.

Discussion

The first aim of the present study was to investigate the frequency of dropout in trauma-focused, guideline-recommended interventions for PTSD in routine clinical care. 15.38% of patients unilaterally decided to prematurely terminate a started PTSD treatment. The dropout rate found in our study was considerably lower than previous estimates in routine clinical care. This applies for a sample of veterans (e.g., 36%, Goetter et al., 2015), as well as for a joint consideration of trauma-focused treatments for PTSD in RCTs and non-RCTs (e.g., 41.5%, Mitchell et al., 2022). The present findings are further accentuated by the fact that the estimated dropout rate is comparable or even slightly lower than mean dropout rates reported in meta-analyses of highly standardized RCTs, e.g., 16% for a wide range of PTSD treatments (Lewis et al., 2020) and 20.9% from guideline-recommended PTSD treatment (Varker et al., 2021). This finding on the low dropout rate is of particular importance as in clinical practice it is a major therapeutic goal to develop not only effective but also acceptable and feasible treatments. A number of possible explanations for the low dropout rate in our study are conceivable. First treatment was delivered in university-based outpatient centers which provide a well-structured treatment approach along with close supervision, while also allowing for some flexibility in treatment provision. Thus, it could be argued that the present setting combines the strengths of both, RCTs and a naturalistic setting. Note, however, that in RCTs across disorders higher dropout rates were found in university-based institutions (Swift & Greenberg, 2012). Second, therapists in training might invest more time and effort to tailor treatment to their patients' needs than it is usually observed in regular care. Third, the manualized TF-CBT provided as a treatment may have been a particularly suitable form of treatment for the PTSD patients who participated in the current study. Conceivable explanations include the modularized phase-based approach with high flexibility in the selected modules per phase and flexibility in the sessions provided per module and phase. It is further conceivable that the specialization of the outpatient centers in PTSD treatment has an additional effect. Forth, we used well defined criteria to operationalize dropout (therapist decision combined with patient-initiated dropout).

The second aim of the study was to investigate predictors of dropout in trauma-focused, guideline-recommended interventions for PTSD in routine clinical care. A multiple logistic regression revealed age and living situation to be significant predictors, with

higher risk of dropout in younger individuals and lower risk of dropout in patients who lived with their parents as opposed to living alone. The finding of younger age being predictive for dropout adds to previous findings on predictors of dropout in the general and PTSD-specific literature (Goetter et al., 2015; Swift & Greenberg, 2012), with only few studies not replicating these findings (e.g., Varker et al., 2021). Note, that all patients in the study were adults (between 18 and 65 years). Possible explanations include the fact that young patients may have more competing time demands (Goetter et al., 2015), treatment may not sufficiently match their needs, or young patients may face a lack of stability in their living environments (de Soet et al., 2024). In addition, it is conceivable that young adults have not yet experienced that PTSD symptoms in most cases do not simply disappear on their own over time (Morina et al., 2014).

To our knowledge, no previous study has investigated the influence of living situation on premature termination of treatment. Note that although patients living with their parents probably tend to be younger, the significant findings on lower risk of dropout in patients who lived with their parents compared to living alone had a unique effect, i.e., when controlling for the influence of age. To explain our findings, it appears important to address the influence of parental support on treatment outcomes. In their review of dropout in adolescents, de Soet et al. (2024) showed that parental approval, participation, and support were associated with lower risk of dropout. Therefore, young patients living with their parents might perceive more parental support and thus dropout becomes less likely than if these patients were living alone. However, more research is needed to understand the influence of living situation on premature termination of treatment.

We also examined the possible role of several clinical variables as predictors of dropout. Results showed that baseline symptom levels and associated clinical variables were overall not predictive of dropout. This is in line with earlier findings (mostly based on data collected using RCT designs) showing that e.g., symptom severity (Varker et al., 2021) or comorbidity (Mitchell et al., 2022; Snoek et al., 2021; Varker et al., 2021) were not predictive of dropout. A notable exception is a study by Mitchell et al. (2022), which did find higher PTSD symptom severity at baseline predicted dropout; however, this was only the case for clinician-rated PTSD severity and not for self-rated PTSD scores. Thus, the role of baseline PTSD symptom severity on dropout needs to be examined in further research focused on a possible role of methodological variables.

With regard to the impact of therapist characteristics on dropout our findings indicate that neither the experience level nor the gender match of the therapist has a significant influence on the dropout rate. This contradicts previous findings on treatment dropout across disorders. There is substantial evidence for the so-called therapist effect, which states that differences between therapist influence dropout rates (Deisenhofer et al., 2024; Saxon et al., 2017; Zimmermann et al., 2017). In addition, research has indicated an effect of therapist experience level on dropout (Roos & Werbart, 2013; Swift & Greenberg, 2012). For PTSD treatment in particular, evidence is sparse, with initial

evidence for a therapist effect on dropout (Sayer et al., 2022). In the present study possible influences of therapist characteristics might have been minimized by the fact that patients were treated in a highly specialized service with close supervision, and the fact that most therapist were at an early-career stage. Therefore, the variability of therapist characteristics may have been rather low in the current study. In line with this reasoning, Deisenhofer et al. (2024) found that the therapist effect on dropout was significantly reduced by such institution effects.

Although it was not the primary focus of the current study, we additionally tested how well the logistic regression model would predict dropout. Taking the given imbalance between dropout and no dropout into account, the model comprising different pretreatment variables was not successful in predicting whether a patient who just started treatment would dropout during the course of treatment. Our results are in line with Vöhringer et al. (2020) who reported poor results on the discriminative power of pretreatment variables to distinguish between dropouts and completers. However, Bremer-Hoeve et al. (2023) were able to predict dropout in PTSD treatment using machine learning techniques.

In sum, only very few variables assessed in the current study were significant predictors of dropout, and the overall model could not predict dropout to a practically useful level. This is broadly in line with the majority of earlier findings. Thus, therapists and researchers should be cautious about making confident predictions about retention based on baseline data.

Limitations

This study has a number of important strengths. One major strength is the naturalistic setting of the study, which allows for flexibility and variance in the trauma-focused, guideline-recommended treatment provided. In addition, the naturalistic setting contributes significantly to an increase in external validity and generalizability of the results to clinical practice. Nevertheless, there are a number of noteworthy limitations. First, the number of participants included in the analysis was limited, potentially leading to reduced statistical power. Even though we combined data from three outpatient centers, we had to exclude a substantial number of participants. This was due to the strict inclusion criteria regarding PTSD diagnosis and missing data for the assessment of dropout despite extensive file analysis. Second, treatments were not standardized but allowed for some flexibility based on a manual delineating key treatment principle. On the one hand, this can be regarded as a strength of the study as it is typical for routine clinical practice, where manuals are usually less strictly applied than in RCT research. On the other hand, however, we cannot rule out the possibility that the variability in the composition and timing of the use of different treatment modules may have obscured effects of certain variables in predicting dropout, as therapists may have counter-acted these variables in treatment. Third, results could be limited by the method used to

operationalize dropout. Forth, the uncontrolled study design allows a more naturalistic investigation of dropout. However, in contrast to an RCT design the internal validity of effects of different variables on dropout is low. Specifically, it remains unclear whether confounding variables that were not controlled may have influenced on the occurrence of dropout. Last, although we examined a wide range of variables, potentially important aspects are missing in our dataset. These include type of trauma experienced, treatment characteristics (e.g., session frequency), and patterns of change during treatment (e.g., rate of improvement).

Conclusion and Future Directions

In conclusion, this study provides important knowledge about the dropout rate and predictors of dropout in trauma-focused, guideline-recommended interventions for PTSD in routine clinical care. Results show that the dropout rate in this naturalistic study was comparable to dropout rates found in RCTs. In addition, two baseline predictors of dropout were identified, suggesting that young adults with PTSD may need close, supportive care, especially when they are no longer living with their parents. Therapists can act as supportive guides, build and strengthen hope (Swift & Greenberg, 2012), and be aware of urgent crises and the social needs of their young patients.

Possibly most importantly, however, our findings replicate earlier results showing that identifying patients at risk of dropping out of treatment early-on by baseline variables is challenging and currently not possible at a practically useful level. A number of implications can be drawn from this finding. First, from an applied perspective, these findings contradict widespread clinical beliefs about trauma-focused interventions being less acceptable to patients with high symptom severities, high comorbidity, or complex symptom presentations (e.g., emotion dysregulation, dissociation, interpersonal difficulties). Neither earlier research nor our current findings suggest that patients with these particularly severe and/or complex presentations are more likely to drop out of treatment. However, larger samples may provide more power and enable us to examine even a broader scope of potential predictor variables with modern machine learning approaches (see Taubitz et al., 2022). Second, the cumulated findings may suggest that it is necessary to look beyond pretreatment factors when predicting dropout and to additionally include variables investigating processes occurring in the course of treatment. For example, Zandberg et al. (2016) found that the rates of symptom change had a significant influence on dropout in patients with comorbid PTSD and alcohol dependence. Patients with low baseline symptom severity showed low risk for dropout in slow improvement and higher risk in fast improvement. When baseline symptom severity was high, the effect was u-shaped, with high risk of dropout in both slow and fast improvement (Zandberg et al., 2016). Third, further research should focus on investigating additional variables characterizing the treatment process, in particular the frequency of sessions provided. In a large-scale meta-analysis Hoppen et al. (2023) showed lower dropout

rates for trauma-focused treatments delivered in high intensity. These findings are in line with Levinson et al.'s (2022) meta-analytical findings on dropout from PE provided in an outpatient setting. Finally, as earlier evidence has been inconsistent, we followed an exploratory research approach. Therefore, further studies are needed to test specific hypotheses based on theory. In addition, it appears recommendable to systematically assess subjective reasons from the patients' perspective (Vöhringer et al., 2020).

Expanding research into dropout from PTSD treatment in these ways appears highly relevant since dropout continues to be an important clinical challenge preventing a considerable subgroup of treatment-seeking PTSD sufferers from receiving effective treatment. A better understanding of predictors of – and ultimately causal factors involved in – dropout may ultimately help to develop preventive strategies to reduce dropout and keep patients with severe symptoms in effective treatment.

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

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: The authors report there are no competing interests to declare.

Ethics Statement: The studies were approved by the local ethics committees at the LMU Munich, University of Münster, and the University of Mannheim. Written informed consent was obtained for all participants.

Reporting Guidelines: We report how we determined our sample size, all data exclusions (if any), all manipulations, and all measures in the study, and we follow JARS.

Related Versions: This publication forms part of the doctoral thesis of Verena Semmlinger at the LMU Munich. Semmlinger, V. (2024). *The complexity of treatment failure – Prevalence and predictors of dropout and non-response in psychological treatment for traumatized populations* [Doctoral dissertation, LMU Munich]. Electronic Theses Repository of LMU Munich. https://doi.org/10.5282/edoc.34387

Preregistration: This study's design and its analysis were not pre-registered.

Social Media Accounts: @thomasehring.bsky.social

Data Availability: The authors have no permission to share the data. The code is available upon reasonable request.

Supplementary Materials

The Supplementary Materials contain a full correlation matrix of all variables studied (see Semmlinger et al., 2025S).

Index of Supplementary Materials

Semmlinger, V., Takano, K., Wolkenstein, L., Krüger-Gottschalk, A., Kuck, S., Dyer, A., Pittig, A., Alpers, G. W., & Ehring, T. (2025S). *Supplementary materials to "Dropout from trauma-focused treatment for PTSD in a naturalistic setting"* [Full correlation matrix of all variables studied]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.15955

References

Alpert, E., Hayes, A. M., Barnes, J. B., & Sloan, D. M. (2020). Predictors of dropout in cognitive processing therapy for PTSD: An examination of trauma narrative content. *Behavior Therapy, 51*(5), 774–788. https://doi.org/10.1016/j.beth.2019.11.003

Angelakis, S., & Nixon, R. D. V. (2015). The comorbidity of PTSD and MDD: Implications for clinical practice and future research. *Behaviour Change, 32*(1), 1–25. https://doi.org/10.1017/bec.2014.26

Barrett, M. S., Chua, W.-J., Crits-Christoph, P., Gibbons, M. B., Casiano, D., & Thompson, D. (2008). Early withdrawal from mental health treatment: Implications for psychotherapy practice. *Psychotherapy, 45*(2), 247–267. https://doi.org/10.1037/0033-3204.45.2.247

Belleau, E. L., Chin, E. G., Wanklyn, S. G., Zambrano-Vazquez, L., Schumacher, J. A., & Coffey, S. F. (2017). Pre-treatment predictors of dropout from prolonged exposure therapy in patients with chronic posttraumatic stress disorder and comorbid substance use disorders. *Behaviour Research and Therapy, 91*, 43–50. https://doi.org/10.1016/j.brat.2017.01.011

Bernstein, D. P., Stein, J. A., Newcomb, M. D., Walker, E., Pogge, D., Ahluvalia, T., Stokes, J., Handelsman, L., Medrano, M., Desmond, D., & Zule, W. (2003). Development and validation of a brief screening version of the Childhood Trauma Questionnaire. *Child Abuse & Neglect, 27*(2), 169–190. https://doi.org/10.1016/S0145-2134(02)00541-0

Bernstein, E. M., & Putnam, F. W. (1986). Development, reliability, and validity of a dissociation scale. *The Journal of Nervous and Mental Disease, 174*(12), 727–735. https://doi.org/10.1097/00005053-198612000-00004

Bhaskaran, K., & Smeeth, L. (2014). What is the difference between missing completely at random and missing at random? *International Journal of Epidemiology, 43*(4), 1336–1339. https://doi.org/10.1093/ije/dyu080

Bohus, M., Kleindienst, N., Hahn, C., Müller-Engelmann, M., Ludäscher, P., Steil, R., Fydrich, T., Kuehner, C., Resick, P. A., Stiglmayr, C., Schmahl, C., & Priebe, K. (2020). Dialectical behavior therapy for posttraumatic stress disorder (DBT-PTSD) compared with cognitive processing therapy (CPT) in complex presentations of PTSD in women survivors of childhood abuse: A randomized clinical trial. *JAMA Psychiatry, 77*(12), 1235–1245. https://doi.org/10.1001/jamapsychiatry.2020.2148

Bremer-Hoeve, S., van Vliet, N. I., van Bronswijk, S. C., Huntjens, R. J. C., de Jongh, A., & van Dijk, M. K. (2023). Predictors of treatment dropout in patients with posttraumatic stress disorder due

to childhood abuse. *Frontiers in Psychiatry, 14*, Article 1194669. https://doi.org/10.3389/fpsyt.2023.1194669

Deisenhofer, A.-K., Hehlmann, M. I., Rubel, J. A., Lutz, W., Schwartz, B., Bräscher, A.-K., Christiansen, H., Fehm, L., Glombiewski, J. A., Heider, J., Helbig-Lang, S., Hermann, A., Hoyer, J., In-Albon, T., Lincoln, T., Margraf, J., Risch, A. K., Schöttke, H., Schulze, L., . . . Odyniec, P. (2024). Love yourself as a therapist, doubt yourself as an institution? Therapist and institution effects on outcome, treatment length, and dropout. *Psychotherapy Research*. Advance online publication. https://doi.org/10.1080/10503307.2024.2352749

de Soet, R., Vermeiren, R. R. J. M., Bansema, C. H., van Ewijk, H., Nijland, L., & Nooteboom, L. A. (2024). Drop-out and ineffective treatment in youth with severe and enduring mental health problems: A systematic review. *European Child & Adolescent Psychiatry, 33*(10), 3305–3319. https://doi.org/10.1007/s00787-023-02182-z

Dunmore, E., Clark, D. M., & Ehlers, A. (1999). Cognitive factors involved in the onset and maintenance of posttraumatic stress disorder (PTSD) after physical or sexual assault. *Behaviour Research and Therapy, 37*(9), 809–829. https://doi.org/10.1016/S0005-7967(98)00181-8

Dunmore, E., Clark, D. M., & Ehlers, A. (2001). A prospective investigation of the role of cognitive factors in persistent posttraumatic stress disorder (PTSD) after physical or sexual assault. *Behaviour Research and Therapy, 39*(9), 1063–1084. https://doi.org/10.1016/S0005-7967(00)00088-7

Ehlers, A. (1999). *Posttraumatische Belastungsstörung* [Posttraumatic stress disorder]. Hogrefe.

Ehlers, A., & Boos, A. (2000). Fragebogen zu Gedanken nach traumatischen Erlebnissen (PTCI) [Posttraumatic Cognition Inventory]. In A. Ehlers (Ed.), *Posttraumatische Belastungsstörungen* [Posttraumatic stress disorder] (pp. 92-94). Hogrefe.

Ehlers, A., & Wild, J. (2022). Cognitive therapy for PTSD: Updating memories and meanings of trauma. In U. Schnyder & M. Cloitre (Eds.), *Evidence based treatments for trauma-related psychological disorders* (pp. 181-210). Springer, Cham. https://doi.org/10.1007/978-3-030-97802-0_9

Ehring, T. (2019). Kognitive Verhaltenstherapie [Cognitive behavioural therapy]. In A. Maercker (Ed.), *Traumafolgestörungen* [Trauma-related disorders]. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-58470-5_13

Ehring, T., Fischer, S., Schnülle, J., Bösterling, A., & Tuschen-Caffier, B. (2008). Characteristics of emotion regulation in recovered depressed versus never depressed individuals. *Personality and Individual Differences, 44*(7), 1574–1584. https://doi.org/10.1016/j.paid.2008.01.013

First, M. B., Williams, J. B. W., Karg, R. S., & Spitzer, R. L. (2016). *Structured Clinical Interview for DSM-5 Disorders – Clinician Version*. Hogrefe.

First, M. B., Williams, J. B. W., Smith Benjamin, L., & Spitzer, R. L. (2016). *Structured Clinical Interview for DSM-5 – Personality Disorders*. Hogrefe.

Foa, E. B., Ehlers, A., Clark, D. M., Tolin, D. F., & Orsillo, S. M. (1999). The Posttraumatic Cognitions Inventory (PTCI): Development and validation. *Psychological Assessment, 11*(3), 303–314. https://doi.org/10.1037/1040-3590.11.3.303

Frías, Á., & Palma, C. (2015). Comorbidity between post-traumatic stress disorder and borderline personality disorder: A review. *Psychopathology, 48*(1), 1–10. https://doi.org/10.1159/000363145

Fydrich, T., Renneberg, B., Schmitz, B., & Wittchen, H.-U. (1997). *SKID II. Strukturiertes Klinisches Interview für DSM-IV, Achse II: Persönlichkeitsstörungen. Eine deutschsprachige, erweiterte Bearbeitung der amerikanischen Originalversion des SKID-II von: M.B. First, R.L. Spitzer, M. Gibbon, J.B.W. Williams, L. Benjamin* [SKID II. Structured Clinical Interview for DSM-IV, Axis II: Personality Disorders. A German-language, edited version of the original American version of the SKID-II by: M.B. First, R.L. Spitzer, M. Gibbon, J.B.W. Williams, L. Benjamin]. Hogrefe.

Garcia, H. A., Kelley, L. P., Rentz, T. O., & Lee, S. (2011). Pretreatment predictors of dropout from cognitive behavioral therapy for PTSD in Iraq and Afghanistan war veterans. *Psychological Services, 8*(1), 1–11. https://doi.org/10.1037/a0022705

Gilmore, A. K., Lopez, C., Muzzy, W., Brown, W. J., Grubaugh, A., Oesterle, D. W., & Acierno, R. (2020). Emotion dysregulation predicts dropout from prolonged exposure treatment among women veterans with military sexual trauma-related posttraumatic stress disorder. *Women's Health Issues, 30*(6), 462–469. https://doi.org/10.1016/j.whi.2020.07.004

Goetter, E. M., Bui, E., Ojserkis, R. A., Zakarian, R. J., Brendel, R. W., & Simon, N. M. (2015). A systematic review of dropout from psychotherapy for posttraumatic stress disorder among Iraq and Afghanistan combat veterans. *Journal of Traumatic Stress, 28*(5), 401–409. https://doi.org/10.1002/jts.22038

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41–54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Hagenaars, M. A., van Minnen, A., & Hoogduin, K. A. L. (2010). The impact of dissociation and depression on the efficacy of prolonged exposure treatment for PTSD. *Behaviour Research and Therapy, 48*(1), 19–27. https://doi.org/10.1016/j.brat.2009.09.001

Hinton, E., Steel, Z., Hilbrink, D., & Berle, D. (2022). Anger and predictors of drop-out from PTSD treatment of veterans and first responders. *Behavioural and Cognitive Psychotherapy, 50*(2), 237–251. https://doi.org/10.1017/S1352465821000382

Hoppen, T. H., Jehn, M., Holling, H., Mutz, J., Kip, A., Morina, N., Bisson, J. I., & Olff, M. (2023). The efficacy and acceptability of psychological interventions for adult PTSD: A network and pairwise meta-analysis of randomized controlled trials. *Journal of Consulting and Clinical Psychology, 91*(8), 445–461. https://doi.org/10.1037/ccp0000809

Horowitz, L., Alden, L., Wiggins, J., & Pincus, A. (2000). *Inventory of interpersonal problems* (IIP-32/IIP-64). London, United Kingdom: Psychological Corporation.

Imel, Z. E., Laska, K., Jakupcak, M., & Simpson, T. L. (2013). Meta-analysis of dropout in treatments for posttraumatic stress disorder. *Journal of Consulting and Clinical Psychology, 81*(3), 394–404. https://doi.org/10.1037/a0031474

Klinitzke, G., Romppel, M., Häuser, W., Brähler, E., & Glaesmer, H. (2012). The German Version of the Childhood Trauma Questionnaire (CTQ): Psychometric characteristics in a representative

sample of the general population. *Psychotherapie, Psychosomatik, Medizinische Psychologie, 62*(2), 47–51. https://doi.org/10.1055/s-0031-1295495

Kropko, J., Goodrich, B., Gelman, A., & Hill, J. (2014). Multiple imputation for continuous and categorical data: Comparing joint multivariate normal and conditional approaches. *Political Analysis, 22*(4), 497–519. https://doi.org/10.1093/pan/mpu007

Krüger-Gottschalk, A., Alpers, G., Pittig, A., Kuck, S., Dyer, A., & Ehring, T. (2024). *Effectiveness of phase-based trauma-focused cognitive behavioural therapy in a routine clinical setting* [Unpublished manuscript]. Institute of Psychology, University of Muenster.

Krüger-Gottschalk, A., Knaevelsrud, C., Rau, H., Dyer, A., Schäfer, I., Schellong, J., & Ehring, T. (2017). The German version of the Posttraumatic Stress Disorder Checklist for DSM-5 (PCL-5): Psychometric properties and diagnostic utility. *BMC Psychiatry, 17*(1), Article 379. https://doi.org/10.1186/s12888-017-1541-6

Kruse, J., Svaldi, J., Tuschen-Caffier, B., Berking, M., & Ehring, T. (2024). *Development and validation of the Difficulties in Emotion Regulation Questionnaire – German version.* [Unpublished manuscript]. University of Muenster.

Leichsenring, F. (2004). Randomized controlled versus naturalistic studies: A new research agenda. *Bulletin of the Menninger Clinic, 68*(2), 137–151. https://doi.org/10.1521/bumc.68.2.137.35952

Levinson, D. B., Halverson, T. F., Wilson, S. M., & Fu, R. (2022). Less dropout from prolonged exposure sessions prescribed at least twice weekly: A meta-analysis and systematic review of randomized controlled trials. *Journal of Traumatic Stress, 35*(4), 1047–1059. https://doi.org/10.1002/jts.22822

Lewis, C., Roberts, N. P., Gibson, S., & Bisson, J. I. (2020). Dropout from psychological therapies for post-traumatic stress disorder (PTSD) in adults: Systematic review and meta-analysis. *European Journal of Psychotraumatology, 11*(1), Article 1709709. https://doi.org/10.1080/20008198.2019.1709709

Mavranezouli, I., Megnin-Viggars, O., Daly, C., Dias, S., Welton, N. J., Stockton, S., Bhutani, G., Grey, N., Leach, J., Greenberg, N., Katona, C., El-Leithy, S., & Pilling, S. (2020). Psychological treatments for post-traumatic stress disorder in adults: A network meta-analysis. *Psychological Medicine, 50*(4), 542–555. https://doi.org/10.1017/S0033291720000070

McDonagh, A., Friedman, M., McHugo, G., Ford, J., Sengupta, A., Mueser, K., Demment, C. C., Fournier, D., Schnurr, P. P., & Descamps, M. (2005). Randomized trial of cognitive-behavioral therapy for chronic posttraumatic stress disorder in adult female survivors of childhood sexual abuse. *Journal of Consulting and Clinical Psychology, 73*(3), 515–524. https://doi.org/10.1037/0022-006X.73.3.515

Miles, S. R., & Thompson, K. E. (2016). Childhood trauma and posttraumatic stress disorder in a real-world Veterans Affairs clinic: Examining treatment preferences and dropout. *Psychological Trauma: Theory, Research, Practice, and Policy, 8*(4), 464–467. https://doi.org/10.1037/tra0000132

Mitchell, S., Mitchell, R., Shannon, C., Dorahy, M., & Hanna, D. (2022). Effects of baseline psychological symptom severity on dropout from trauma-focused cognitive behavior therapy

for posttraumatic stress disorder: A meta-analysis. *Traumatology, 29*(2), 112–124. https://doi.org/10.1037/trm0000404

Morina, N., Wicherts, J. M., Lobbrecht, J., & Priebe, S. (2014). Remission from post-traumatic stress disorder in adults: A systematic review and meta-analysis of long term outcome studies. *Clinical Psychology Review, 34*(3), 249–255. https://doi.org/10.1016/j.cpr.2014.03.002

Müller, J., Wessa, M., Rabe, S., Dörfel, D., Knaevelsrud, C., Flor, H., Maercker, A., & Karl, A. (2010). Psychometric properties of the Posttraumatic Cognitions Inventory (PTCI) in a German sample of individuals with a history of trauma. *Psychological Trauma: Theory, Research, Practice, and Policy, 2*(2), 116–125. https://doi.org/10.1037/a0018603

Resick, P. A., Suvak, M. K., & Wells, S. Y. (2014). The impact of childhood abuse among women with assault-related PTSD receiving short-term cognitive-behavioral therapy. *Journal of Traumatic Stress, 27*(5), 558–567. https://doi.org/10.1002/jts.21951

Rizvi, S. L., Vogt, D. S., & Resick, P. A. (2009). Cognitive and affective predictors of treatment outcome in cognitive processing therapy and prolonged exposure for posttraumatic stress disorder. *Behaviour Research and Therapy, 47*(9), 737–743. https://doi.org/10.1016/j.brat.2009.06.003

Roos, J., & Werbart, A. (2013). Therapist and relationship factors influencing dropout from individual psychotherapy: A literature review. *Psychotherapy Research, 23*(4), 394–418. https://doi.org/10.1080/10503307.2013.775528

Rubin, D. B. (1976). Inference and missing data. *Biometrika, 63*(3), 581–592. https://doi.org/10.1093/biomet/63.3.581

Saxon, D., Barkham, M., Foster, A., & Parry, G. (2017). The contribution of therapist effects to patient dropout and deterioration in the psychological therapies. *Clinical Psychology & Psychotherapy, 24*(3), 575–588. https://doi.org/10.1002/cpp.2028

Sayer, N. A., Wiltsey-Stirman, S., Rosen, C. S., Bernardy, N. C., Spoont, M. R., Kehle-Forbes, S. M., Eftekhari, A., Chard, K. M., & Nelson, D. B. (2022). Investigation of therapist effects on patient engagement in evidence-based psychotherapies for posttraumatic stress disorder in the Veterans Health Administration. *Journal of Traumatic Stress, 35*(1), 66–77. https://doi.org/10.1002/jts.22679

Schindler, A. C., Hiller, W., & Witthöft, M. (2011). Benchmarking of cognitive-behavioral therapy for depression in efficacy and effectiveness studies–How do exclusion criteria affect treatment outcome? *Psychotherapy Research, 21*(6), 644–657. https://doi.org/10.1080/10503307.2011.602750

Schnyder, U. (2013). *German version of the Clinician-Administered PTSD Scale for DSM-5 (CAPS-5)* [Unpublished manuscript]. University Hospital Zurich.

Schumm, H., Krüger-Gottschalk, A., Dyer, A., Pittig, A., Cludius, B., Takano, K., Alpers, G. W., & Ehring, T. (2022). Mechanisms of change in trauma-focused treatment for PTSD: The role of rumination. *Behaviour Research and Therapy, 148*, Article 104009. https://doi.org/10.1016/j.brat.2021.104009

Schumm, H., Krüger-Gottschalk, A., Ehring, T., Dyer, A., Pittig, A., Takano, K., Alpers, G. W., & Cludius, B. (2023). Do changes in dysfunctional posttraumatic cognitions differentially predict

PTSD symptom clusters? *Journal of Consulting and Clinical Psychology, 91*(7), 438–444. https://doi.org/10.1037/ccp0000817

Semmlinger, V., & Ehring, T. (2022). Predicting and preventing dropout in research, assessment and treatment with refugees. *Clinical Psychology & Psychotherapy, 29*(3), 767–782. https://doi.org/10.1002/cpp.2672

Shayani, D. R., Canale, C. A., Sloan, D. M., & Hayes, A. M. (2023). Predictors of dropout in cognitive processing therapy for PTSD: An examination of in-session treatment processes. *Behaviour Research and Therapy, 171*, Article 104428. https://doi.org/10.1016/j.brat.2023.104428

Shnaider, P., Boyd, J. E., Cameron, D. H., & McCabe, R. E. (2022). The relationship between emotion regulation difficulties and PTSD outcomes during group cognitive processing therapy for PTSD. *Psychological Services, 19*(4), 751–759. https://doi.org/10.1037/ser0000546

Snoek, A., Nederstigt, J., Ciharova, M., Sijbrandij, M., Lok, A., Cuijpers, P., & Thomaes, K. (2021). Impact of comorbid personality disorders on psychotherapy for post-traumatic stress disorder: Systematic review and meta-analysis. *European Journal of Psychotraumatology, 12*(1), Article 1929753. https://doi.org/10.1080/20008198.2021.1929753

Spitzer, C., Mestel, R., Klingelhöfer, J., Gänsicke, M., & Freyberger, H. J. (2004). Screening and measurement of change of dissociative psychopathology: Psychometric properties. *PPmP – Psychotherapie · Psychosomatik · Medizinische Psychologie, 54*(3–4), 165–172. https://doi.org/10.1055/s-2003-814783

Steindl, S. R., Young, R. M., Creamer, M., & Crompton, D. (2003). Hazardous alcohol use and treatment outcome in male combat veterans with posttraumatic stress disorder. *Journal of Traumatic Stress, 16*(1), 27–34. https://doi.org/10.1023/A:1022055110238

Swift, J. K., Callahan, J., & Levine, J. C. (2009). Using clinically significant change to identify premature termination. *Psychotherapy, 46*(3), 328–335. https://doi.org/10.1037/a0017003

Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology, 80*(4), 547–559. https://doi.org/10.1037/a0028226

Swift, J. K., Greenberg, R., Whipple, J., & Kominiak, N. (2012). Practice recommendations for reducing premature termination in therapy. *Professional Psychology, Research and Practice, 43*(4), 379–387. https://doi.org/10.1037/a0028291

Taubitz, F.-S., Büdenbender, B., & Alpers, G. W. (2022). What the future holds: Machine learning to predict success in psychotherapy. *Behaviour Research and Therapy, 156*, Article 104116. https://doi.org/10.1016/j.brat.2022.104116

Thomas, A., Brähler, E., & Strauß, B. (2011). IIP-32: Entwicklung, Validierung und Normierung einer Kurzform des Inventars zur Erfassung interpersonaler Probleme [IIP-32: Development, validation and standardization of the short form of the inventory of interpersonal problems]. *Diagnostica, 57*(2), 68–83. https://doi.org/10.1026/0012-1924/a000034

van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software, 45*(3, SE-Articles), 1–67. https://doi.org/10.18637/jss.v045.i03

van Minnen, A., Arntz, A., & Keijsers, G. P. J. (2002). Prolonged exposure in patients with chronic PTSD: Predictors of treatment outcome and dropout. *Behaviour Research and Therapy, 40*(4), 439–457. https://doi.org/10.1016/S0005-7967(01)00024-9

Varker, T., Jones, K. A., Arjmand, H.-A., Hinton, M., Hiles, S. A., Freijah, I., Forbes, D., Kartal, D., Phelps, A., Bryant, R. A., McFarlane, A., Hopwood, M., & O'Donnell, M. (2021). Dropout from guideline-recommended psychological treatments for posttraumatic stress disorder: A systematic review and meta-analysis. *Journal of Affective Disorders Reports, 4*, Article 100093. https://doi.org/10.1016/j.jadr.2021.100093

Vöhringer, M., Knaevelsrud, C., Wagner, B., Slotta, M., Schmidt, A., Stammel, N., & Böttche, M. (2020). Should I stay or must I go? Predictors of dropout in an internet-based psychotherapy programme for posttraumatic stress disorder in Arabic. *European Journal of Psychotraumatology, 11*(1), Article 1706297. https://doi.org/10.1080/20008198.2019.1706297

Weathers, F. W., Blake, D., Schnurr, P., Kaloupek, D., Marx, B., & Keane, T. (2013). *The clinician-administered PTSD scale for DSM-5 (CAPS-5).* https://www.ptsd.va.gov/professional/assessment/adult-int/caps.asp

Weathers, F. W., Bovin, M. J., Lee, D. J., Sloan, D. M., Schnurr, P. P., Kaloupek, D. G., Keane, T. M., & Marx, B. P. (2018). The Clinician-Administered PTSD Scale for DSM-5 (CAPS-5): Development and initial psychometric evaluation in military veterans. *Psychological Assessment, 30*(3), 383–395. https://doi.org/10.1037/pas0000486

Weathers, F. W., Litz, B. T., Keane, T. M., Palmieri, P. A., Marx, B. P., & Schnurr, P. P. (2013). *The PTSD checklist for DSM-5 (PCL-5).* Scale available from the National Center for PTSD at www.ptsd.va.gov.

Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout. *Professional Psychology, Research and Practice, 24*(2), 190–195. https://doi.org/10.1037/0735-7028.24.2.190

Wittchen, H.-U., Zaudig, M., & Fydrich, T. (1997). *SKID. Strukturiertes Klinisches Interview für DSM-IV. Achse I und II. Handanweisung* [SKID. Structured Clinical Interview for DSM-IV. Axis I and II. Manual instructions]. Hogrefe.

Zandberg, L. J., Rosenfield, D., Alpert, E., McLean, C. P., & Foa, E. B. (2016). Predictors of dropout in concurrent treatment of posttraumatic stress disorder and alcohol dependence: Rate of improvement matters. *Behaviour Research and Therapy, 80*, 1–9. https://doi.org/10.1016/j.brat.2016.02.005

Zayfert, C., DeViva, J. C., Becker, C. B., Pike, J. L., Gillock, K. L., & Hayes, S. A. (2005). Exposure utilization and completion of cognitive behavioral therapy for PTSD in a "real world" clinical practice. *Journal of Traumatic Stress, 18*(6), 637–645. https://doi.org/10.1002/jts.20072

Zimmermann, D., Rubel, J., Page, A. C., & Lutz, W. (2017). Therapist effects on and predictors of non-consensual dropout in psychotherapy. *Clinical Psychology & Psychotherapy, 24*(2), 312–321. https://doi.org/10.1002/cpp.2022

EACLIPT

Clinical Psychology in Europe (CPE) is the official journal of the European Association of Clinical Psychology and Psychological Treatment (EACLIPT).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

International Prolonged Grief Disorder Scale Addendum for Refugees and Displaced People (IPGDS-ARD): A Study of Arabic-Speaking Bereaved Refugees

Clare Killikelly[1,2], Alexandra Reymond[1], Anaïs Aeschlimann[1], Andreas Maercker[1], Eva Heim[3]

[1] *Department of Psychology, University of Zurich, Zurich, Switzerland.* **[2]** *Department of Psychiatry, University of British Columbia, Vancouver, Canada.* **[3]** *Institute of Psychology, University of Lausanne, Lausanne, Switzerland.*

Clinical Psychology in Europe, 2025, Vol. 7(1), Article e11435, https://doi.org/10.32872/cpe.11435

Received: 2023-02-23 • **Accepted:** 2024-07-02 • **Published (VoR):** 2025-02-28

Handling Editor: Cornelia Weise, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Corresponding Author: Clare Killikelly, Department of Psychology, University of Zurich, Binzmuehlestrasse 14/17, CH-8050 Zurich, Switzerland. E-mail: c.killikelly@psychologie.uzh.ch

Supplementary Materials: Materials [see Index of Supplementary Materials]

Abstract

Background: Prolonged grief disorder (PGD) is a new and significant addition to the ICD-11 WHO disease classification system and the DSM 5-TR. As a new disorder, it stands to improve diagnostic precision, enhance communication among health professionals and patients, provide better access to care and lead to effective treatments and intervention. However, it remains to be determined if the new diagnostic criteria for PGD are applicable to different cultural groups.

Method: Here we sought to adapt the International Prolonged Grief Disorder Scale for refugees and displaced people. We conducted two focus groups with clinicians and health care workers and six cognitive interviews with bereaved Arabic-speaking refugees.

Results: This formative research resulted in an addendum (comprised of three new scales) to the IPGDS aimed to aid with treatment planning: the 42 item Addendum for Refugees and Displaced people (IPGDS-ARD). Here we present the steps for scale augmentation based on cultural considerations, a detailed description of clinical utility, feasibility and content validity established at each step, and an analysis of the percent of change in content at each step.

Conclusion: We conclude that the presented method of scale augmentation is a feasible and efficient approach that led to a culturally relevant, clinically useful addendum to an existing PGD questionnaire.

Keywords

grief, bereavement, prolonged grief disorder, cross-cultural relevance

Highlights

- This study is one of the first formative research studies on culturally relevant items for treatment planning with bereaved refugees.
- The additional scales developed seek to unpick the complex experiences of ambiguous loss, adjustment to host country and grief.
- With this work we present a detailed description of qualitative research methods for scale augmentation for bereaved refugees.

Researchers and clinicians are increasingly confronted with a difficult and exciting question: to what extent does cultural background contribute to the presentation, chronicity, and treatment of mental health disorders? The WHO's ICD-11 and the DSM-TR 5 now include cultural caveats in their definitions of mental disorders (Boelen et al., 2020; Killikelly & Maercker, 2017; Prigerson et al., 2021). For example, the new diagnostic definition of prolonged grief disorder (PGD) can only be diagnosed if symptoms persist for a longer period of time or are more intense than would be expected in the individuals' culture and context (Maercker et al., 2013). This novel addition to diagnostics catalyzes important discussions about the role of culture in mental disorder presentation and treatment. However, there are several challenges that have so far been overlooked. It is not clear how to establish the cultural norms of an individual or culture and whether these have been violated by disorder. For example, in the case of grief, cultural norms prescribe a mourning period and if grieving persists beyond the culturally expected mourning period that cultural norm has been violated. Clinicians, researchers, and health care workers on the frontline of mental health assessment are perplexed by the proposition that culture can quickly and easily be assessed in a diagnostic setting (Stelzer et al., 2020). The new disorder definition of PGD provides a unique opportunity to explore the assessment of a mental health disorder with its emphasis on culture and global applicability (First et al., 2015). Core symptoms of PGD include longing for and preoccupation with the deceased, significant emotional distress and significant functional impairment that persist beyond half a year after the death of a loved one (Killikelly & Maercker, 2017). Symptoms can differ in duration and expression according to the culture, religion, social status or gender of the bereaved (Rosenblatt, 2008). Therefore, although six months seem to be a good approximation of a grieving process, this duration is not exclusively limited. Indeed, the main distinction in diagnostic criteria between the newly proposed ICD-11 and DSM 5-TR PGD is the time criteria: the ICD-11 purports that symptoms must persist for more than six months, and the DSM suggests more than 12 months. To date, when questions about the duration of symptoms arise, clinicians are advised to use cultural norms in diagnostic decision making (O'Connor et al., 2015). Currently there

is only one PGD assessment questionnaire that includes an item examining the role of culture. The International ICD-11 Prolonged Grief Disorder Scale (IPGDS) was developed based on key ICD-11 items from the Prolonged Grief Disorder-13 (PG-13; (Prigerson & Maciejewski, 2008) and the Structured Clinical Interview for Complicated Grief (SCI-CG; Bui et al., 2015; Killikelly et al., 2020). This questionnaire was developed in two parts. Part one examines the specific diagnostic items of the ICD-11 (13 items) and includes one item related to the cultural caveat (Item 14 *My grief would be considered worse [e.g., more intense, severe and/or of longer duration] than for others from my community or culture).* Part one of this scale may be used to establish a preliminary diagnosis of PGD according to established diagnostic features. Part two of the scale is a catalogue of possible grief symptoms that may have a stronger cultural fit depending on the individual assessed, (for example Item 4: I had a strong loud emotional outburst after the loss) (Killikelly & Maercker, 2023). Part two provides the patient and clinician a more in-depth assessment tool that may help with treatment planning, therapeutic rapport and further symptom delineation but is not used for diagnostics.

This two-part assessment framework is modeled on the Harvard Trauma Questionnaire (Mollica et al., 1992). The HTQ was developed in several parts and contains a combination of checklist and open-ended questions about traumatic events and emotional symptoms that are unique to a particular place and context. The HTQ is intended to be used in clinician interviews and not as a self-report. We have built on this framework by including a standard scale of well-known PGD symptoms alongside a supplementary assessment of additional symptoms that may be culturally relevant and particularly relevant for treatment planning. In addition, we provide a new simplified framework for how to adapt the IPGDS to different cultural groups. Presently, the IPGDS has been used and adapted to assess PGD and wider symptoms of grief in German-speaking, Chinese, Japanese, Syrian bereaved and a group of Arabic speaking migrants (Killikelly & Maercker, 2023). The standard scale and the adapted cultural supplement have been psychometrically validated in German-speaking and Chinese bereaved and preliminary validated in a group of Swiss and Canadian migrants from a variety of backgrounds (Killikelly et al., 2020). The cultural supplement was found to be valid and reliable for use in each of these groups and in each case items were augmented with specific culturally relevant items. For example, Chinese participants requested and strongly endorsed Item 1 of the cultural supplement 'I experience strong physical problems since the loss (e.g., headache, problems with appetite)'.

Humanitarian migrants comprise refugees, asylum seekers or displaced people with a population rising to more than 70 million around the world (UNHCR, 2017). This group is affected by several challenges that have a negative impact on mental health including a high rate of traumatic events (Bhugra & Becker, 1999). The most common disorders are posttraumatic stress disorder (PTSD), depression and anxiety disorder (Fazel et al., 2005). Concerning PGD, Killikelly et al. (2018) found that refugees are significantly more

affected than the normal population. For example up to 54% of refugees may experience symptoms of PGD while less than 10% of the general population (Lechner-Meichsner et al., 2024; Rosner et al., 2021). Research has shown that ethnic minorities, refugees and immigrants are more likely to be misdiagnosed with mental health disorders than patients from the main culture (Bäärnhielm et al., 2015). Indeed, patients and clinicians' differences such as culture, gender, language, religion and ages can lead to a misunderstanding of the illness (Lewis-Fernández & Kirmayer, 2019). Currently there is an urgent need for a measure of PGD that is relevant and acceptable for refugees and displaced people. The main aim of this study is to culturally refine and develop a PGD addendum for refugees and displaced people that considers the distinctive experiences of refugees. This addendum may be used to supplement to the standard ICD-11 PGD diagnostic items (IPGDS standard scale) in the effort to improve therapeutic rapport, identify missing symptoms or features and improve treatment and care planning with refugees. As a proof of concept, we have explored symptoms and experiences of grief in a small group of Arabic-speaking refugees.

Augmentation of the IPGDS for Arabic-Speaking Refugees

According to Boateng et al. (2018) there are three phases to scale augmentation (item development, scale development, scale evaluation) including nine steps to create and validate a scale. This project has developed a simplified method that focuses on the two steps of the "item development" phase (identification of domain and item generation, content validity), and on the first step of the "scale development" phase, which consists in the pre-testing of questions to distill the most relevant and important information for cultural relevance (see Figure 1).

Two focus groups (FGs) were conducted: the first focused on item generation and adapting the content of the questionnaire and a second focused on determining the feasibility and clinical utility of the questionnaire. The following research questions were addressed in the focus groups:

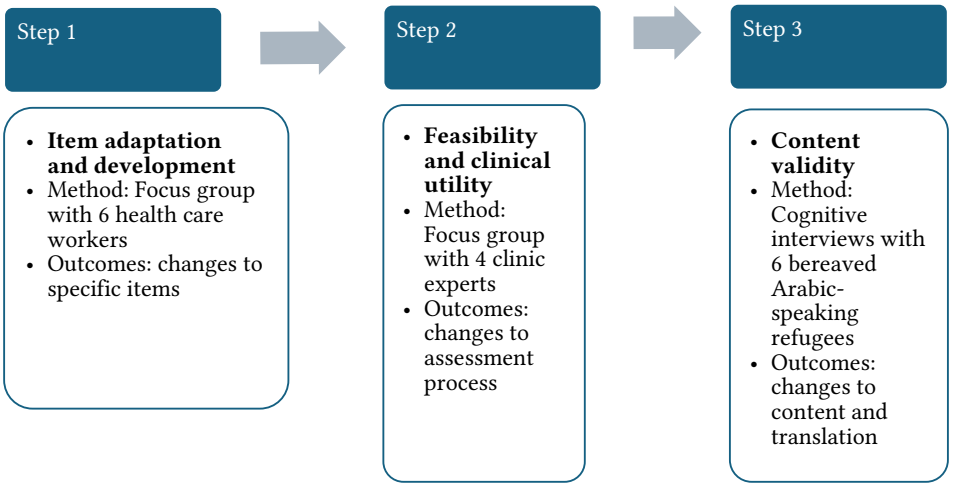
1. What content (specific items) from the existing IPGD scale is missing / needs to be added for bereaved Arabic-speaking refugees?
2. Is the IPGDS feasible for use with Arabic-speaking refugees in different clinical settings? Is the IPGDS for refugees clinically useful?

The World Health Organization (WHO) Procedure of translation and adaptation of instruments was followed to translate the IPGDS for refugees in Arabic (World Health Organization, 2019). To establish the preliminary content validity of the questionnaire in Arabic, six cognitive interviews (CIs) were conducted with bereaved Arabic-speaking refugees (5 Syrian, 1 Iraqi) to evaluate the possible sources of errors in the questionnaire and acceptability of the items (Beatty & Willis, 2007). The following research question was examined in the CI's:

3. Is the content of the IPGDS for Arabic-speaking refugees valid and acceptable?

Figure 1

Iterative Process of Cultural Augmentation and Testing



Method

Procedure

Ethical approval was obtained from the UZH Faculty of Arts and Social Sciences (Grant No. 19.10.4.)

Measure: IPGDS

This scale comprises 13 previously used items integrating the PG-13 (Prigerson & Maciejewski, 2008) and the SCI-CG (Bui et al., 2015). The participants indicated how often they felt preoccupation, yearning and symptoms of emotional distress over the past month because of loss of a loved one, using a 5-point scale: 1 = almost never (less than once a month), 2 = rarely (monthly), 3 = sometimes (weekly), 4 = often (daily), and 5 = always (several times a day). An impairment item as well as screening items (for the length of time since bereavement and the violation of socio-cultural norms) were also included. As mentioned above Item 14 assesses the cultural caveat. The original version of the cultural supplement consisted of 19 items developed from key informant interviews with health care workers and bereaved individuals from Europe and China (Stelzer et al., 2020). Psychometric analysis confirmed the internal consistency, concurrent and criterion validity. In this current study we aim to develop a new cultural supplement,

now referred to as the addendum, from focus groups and interviews with bereaved refugees and migrants.

Step 1: Item Generation and Augmentation of IPGDS for Refugees Using a Focus Group

The first FG was organized to generate new possible questionnaire items and discuss the adaptation of the IPGDS for Arabic-speaking refugees. Five professional health care workers were invited to discuss the questions of the IPGDS for refugees in general and evaluate items' relevance for the bereaved Arabic-speaking refugees in particular. This first meeting took place in December 2019 at the University of Zürich. The meeting lasted 2 hours and was facilitated by CK, Post Doctoral Researcher and Clinical psychologist with support from Master's students AR and AA. The purpose of the meeting was explained and the experts were asked to share their personal experiences of working clinically with bereaved refugees, to assess what should be added or could be missing from the standard IPGDS questionnaire, and to share their feedback/comments. The meeting was audio recorded to collect the data and transcribed with MAXQDA (version 2020). In line with the first step of scale adaptation "identification of domain and item generation", a draft catalogue of possible grief symptoms was collected based on the FG recommendations and with input from existing literature (Hassan et al., 2016; Kokou-Kpolou et al., 2017; Vromans et al., 2012). These included 57 new possible items were added to the 14 standard items of the International Prolonged Grief Disorder scale (IPGDS) (Killikelly et al., 2020).

Step 2: Feasibility and Acceptability of the IPGDS for Refugees Using a Second Focus Group

The second phase of the scale development aimed to assess the feasibility and clinical utility of the questionnaire. For this purpose, a second FG was organized with four other professional health care workers who evaluated the items to determine if the scales would be feasible for use in the clinic setting and if they would be clinically useful. The second FG took place at local psychosomatic clinic at the University Hospital Zurich in December 2019. This focus group lasted one hour. It was facilitated by CK and supported by AR and AA. The purpose of the focus group was explained, participants were asked to provide feedback on the scales, particularly considering the treatment of refugee patients who were suffering from grief.

Step 3: Preliminary Content Validity of IPGDS in Refugees Using Cognitive Interviewing

Cognitive interviews (CIs) can be used to clarify how items could be understood or how participants will answer to specific questions (Drennan, 2003). The main goal is to investigate how a participant arrives at an answer instead of the answer itself. There

are two different alternatives or paradigms to conduct CIs, the "think-aloud" method and "probing", which both aim to gather information that can't be seen in the questionnaire (Beatty & Willis, 2007). In the "think-aloud" method, the interviewer asks the participant to answer the questions by explicitly thinking out loud. It means that the interviewer asks the questions to the participant and looks at how they arrive at their answer. The other method, "verbal probing", consists of asking the participant the questions and then to elaborate on their answer. Probes can be used to ask about comprehension or interpretation of the questions, to paraphrase the questions, to ask about confidence judgment, to recall things, and to ask about very specific things as well as more general thoughts (Beatty & Willis, 2007). Both "thinking-aloud" and "verbal probing" methods are often used together in CIs as was done in the current study ((Willis & Artino, 2013).

Following the forward and back translation of the questions (WHO procedure), five CIs were conducted to pre-test the IPGDS for refugees. The interviews started with a presentation of the interviewer (AR Master's student) and the translator. A qualified clinical psychologist was always on hand (CK) in case participants became distressed. Participants were also provided with a list of local resources and psychological services. All five interviews lasted between 60 and 90 minutes. The interviewer explained the purpose of the study and of the interview and gave instructions on how to answer the questions according to the "think-aloud" method. The interviewer used probes to go deeper when insufficient information was given. After five interviews the IPGDS was adapted. A final CI with a sixth person was then conducted to pre-test this new version of the questionnaire. At the end of the interviews, participants received 30 Swiss francs for their participation.

Recruitment and Participants

Focus Group and Cognitive Interviews

The recruitment of professional health care workers for the FGs took place in November and December 2019. Clinicians were mostly recruited from health clinics that specialize in the treatment of migrant and refugee communities, private practices and NGOs in the German-speaking part of Switzerland. The main criteria were "having experience working with refugees/migrants/asylum seekers for at least one year and having professional experience with grief". The first FG consisted of six health care workers and the second of four clinicians, all of whom specialized in the treatment of trauma and grief in refugee populations.

The recruitment of participants for the CIs started in early January and lasted until the end of February 2020. Participants were first recruited through NGOs, language schools, associations, and refugee housing. The snowballing method was used for further recruitment. Inclusion criteria included speaking Arabic, the ability to provide written informed consent and having lost a loved one at least six months prior to the interview. People with severe mental health disorders (e.g., a diagnosis of major depression or

current diagnosis of schizophrenia), imminent risk of suicide, or currently receiving treatment from a psychiatric in-patient unit were excluded from the study. Five participants were from Syria and one from Iraq (Table 1).

Table 1

Demographic Information of Cognitive Interview Participants

Category	P1	P2	P3	P4	P5	P6
Gender	Female	Female	Male	Male	Male	Female
Age	28	48	23	48	46	70
Ethnicity	Syrian	Syrian	Syrian/Kurd	Syrian/Kurd	Syrian	Iraq
Residency status	Residency	Asylum Seeker	Waiting to have the refugee status	Temporary Visa	Residency	Asylum Seeker
Cause of migration	Study	War refugee	Political refugee War refugee	Political refugee War refugee	Political refugee	War refugee
Relation to the deceased	Close friend	Parent	Cousin Friends	Parent	Brother	Parent
Cause of the death	Homicide	Natural cause	Homicides	Natural cause	Homicide	Natural cause
Missing person	No	No	No	Yes	No	Yes

Analyses

The FG and CI data was audio recorded and then transcribed with MAXQDA (version 2020). The framework method of qualitative analysis was used as this method is frequently used for semi-structured interview transcripts (Gale et al., 2013). The framework approach involves the categorization and organization of the qualitative data into a matrix to reduce and summarize the data with the aim to answer the research questions and to generate themes. The framework analysis was conducted according to the seven steps proposed by Gale et al. (2013): “transcription”, “familiarization with the interview”, “coding”, “developing a working analytical framework”, applying the analytical framework”, “charting data into the framework matrix” and “interpreting the data” using Microsoft excel. Regarding the first step of the analysis, all FGs and CIs were transcribed using MAXQDA software. FGs were conducted in English and transcribed into English. CIs were conducted in English with an Arabic speaking translator and transcribed into English. This study received ethical approval from the University of Zurich.

Results

The results from this study are presented in two formats. The first type of results are the new questionnaire items and the structure for the new measure of grief. The step-by-step questionnaire adaptation and restructuring is presented in Appendix 1 (see Killikelly

et al., 2025S). Second, the results from the qualitative/descriptive data analysis of the FG and CI are presented. These results provide an in-depth rationale supporting the amendment and additions to the IPGDS. They are organized into conceptual themes of clinical utility, feasibility, and content validity. The results of the framework analysis for the two FGs (combined) and CIs are presented separately below.

Focus Groups

The first focus group resulted in a new structure of the IPGDS for Arabic-speaking refugees (see Killikelly et al., 2025S, Appendix 1, Step 1 FG1 outcomes). The cultural supplement was replaced with three new scales a) *Loss of homeland* and b) *Refugee adjustment and impact on grief* and c) *Culturally specific items*. The second focus group resulted in restructuring the scales to replace the *Loss of homeland* section with an *Ambiguous loss* section (see Killikelly et al., 2025S, Appendix 1, Step 2 FG2 outcomes). Framework analysis conducted on the two FGs revealed three major categories with several corresponding subthemes (see Table 2).

Table 2

Overview of Framework Analysis Results: Categories and Themes Resulting From FG and CI

Sample / Category	Themes
Focus groups	
Category 1: Item generation and content adaptation	Content adaptation New content
Category 2: Clinical utility	Systematic assessment Treatment priorities
Category 3: Feasibility	Duration Cultural differences
Cognitive interviews	
Content validity: *Category 1: Sources of response error*	Difficulties with response options Inapplicable items Language and meaning
Content validity: *Category 2: Expected distress reactions*	Strong emotions Refusal to respond

Category 1: Item Generation and Content Adaptation

Clinicians reviewed the content of the existing IPGDS items and made suggestions for new content and for adaptation.

Theme: Content Adaptation — Clinicians made several concrete suggestions for how the existing IPGDS could be adapted to the refugee context. Clinicians suggested including a second step in the assessment process which would include assessing grief for the homeland. Several clinicians identified how refugees may experience grief for the loss of a person but also for the loss of their community and culture (homeland).

> **FG2:** *Is it the same feeling "grieving for someone who is deceased" and "grieving for someone who we are just separated from"?*

Regarding PGD, grieving while being a refugee (how being a refugee can have an impact on the grieving and vice versa) was a main topic. According to the participants, these two experiences overlap and are hardly separable from each other.

> **FG1:** *I find it extremely complicated because refugees have an accumulated grief. And that kind of grief we have, it's connected socially, politically or religiously and normally we learn to suppress that and to deal with life as it is. [of note this clinician has a Syrian background so could speak as a cultural broker].*

Other additions include revisions to the wording and structure of the scale. Two examples are given below:

> **FG1:** *I have another question: "my grief is worse because", is it possible to make it the other way? "My grief would better if"?*

> **FG2:** *Will you integrate multiple losses? Because all these scales and criteria have been developed for people who face one loss. In the refugee context, they've lost brother and mother and friend.*

Theme: New Content — Clinicians also suggested specific items that should be added to the existing ICD-11 criteria including hopelessness, arousal, rituals, and conflicting symptoms. Fifty-seven new items were ultimately added to the catalogue of possible symptoms. These new items included the loss of homeland and the unique challenges experienced by refugees.

> **FG1:** *They are kind of preoccupied and at the same time they avoid all reminders. They can have both at the same time. It seems like paradoxical (...).*

> **FG1:** *This is a kind of avoidance, you kind of shut, you close it, life is over because someone has died. Sometimes I have the feeling when I'm with this lady as if she resigned, she gives up.*

Clinicians also mentioned typical circumstances that refugees are confronted with that should be assessed such as multiple losses, ambiguous losses (missing persons) or the lack of resources in the host country.

> **FG1:** *One was ambiguous because the husband just was gone, gone until now no one knows what happened to him, if he's dead or he lives with another family somewhere happily or whatever, no one knows.*

Category 2: Clinical Utility

Clinicians provided insight into how the assessment measure would be useful in the clinic environment.

Theme: Systematic Assessment — Clinicians identified that grief is a common complaint amongst refugees and that a questionnaire to assess it in refugees in a more systematic way is missing. They mentioned how they could use the IPGDS for refugees with their patients, for example:

> **FG2:** *It's not easy for many of our patients to differentiate between their feelings. For many, stress (or distress) is the most we can get from them. Anger, guilt, shame and grief, that's too much for many of them, so that makes it more difficult to distinguish.*

Theme: Treatment Priorities — Clinicians reported the need to have a hierarchy of what's the most important or acute symptom. Refugees often suffer from many problems such as different types of pre and post migration stressors and various types of symptoms. Clinicians must set treatment priorities in terms of what should be treated first and adjust treatment planning accordingly. For example:

> **FG2:** *So that's surely an important issue, what is the most prominent symptom, what is causing the most suffering.*

Clinicians saw an added value in using pre-post measures of the IPGDS, as this would help to track treatment progress and to indicate which symptoms of PGD would have been improved after a certain treatment.

> **FG1:** *What is the main fact [...], is it less sadness or acceptance? What is our success?*

Category 3: Feasibility

Clinicians provided specific feedback on how feasible and achievable a grief assessment for refugees would be in their clinic environments.

Theme: Duration — The most common criticism of the questionnaire was the duration. Many clinicians identified that there were too many items and the process may take too long in the clinic. They also pointed out the fact that as many refugees do not speak the language of their host country, clinicians must work with translators which adds extra time.

> **FG1:** *So even a simple BDI questionnaire takes about one hour to be translated, so the shorter the better.*

All of the clinicians agreed that the IPGDS should be shortened but they also argued that it should be able to capture important content. Clinicians therefore suggested having two separated scales in the IPGDS for refugees, one for the loss of a person and another one for the loss of a homeland. The different results on those scales could be compared:

> **FG1:** *(…) then we could apparently differentiate grief or sadness related to a loss of home country or culture and to tear that apart and grief related to the loss of a loved one. That's not the same thing and almost all refugees of course suffer from grief, so to speak related to the loss of their home country or culture (…).*

Theme: Cultural Differences — Clinicians identified one of the biggest challenges in working with refugees as differences in belief systems and culture. Indeed, clinicians expressed how each of their patients have their own way of expressing their problems or symptoms regarding their cultures or beliefs that are most of the time different from their own. They mentioned how their patients may fear being misunderstood due to cultural differences, and therefore may fear disclosing symptoms.

> **FG1:** *I have one Kurdish patient and he says "I see in Switzerland you grieve for 10 days or two weeks, but I would do it for one year. People are expecting me to move on but I'm not ready at the moment".*

Another common problem may be patients' non-acceptance of the problem and, therefore, a non-adherence to treatment. This can make diagnostic and treatment decisions difficult. In the case of PGD, participants also pointed out the multiple ways in which the disorder can manifest itself and, again, this makes its diagnosis difficult.

The use of a reference point or a cultural mediator to understand the culture of the patient was suggested. Indeed, they often use a mediator or someone from the family to help them understand the problem in the patient's context. The other strategies

to develop a common cultural understanding included: to try to step in the patients' shoes, to confront the patient with the problem, to use psychoeducation, and to plan a treatment in advance.

Cognitive Interviewing: Preliminary Content Validity

The results from the six cognitive interviews revealed two main categories with several underlying themes. These results are based on both the 'think aloud' and probing methods used interchangeably throughout the interviews.

Category 1: Sources of Response Error

Theme: Difficulties With Response Options — The response options of the IPGDS include a five-point Likert scale (not at all, rarely, sometimes, often, always). During the interviews, participants rarely used those given options and were more likely to answer with "yes", "no" or with other alternatives. When participants made an effort to answer in the manner requested, confusions or difficulties often occurred. For example, when asking Participant 3 why he chose a specific option, he answered that he chose randomly and didn't know if the response was really correct for him.

> **Interviewer:** *I try to avoid reminders of the deceased or the death as much as possible (like photos or memories).*
>
> **Participant 1:** *absolutely*

Theme: Inapplicable Items — Participants identified items that did not apply to their experience of grief or where not relevant. For example, some questions were asked assuming that participants had not attended the funeral of their loved one. However, in many cases participants had been in the same country as the person who died and could be present at the funeral. In this case the question about inability to attend a funeral or other rituals did not apply. The question about visiting the grave of the deceased person was also deemed inapplicable by all participants. For example:

> **Interviewer:** *I would do anything to feel close to the deceased (e.g., visit their grave every day, sleep next to their picture).*
>
> **Participant 5:** *It depends it's too far away, you can't imagine it's not realistic, because the grave it's too far away, I can't go to Syria.*

Another example is the question on acceptance:

> **Interviewer:** *I have trouble or just don't want to accept the loss.*

> **Participant 1:** *I don't really understand the question [...] because what does it mean or how should you accept a loss at all? [...] Yes, but the second one doesn't make sense for me. What does it mean to accept it. Like to accept, it means that I'm okay with it? You can't ask this question. I don't know, this question is a bit confusing.*

The timeframe for assessment was also questioned. Participants were asked to think about the previous week while answering the questions. Most participants had difficulties answering the questions while relating only to the last 7 days.

Theme: Language and Meaning — Participants also provided feedback on the language translation. For example:

> **Interviewer:** *I have intense feelings of sorrow, related to missing my family and friends.*
>
> **Participant 3:** *It's more an Egyptian word, not high Arabic. [wrote a different word]*

At times participants pointed out when they encountered issues with the wording and how this specific way of writing in Arabic prevented them from answering the question adequately. For example:

> **Interviewer:** *I have trouble or just don't want to accept the loss.*
>
> **Participant 1:** *There is a difference between "I have trouble or suffer from something" and "I can accept something".*

Some word problems affected all participants whereas others were only mentioned by a single person. In this case, it didn't require a change after the revision of the questionnaire.

Category 2: Expected Distress Reactions Triggered During the Interview

Two main themes, strong emotions (blaming and anger) and rejection of the questionnaire item (guilt, non-acceptance of the item) were observed as reactions to certain items. These are expected distress reactions as we expect certain emotional responses are likely to be triggered if these items are clinically valid and important.

Theme: Strong Emotions or Distress — During the interviews, all participants experienced strong emotions or distress at some point. Indeed, one participant felt very emotional when filling out the questions about how her loved one died and when she was asked if the death had been expected or not. Over the course of the interview,

most of the participants became more comfortable and were able to relax. However, two participants remained very emotional during the interview process. For example:

> **Interviewer:** *I'm longing or yearning for the deceased.*
>
> **Participant 6:** *Sure. She is always on my mind. [Crying]. It is weird to talk about her with strangers.*

In most cases, participants could answer the questions without difficulties and without feeling overwhelmed. However, a few items seemed to provoke strong emotions for all participants.

All participants expressed strong blame towards someone or something for the death of their loved one, or for the reason they had to leave their country. Four out of six participants blamed others or the circumstances for the death and all six participants blamed others or the circumstances for the reason they had to leave the country:

> **Interviewer:** *I blame others on the circumstances for the death (like a higher power).*
>
> **Participant 1:** *I blame the higher powers which are directly responsible for the problems, like the regime. Yes.*

Anger was also a strongly expressed emotion. Five out of six participants felt very angry about the loss and, again, all of them reported feeling angry over being separated from their family and friends.

> **Interviewer:** *I'm angry over the death.*
>
> **Participant 6:** *I always say: How did you leave me alone here and leave?*

Theme: Refusal to Respond — It was observed that several items seemed to trigger a rejection or denial response from participants. One participant in particular had difficulties answering certain questions in general, for example:

> **Interviewer:** *I feel that I lost a part of myself.*
>
> **Participant 3:** *I won't answer.*

Nevertheless, for most participants only a few items triggered rejection. Questions about guilt triggered ambivalent responses and seemed to bring up discomfort in all participants. Not only did they all report not feeling guilty for the death of their loved one, but some of them were even shocked or surprised by such a question, as if they would not allow themselves to be guilty:

> **Interviewer:** *I feel guilty about the death or circumstances surrounding the death.*
>
> **Participant 5:** *How should I be guilty of it? No for sure, I'm not guilty. I didn't have to do anything with the death.*

When asked about feeling guilty about being separated from their family and friends, participants were more receptive. Participants acknowledged feelings of guilt.

Importantly, one participant expressed a wish to die. Some items seemed to cause him such distress that he refused to answer. He seemed particularly nervous during the whole interview but items specifically about death or his role in life triggered even more emotions and at times he did not want to answer:

> **Interviewer:** *I want to die in order to be with the deceased.*
>
> **Participant 3:** *I don't want to answer.*

Change Analysis

In addition to the qualitative framework analysis, we examined the percentage of change in the questionnaire items, afforded by each step of the adaptation process. Here the results show that the number of changes at each step was reduced (e.g. from 80% of items changed in Step 1 to 12.5% changed in Step 3). This supports the content validity of the final questionnaire assessed by Arabic-speaking refugees in Step 3 (cognitive interviews). At this final stage only small changes to translation and some clarifications of content were required.

Discussion

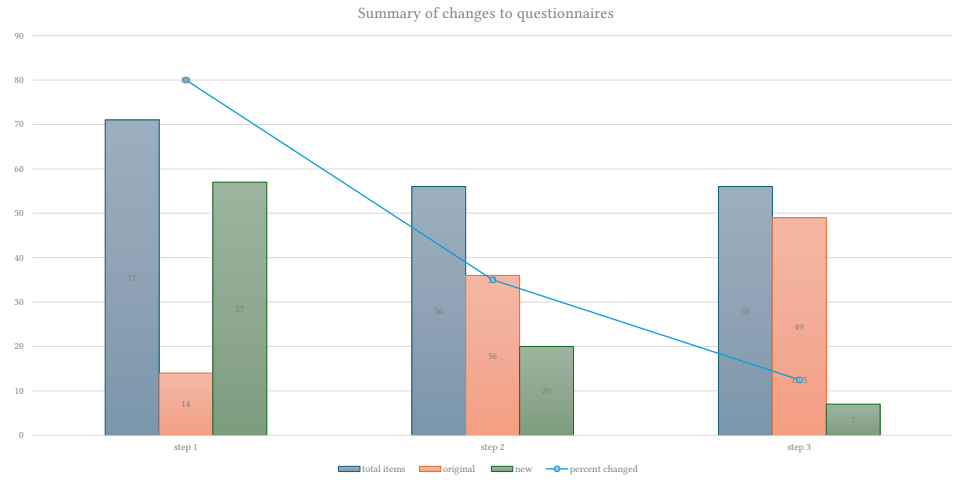
The IPGDS is the only scale that is extended to include culturally specific symptoms of grief alongside items for a prolonged grief diagnosis (Killikelly & Maercker, 2023). So far, there is no questionnaire to assess PGD and grief more generally, taking the refugee experience into account. This formative, proof of concept, research project had two overarching aims 1) to use a step-by-step method of cultural augmentation to develop an IPGDS addendum tailored to Arabic-speaking refugees and displaced persons (IPGDS-ARD) and 2) To provide an in-depth description of the rationale supporting the new content and additional scales in terms of feasibility, clinical utility, and content validity. To this end we conducted two focus groups with experts and six cognitive interviews with bereaved Arabic-speaking refugees. This resulted in a newly culturally relevant IPGDS-ARD questionnaires for bereaved Arabic-speaking refugees. Of note the current study aimed to develop and extend items for bereaved refugees more generally,

however the resulting scales have only been piloted in a small group of Arabic speaking refugees. This will be explored in follow up studies.

In the first step of cultural augmentation, FG 1, the original IPGDS standard scale measuring ICD-11 PGD symptoms was preserved however the cultural supplement questions were replaced with *a) loss of homeland scale, b) refugee adjustment and impact on grief and c) new culturally specific items*. The rationale for the inclusion of these new sub-scales is captured by the themes revealed in the framework analysis: improved clinical utility (systematic assessment, definition of treatment priorities) and feasibility (duration, cultural differences) for use with bereaved refugees. In the second step, FG 2, the sub-scales were adapted again based on clinical utility and feasibility for use in the busy clinic environment. For example, *the loss of homeland* sub-scale was changed to the *ambiguous loss* sub-scale to reflect a more clinically useful phenomenon. In the final step of the adaptation the CI revealed few changes to the content of the questionnaires. Small changes to translation and some re-phrasings of the items were made. Overall, the few changes from the six CI support the content validity of the final IPGDS-ARD questions (see Figure 2).

Figure 2

Percentage of Change in Number of Questionnaire Items After Each Augmentation Step



Culture is a crucial component for better understanding the expression of illness, that is, cultural differences must be understood and considered for an appropriate diagnosis and treatment (Bhugra, 2005; Bhugra & Becker, 1999). In this study clinicians served as cultural brokers to provide a first window of insight into the clinical utility of a grief

measure for refugees. The qualitative analysis on the FGs and CI provided rich data supporting each step of cultural augmentation and restructuring. Clinicians' recommendations and insights reflected main themes of improving the clinical utility, feasibility, and content validity of the questionnaires. These are touchstones of establishing questionnaire acceptability and validity, particularly for different cultural groups. Previous research has used similar methods including focus groups and CI to develop the content of cultural adapted mental health measures. Our previous work has also demonstrated the value of including a wide range of clinician and patient perspectives to establish cultural relevance and clinical utility of the IPGDS. Swiss and Chinese health care workers examined the content of the ICD-11 PGD guidelines and provided in depth qualitative interview data on their perspectives on the clinical utility and global applicability (Stelzer et al., 2020). Important themes were revealed including the role of stigma in preventing help seeking and treatment and the value of including somatic symptoms in the diagnostic guidelines. Interviews with Japanese health care workers confirmed the overall utility of the ICD-11 PGD guidelines but included important possible barriers to clinical assessment such as the role of emotional control and the strong shame attached to seeking grief support (Killikelly et al., 2023). Health care workers can provide vital insight into their own illness beliefs and beliefs about the usefulness of assessments, as well as insight into their patients experiences and illness models.

The outcome of this cultural adaptation is an addendum to the original standard IPGDS scale with three new subscales: *a) ambiguous loss, b) refugee adjustment and impact on grief and c) culturally specific items.* The *ambiguous loss* section will be an extremely valuable measure for refugees and displaced people. Ambiguous loss is increasingly found to be a significant source of mental distress for displaced people (Boss, 2006). It is defined as the loss of a loved one where death is not confirmed (Solheim et al., 2016). Until now there have been no validated scales or systematically developed measures to assess ambiguous loss in refugees. Renner et al. (2021) found that ambiguous loss is associated with higher levels of depression and prolonged grief. We have built on the preliminary findings from the IPGDS-ARD to develop a shorter stand-alone scale, the AL+ (Comtesse et al., 2023).

The section *refugee adjustment and impact on grief* seeks to examine how post migration experiences may hinder the grieving process as an experience specific to refugees and displaced people. Hwang et al. (2008) recommend that cultural and contextual factors are considered in a timely and consistent manner when treatment planning and that this is re-evaluated systematically. This is especially important because many minorities share similar immigration experiences that could be targeted in prevention and treatment programs. Indeed, in a study about the impact of migration on illness experience and help-seeking strategies of patients from Turkey and Bosnia, Gilgen et al. (2005) examined explanatory models to investigate how those patients understood their illness and found that refugees attributed some of their migration experiences as causes of their illness.

Kim et al. (2017) found that traumatic events can impact or diminish the ability to grieve. In their study, participants reported that the psychological sequalae following traumatic events such as torture were more significant and impairing and they saw grief as a less significant problem. The IPGDS-ARD provides clinicians and patients with concrete questions that may help unpick the source of distress and guide further assessment and tailor treatment planning towards grief interventions or towards resource building and support for post migration living stress.

It is important to note that these three new scales should be used in addition to the standard IPGDS scale if a diagnosis is sought. Only the standard IPGDS scale based on the ICD-11 PGD items can be used to confirm a diagnosis. The 3 scales in this new addendum can then be used to guide clinicians to assess and explore other grief related areas of possible concern and distress in refugee groups.

Limitations

There are several limitations in this study. First, the small sample sizes for the focus groups and cognitive interviews indicate that our results should be considered a preliminary examination. A follow up study is underway to examine the psychometric validity of these new subscales in a larger bereaved refugee sample. In addition, the sample did not include a clinical sample of patients with a confirmed diagnosis of PGD. Additional testing is needed to confirm the validity in a clinical sample of bereaved refugees with PGD. Further adaptations may be necessary to ensure the clinical utility of the subscales for different refugee groups. At the moment there are several items in the subscales which increase the administration time. Further reduction and refinement of the items may be required to improve the clinical utility. The analysis of percentage of change after each adaptation step may be biased as the participants in the CI may be hesitant to express criticism to the research team. As a next step it will be vital to further refine the addendum items through patient and clinician debriefing and to examine clinical decision making and rates of PGD diagnosis with and without the support of the addendum (see the method in Lewis-Fernández et al., 2017).

Conclusions and Future Directions

In this study, clinicians carefully and thoughtfully described the difficulties they sometimes experience providing accurate culturally reliable assessment and treatment for patients from different cultural backgrounds with a particular focus of bereaved Arabic-speaking refugees. They particularly emphasized the difficulty in assessing and treating PGD in refugees as patients often present with a myriad of symptoms and stressors. The IPGDS for refugees, along with the newly developed subscales (IPGDS-ARD) were presented to assess PGD and grief experiences more wholistically in Arabic-speaking refugees and with the aim to support other displaced people in the future. The new

IPGDS-ARD will help clinicians diagnose PGD with the standard scale as well as assess the importance and relevance of possible overlapping or co-occurring stressors such as post migration difficulties or ambiguous loss. Due to the challenges with clinical utility including the large number of items, we recommend using a formulation approach to develop a symptom map of the most distressing and clinically relevant symptoms highlighted through both the standard IPGDS scale and the addendum. An example formulation is included in Appendix 2 (see Killikelly et al., 2025S). Additionally, the newly developed and tested Ambiguous Loss Inventory + would directly assess the loss of missing loved ones (Comtesse et al., 2023). Clinicians will then be able to direct treatment to appropriate evidence-based interventions.

Funding: Swiss National Science Foundation (SNF): Mobility fellowship P400PS_191001.

Acknowledgments: We would like to thank our participants for their time and effort in completing the interviews.

Competing Interests: The authors have declared that no competing interests exist.

Ethics Statement: The authors assert that all procedures contributing to this work and the process of informed consent comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. *Ethical Approval was obtained from the University of Zurich, Faculty of Arts and Sciences (Grant No. 19.10.4.).*

Reporting Guidelines: Reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report.

Preregistration: The study was not preregistered.

Related Versions: This manuscript is based on the master's thesis of AR: Reymond, A. (2020). *Development and preliminary testing of the International Prolonged Grief Disorder Scale for refugees* [Unpublished master's thesis, University of Zurich, Department of Psychology].

Data Availability: Data are available upon request.

Supplementary Materials

The Supplementary Materials contain the online appendices for the study (see Killikelly et al., 2025S):

- **Appendix 1. Summary of questionnaire changes after each step, resulting in the Addendum for Refugees and Displaced people (IPGDS-ARD):** Here we present the process of item reduction and generation resulting in the final version of the addendum.
- **Appendix 2. Sample formulation template for assessing grief in displaced people:** Here we present a template formulation for considering treatment planning for working with grief and displaced people.

Index of Supplementary Materials

Killikelly, C., Reymond, A., Aeschlimann, A., Maercker, A., & Heim, E. (2025S). *Supplementary materials to "International Prolonged Grief Disorder Scale Addendum for Refugees and Displaced people (IPGDS-ARD): A study of Arabic-speaking bereaved refugees"* [Online appendices]. PsychOpen GOLD. https://doi.org/10.23668/psycharchives.15956

References

Bäärnhielm, S., Åberg Wistedt, A., & Rosso, M. S. (2015). Revising psychiatric diagnostic categorisation of immigrant patients after using the Cultural Formulation in DSM-IV. *Transcultural Psychiatry, 52*(3), 287–310. https://doi.org/10.1177/1363461514560657

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly, 71*(2), 287–311. https://doi.org/10.1093/poq/nfm006

Bhugra, D. (2005). Cultural identities and cultural congruency: A new model for evaluating mental distress in immigrants. *Acta Psychiatrica Scandinavica, 111*(2), 84–93. https://doi.org/10.1111/j.1600-0447.2004.00454.x

Bhugra, D., & Becker, M. A. (1999). Migration, cultural bereavement and cultural identity. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA), 4*(1), 18–24. https://doi.org/10.1002/(ISSN)2051-5545

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Boelen, P. A., Eisma, M. C., Smid, G. E., & Lenferink, L. I. M. (2020). Prolonged grief disorder in section II of DSM-5: A commentary. *European Journal of Psychotraumatology, 11*(1), Article 1771008. https://doi.org/10.1080/20008198.2020.1771008

Boss, P. (2006). *Loss, trauma, and resilience: Therapeutic work with ambiguous loss*. W. W. Norton.

Bui, E., Mauro, C., Robinaugh, D. J., Skritskaya, N. A., Wang, Y. J., Gribbin, C., Ghesquiere, A., Horenstein, A., Duan, N. H., Reynolds, C., Zisook, S., Simon, N. M., & Shear, K. (2015). The Structured Clinical Interview for Complicated Grief: Reliability, validity, and exploratory factor analysis. *Depression and Anxiety, 32*(7), 485–492. https://doi.org/10.1002/da.22385

Comtesse, H., Killikelly, C., Hengst, S. M. C., Lenferink, L. I. M., de la Rie, S. M., Boelen, P. A., & Smid, G. E. (2023). The Ambiguous Loss Inventory Plus (ALI+): Introduction of a measure of psychological reactions to the disappearance of a loved one. *International Journal of Environmental Research and Public Health, 20*(6), Article 5117. https://doi.org/10.3390/ijerph20065117

Drennan, J. (2003). Cognitive interviewing: Verbal data in the design and pretesting of questionnaires. *Journal of Advanced Nursing, 42*(1), 57–63. https://doi.org/10.1046/j.1365-2648.2003.02579.x

Fazel, M., Wheeler, J., & Danesh, J. (2005). Prevalence of serious mental disorder in 7000 refugees resettled in Western countries: A systematic review. *Lancet, 365*(9467), 1309–1314. https://doi.org/10.1016/S0140-6736(05)61027-6

First, M. B., Reed, G. M., Hyman, S. E., & Saxena, S. (2015). The development of the ICD-11 Clinical Descriptions and Diagnostic Guidelines for Mental and Behavioural Disorders. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA), 14*(1), 82–90. https://doi.org/10.1002/wps.20189

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology, 13*(1), Article 117. https://doi.org/10.1186/1471-2288-13-117

Gilgen, D., Maeusezahl, D., Salis Gross, C., Battegay, E., Flubacher, P., Tanner, M., Weiss, M. G., & Hatz, C. (2005). Impact of migration on illness experience and help-seeking strategies of patients from Turkey and Bosnia in primary health care in Basel. *Health & Place, 11*(3), 261–273. https://doi.org/10.1016/j.healthplace.2004.04.002

Hassan, G., Ventevogel, P., Jefee-Bahloul, H., Barkil-Oteo, A., & Kirmayer, L. J. (2016). Mental health and psychosocial wellbeing of Syrians affected by armed conflict. *Epidemiology and Psychiatric Sciences, 25*(2), 129–141. https://doi.org/10.1017/S2045796016000044

Hwang, W.-C., Myers, H. F., Abe-Kim, J., & Ting, J. Y. (2008). A conceptual paradigm for understanding culture's impact on mental health: The Cultural Influences on Mental Health (CIMH) model. *Clinical Psychology Review, 28*(2), 211–227. https://doi.org/10.1016/j.cpr.2007.05.001

Killikelly, C., Bauer, S., & Maercker, A. (2018). The assessment of grief in refugees and post-conflict survivors: A narrative review of etic and emic research. *Frontiers in Psychology, 9*, Article 1957. https://doi.org/10.3389/fpsyg.2018.01957

Killikelly, C., Hasenöhrl, A., Stelzer, E. M., & Maercker, A. (2023). The new ICD-11 prolonged grief disorder guidelines in Japan: Findings and implications from key informant interviews. *Culture, Medicine and Psychiatry, 47*(2), 519–542. https://doi.org/10.1007/s11013-022-09781-6

Killikelly, C., & Maercker, A. (2017). Prolonged grief disorder for ICD-11: The primacy of clinical utility and international applicability. *European Journal of Psychotraumatology, 8*(Sup6), Article 1476441. https://doi.org/10.1080/20008198.2018.1476441

Killikelly, C., & Maercker, A. (2023). The cultural supplement: A new method for assessing culturally relevant prolonged grief disorder symptoms. *Clinical Psychology in Europe, 5*(1), Article e7655. https://doi.org/10.32872/cpe.7655

Killikelly, C., Zhou, N., Merzhvynska, M., Stelzer, E. M., Dotschung, T., Rohner, S., Sun, L. H., & Maercker, A. (2020). Development of the International Prolonged Grief Disorder Scale for the ICD-11: Measurement of core symptoms and culture items adapted for Chinese and German-speaking samples. *Journal of Affective Disorders, 277*, 568–576. https://doi.org/10.1016/j.jad.2020.08.057

Kim, J., Tol, W. A., Shrestha, A., Kafle, H. M., Rayamajhi, R., Luitel, N. P., Thapa, L., & Surkan, P. J. (2017). Persistent complex bereavement disorder and culture: Early and prolonged grief in Nepali widows. *Psychiatry, 80*(1), 1–16. https://doi.org/10.1080/00332747.2016.1213560

Kokou-Kpolou, K., Mbassa Menick, D., Moukouta, C. S., Baugnet, L., & Kpelly, D. E. (2017). A cross-cultural approach to complicated grief reactions among Togo–Western African immigrants in Europe. *Journal of Cross-Cultural Psychology, 48*(8), 1247–1262. https://doi.org/10.1177/0022022117721972

Lechner-Meichsner, F., Comtesse, H., & Olk, M. (2024). Prevalence, comorbidities, and factors associated with prolonged grief disorder, posttraumatic stress disorder and complex posttraumatic stress disorder in refugees: A systematic review. *Conflict and Health, 18*(1), Article 32. https://doi.org/10.1186/s13031-024-00586-5

Lewis-Fernández, R., Aggarwal, N. K., Lam, P. C., Galfalvy, H., Weiss, M. G., Kirmayer, L. J., Paralikar, V., Deshpande, S. N., Díaz, E., Nicasio, A. V., Boiler, M., Alarcón, R. D., Rohlof, H., Groen, S., Van Dijk, R. C. J., Jadhav, S., Sarmukaddam, S., Ndetei, D., Scalco, M. Z., . . . Vega-Dienstmaier, J. M. (2017). Feasibility, acceptability and clinical utility of the Cultural Formulation Interview: Mixed-methods results from the DSM-5 international field trial. *The British Journal of Psychiatry, 210*(4), 290–297. https://doi.org/10.1192/bjp.bp.116.193862

Lewis-Fernández, R., & Kirmayer, L. J. (2019). Cultural concepts of distress and psychiatric disorders: Understanding symptom experience and expression in context. *Transcultural Psychiatry, 56*(4), 786–803. https://doi.org/10.1177/1363461519861795

Maercker, A., Brewin, C. R., Bryant, R. A., Cloitre, M., Reed, G. M., Van Ommeren, M., Humayun, A., Jones, L. M., Kagee, A., Llosa, A. E., Rousseau, C., Somasundaram, D. J., Souza, R., Suzuki, Y., Weissbecker, I., Wessely, S. C., First, M. B., & Saxena, S. (2013). Proposals for mental disorders specifically associated with stress in the International Classification of Diseases-11. *Lancet, 381*(9878), 1683–1685. https://doi.org/10.1016/S0140-6736(12)62191-6

Mollica, R. F., Caspi-Yavin, Y., Bollini, P., Truong, T., Tor, S., & Lavelle, J. (1992). The Harvard Trauma Questionnaire: Validating a cross-cultural instrument for measuring torture, trauma, and posttraumatic stress disorder in Indochinese refugees. *The Journal of Nervous and Mental Disease, 180*(2), 111–116. https://doi.org/10.1097/00005053-199202000-00008

O'Connor, M., Nickerson, A., Aderka, I. M., & Bryant, R. A. (2015). The temporal relationship between change in symptoms of prolonged grief and posttraumatic stress following old age spousal bereavement. *Depression and Anxiety, 32*(5), 335–340. https://doi.org/10.1002/da.22349

Prigerson, H. G., Boelen, P. A., Xu, J., Smith, K. V., & Maciejewski, P. K. (2021). Validation of the new DSM-5-TR criteria for prolonged grief disorder and the PG-13-Revised (PG-13-R) scale. *World Psychiatry: Official Journal of the World Psychiatric Association (WPA), 20*(1), 96–106. https://doi.org/10.1002/wps.20823

Prigerson, H. G., Vanderwerker, L. C., & Maciejewski, P. K. (2008). A case for inclusion of prolonged grief disorder in DSM-V. In M. S. Stroebe, R. O. Hansson, H. Schut, & W. Stroebe (Eds.), *Handbook of bereavement research and practice: Advances in theory and intervention* (pp. 165–186). American Psychological Association. https://doi.org/10.1037/14498-008

Renner, A., Jäckle, D., Nagl, M., Plexnies, A., Röhr, S., Löbner, M., Grochtdreis, T., Dams, J., König, H. H., Riedel-Heller, S., & Kersting, A. (2021). Traumatized Syrian refugees with ambiguous loss: Predictors of mental distress. *International Journal of Environmental Research and Public Health, 18*(8), Article 3865. https://doi.org/10.3390/ijerph18083865

Rosenblatt, P. C. (2008). Grief across cultures: A review and research agenda. In M. Stroebe, R. Hansson, H. Schut, & W. Stroebe (Eds.), *Handbook of bereavement research and practice: Advances in theory and intervention* (pp. 207–222). American Psychological Association. https://doi.org/10.1037/14498-010

Rosner, R., Comtesse, H., Vogel, A., & Doering, B. K. (2021). Prevalence of prolonged grief disorder. *Journal of Affective Disorders, 287*, 301–307. https://doi.org/10.1016/j.jad.2021.03.058

Solheim, C., Zaid, S., & Ballard, J. (2016). Ambiguous loss experienced by transnational Mexican immigrant families. *Family Process, 55*(2), 338–353. https://doi.org/10.1111/famp.12130

Stelzer, E. M., Zhou, N., Merzhvynska, M., Rohner, S., Sun, H., Wagner, B., Maercker, A., & Killikelly, C. (2020). Clinical utility and global applicability of prolonged grief disorder in the ICD-11 from the Perspective of Chinese and German-speaking health care professionals. *Psychopathology, 53*(1), 8–22. https://doi.org/10.1159/000505074

UNHCR. (2017). *UNHCR – Figures at a Glance.* http://www.unhcr.org/figures-at-a-glance.html

Vromans, L., Schweitzer, R. D., & Brough, M. (2012). The Multidimensional Loss Scale: Validating a cross-cultural instrument for measuring loss. *The Journal of Nervous and Mental Disease, 200*(4), 349–357. https://doi.org/10.1097/NMD.0b013e31824cc458

Willis, G. B., & Artino, A. R., Jr. (2013). What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. *Journal of Graduate Medical Education, 5*(3), 353–356. https://doi.org/10.4300/JGME-D-13-00154.1

World Health Organization. (2019). *Process of translation and adaptation of instruments.* https://terrance.who.int/mediacentre/data/WHODAS/Guidelines/WHODAS%202.0%20Translation%20guidelines.pdf

Clinical Psychology in Europe

PsychOpen GOLD

Research Articles

Check for updates

Effectiveness of Empower-Grief for Relatives of Palliative Care Patients: Protocol for an Exploratory Randomized Controlled Trial

David D. Neto [1,2], Alexandra Coelho [1,2], Sara Albuquerque [3], Ana Nunes da Silva [4,5]

[1] *School of Psychology, ISPA – Instituto Universitário, Lisbon, Portugal.* **[2]** *APPsyCI – Applied Psychology Research Center Capabilities & Inclusion, Lisbon, Portugal.* **[3]** *HEI-Lab: Digital Human-Environment Interaction Labs, Lusófona University, Lisbon, Portugal.* **[4]** *Faculdade de Psicologia, Universidade de Lisboa, Lisbon, Portugal.* **[5]** *CICPSI, Lisbon, Portugal.*

Clinical Psychology in Europe, 2025, Vol. 7(1), Article e14307, https://doi.org/10.32872/cpe.14307

Received: 2024-04-02 • **Accepted:** 2024-10-02 • **Published (VoR):** 2025-02-28

Handling Editor: Winfried Rief, Philipps-University of Marburg, Marburg, Germany

Corresponding Author: David D. Neto, ISPA – Instituto Universitário, Rua Jardim do Tabaco 42, 1100-081, Lisbon, Portugal. Phone: +351 218811700. E-mail: dneto@ispa.pt

Abstract

Background: Grief reactions of relatives of palliative care patients are seldom addressed. Most interventions focus on Prolonged Grief Disorder (PGD) and not on its prevention. This is particularly relevant in palliative care, in which death is the result of a difficult period of a terminal illness, making caregivers particularly vulnerable to psychological distress. The purpose of the present exploratory trial is to test the efficacy of a selective intervention (Empower-Grief) for the initial problematic grief reactions and to study potential predictors of adherence and efficacy.

Method: This is an exploratory Randomized Controlled Trial (RCT) studying Empower-Grief compared with Treatment as Usual (TAU). Participants will be relatives or caregivers of palliative and oncological patients with initial indicators of risk of developing PGD and will be randomly allocated to Empower-Grief and TAU. Participants will be assessed prior, at the end and six months after the intervention. The primary outcome considered will be symptoms of PGD. The assessment includes measures of anxiety and depression, coping, attachment, psychological flexibility, posttraumatic growth, social support and therapeutic alliance.

Results: The trial is ongoing. Forty-four participants will be invited to participate.

Conclusion: This study addresses the need for the development of empirically grounded and feasible interventions aimed at dealing with initial problematic reactions in grief, exploring

potential predictors and possible venues for personalizing intervention and understanding the mechanism through which these interventions operate.

Keywords

prolonged grief disorder, palliative care, Empower-Grief, psychological intervention, bereavement

Highlights

- Caregivers of deceased cancer patients are particularly vulnerable to Prolonged Grief Disorder.
- There is a significant need to develop low-intensity interventions for emotional reactions to grief.
- The current protocol aims to test the efficacy of Empower-Grief and to identify predictors of outcome.

Background

Prolonged Grief Disorder (PGD), recognized as a mental illness in the ICD-11 (World Health Organization, 2019) and more recently in the DSM-5-TR (American Psychiatric Association, 2022), consists of intense grief along with significant social and occupational dysfunction persisting for an extended period (with the inclusion in the DSM-5-TR, the temporal criterion was extended from 6 to 12 months).

Caregivers of patients with advanced cancer are particularly vulnerable to elevated psychological distress. In a caregiver sample, PGD prevalence was 40% at six months, 28% at 13 months, and 27% at 18 months (Guldin et al., 2012), thus supporting the idea that end-of-life caregiving increases the risk of mental health disturbance. High levels of PGD were also found in a Portuguese population of palliative care caregivers (Coelho et al., 2022).

Considering the elevated prevalence of disordered reactions, the goal of providing different interventions according to the risk of developing PGD has the potential to provide care to more individuals. These interventions range from psychoeducation (i.e., universal interventions) to low-intensity selective and high-intensity indicative interventions. Very few low-intensity interventions have been proposed. Lichtenthal and colleagues (Lichtenthal et al., 2022) developed a brief manualized cognitive-behavioural, acceptance-based intervention for critically ill patients´ caregivers called EMPOWER (Enhancing and Mobilizing the POtential for Wellness and Resilience). EMPOWER is a psychological intervention composed of six modules based on cognitive-behavioural and acceptance-based interventions: 1) initial assessment, rationale and adherence to the intervention; 2) resources for stabilization; 3) psychoeducation on grief and the cognitive-behavioural model; 4) promoting experiential acceptance; 5) imagined dialogue; 6) coping training. This intervention has been shown to be feasible, acceptable and effective in reducing

psychological symptoms, including PGD, depression and anxiety. This program, designed initially for intensive care contexts, emphasizes the detrimental role of experiential avoidance (i.e., the tendency to avoid unpleasant feelings) in decision-making and the grieving process. As opposed to acceptance, this coping mechanism has been associated with higher anticipatory grief (Davis et al., 2017) and the persistence of complicated grief in palliative care (Eisma & Stroebe, 2021).

The majority of the family caregivers will require selective intervention, oriented at people identified as being at risk of developing PGD. The differentiation of interventions allows combining need-based timely interventions to prevent PGD while increasing access through rationalization of service delivery. Despite the recommendations, no studies have matched the intensity of interventions to the severity of psychological reaction to death in preventing prolonged grief. The present research intends to implement and evaluate the effectiveness of the Empower-Grief intervention (Coelho et al., 2024; Lichtenthal et al., 2022) in the population of family caregivers of patients accompanied in the Palliative Medicine Unit of the Centro Hospitalar Universitário Lisboa Norte (CHULN). Empower-Grief is an adaptation of the original EMPOWER to the current palliative care context and post-mortem stage. It consists of the same six modules of the original intervention, but a 50-minute session is devoted to each of the modules, with two booster sessions done after treatment. It is a manualized treatment in which each session is structured.

This research project has two objectives. First, to evaluate the effectiveness of Empower-Grief Intervention compared to treatment as usual (TAU) in terms of symptoms of prolonged grief and psychological distress in relatives or caregivers of palliative and oncological patients. The second objective is to identify predictors, among the factors consistently found in the literature, of response to intervention. This second objective is crucial, considering the low-intensity level of the intervention. In addition to the risk of prolonged grief, other factors such as prior mental health, the nature of the death, social support, attachment style and psychological flexibility may be relevant in explaining the response to this intervention and serve as essential variables in adapting the intervention to the needs of the participants. This research is crucial to influence national policy towards a greater emphasis on prevention and early intervention, making the allocation of cost-effective bereavement support services the most efficient and sustainable approach for a significant public health impact in bereavement care.

Method

Design

The present study is an exploratory Randomized Controlled Trial (RCT) with two parallel groups (Figure 1) comparing the experimental treatment conditions, Empower-Grief, and

treatment as usual (TAU) in a medical centre. All study procedures have been approved by the local and central institutional review boards. The project was approved by the ethics committee of ISPA (I-138-2-24).

Setting and Participants

Treatment is offered at CHULN Lisboa Norte to family members or other caregivers of patients, followed by the palliative care unit and oncology service. Inclusion criteria involve individuals more than 18 years old, having experienced the death of a close person (e.g., relative, partner, friend) due to cancer in a palliative or oncological context from three to 12 months, and sufficient cognitive abilities and proficiency in the Portuguese language who provide written informed consent. Participants are invited from the service registry, where the reference caregiver is identified. Exclusion criteria include individuals reporting a diagnosis of pre-existing severe or active mental disorder predating the loss (e.g., schizophrenia, bipolar disorder, major depression). Participants currently undergoing psychological intervention will also be excluded. While medication will be monitored, it will not be an exclusion criterion.

Recruitment, Enrollment, and Randomization

Family members will be contacted by phone from the CHULN. The number of participants who refuse to participate, change their address, or are unreachable will be registered. Unreachable participants are considered those who are not reached after one month of attempts. Second, the protocol of contact is established and will be revised to address any concerns participants may have. Non-participation data will be controlled monthly (weekly at the beginning of the data-gathering period).

Participants will undergo assessment for PGD risk and then be randomly assigned to one of the two conditions. Those considered at risk, using the cut-off values (i.e., 7 or more) of Risk Assessment for Grief (Ministry of Health, 2017) to be at a moderate level, will be invited to additional intervention and participation in the current study. All participants will provide informed consent before initiating one of the two interventions. The informed consent process will cover the study's purpose, procedures, potential risks, and benefits, and it will be obtained from each participant before any data collection or randomization occurs. Participants will be informed that their data will be anonymized and securely stored.

The analysis follows an intention-to-treat principle, ensuring the inclusion of every randomized participant. Participants who do not meet the criteria will be carefully referred to an appropriate service. In cases where participants refuse, do not adhere to, or do not benefit from a particular intervention, they will be offered the intervention at a higher level in the referred services.

Randomization is independently conducted by a research assistant. Participants will be randomly assigned to Empower-Grief vs. TAU. Randomization occurs after a participant meets eligibility criteria, provides consent, and undergoes risk assessment. A random sequence of numbers or allocation codes will be generated using a randomization tool. This sequence will determine the allocation of participants to the two intervention groups. The participants will be allocated to each branch of the RCT through block randomization. Clinicians will not be aware of any research assessments, including the initial risk assessment, and will not be involved in assessing participants' outcomes.

Throughout the study, the project team will continuously monitor the randomization process to ensure it is executed as planned, with any deviations documented and addressed. The study protocol was written in accordance with the SPIRIT 2013 Statement (Standard Protocol Items: Recommendations for Interventional Trials (Chan et al., 2013).

Strategies to enhance participant retention and minimize attrition will be employed, including regular check-ins, reminders for follow-up assessments, and flexible scheduling options. Also, a feedback system for participants to express concerns or provide input on the study's procedures will be put in place, and this will be used to make necessary adjustments. Throughout their participation, we will emphasize participants' contribution to improving grief support for others and the rigorous privacy and confidentiality measures in place.

Treatment Conditions

TAU

Treatment as usual (TAU) consists of supportive psychotherapy based on a non-structured and integrative method that focuses on developing more adaptive coping strategies and understanding and working on the patient's internal models of self, others, and the world (Winston & Lujack, 2015). For the present investigation, TAU will be considered during the same period as the Empower-Grief intervention – that is, 12 weeks – as the frequency of sessions may differ slightly from case to case. Psychologists with specific training for each intervention will administer both interventions.

Empower-Grief

Empower-Grief (Coelho et al., 2024) is a cognitive-behavioural and acceptance-based intervention divided into six modules adapted from EMPOWER (Lichtenthal et al., 2022) to a post-mortem application. It consists of six in-presence or online 50-minute sessions and two booster sessions 2 and 4 weeks after the final intervention. It is a manualized treatment in which each session has a specific goal: 1) Welcome, initial assessment and adherence to the intervention; 2) Psychoeducation and transmission of resources for stabilization; 3) Psychoeducation on Trauma, Grief and Cognitive-Behavioral Model; 4) Promoting Experiential Acceptance; 5) Imagined dialogue; 6) Coping Training; and two

final boosting sessions. In every session, the impact of the previous consultation and evolution is evaluated.

Treatment Fidelity

Various strategies will be employed to ensure the effective delivery of the intervention and maintain consistency among psychologists. The standardized intervention protocol for "Empower-Grief" (Coelho et al., 2024) will be utilized (manualized intervention).

The implementation team for the Empower-Grief will comprise four psychologists, all with master's degrees in clinical and health psychology and at least one year of clinical experience. This group underwent a comprehensive 20/30-hour training in the Empower-Grief intervention model. In addition to the standardized training provided to all psychologists, model sessions were conducted where experienced psychologists, each with a minimum of 10 to 15 years of clinical experience in grief intervention, demonstrated how to deliver interventions effectively. These experienced psychologists will serve as valuable resources in ensuring the fidelity and quality of intervention delivery across the team.

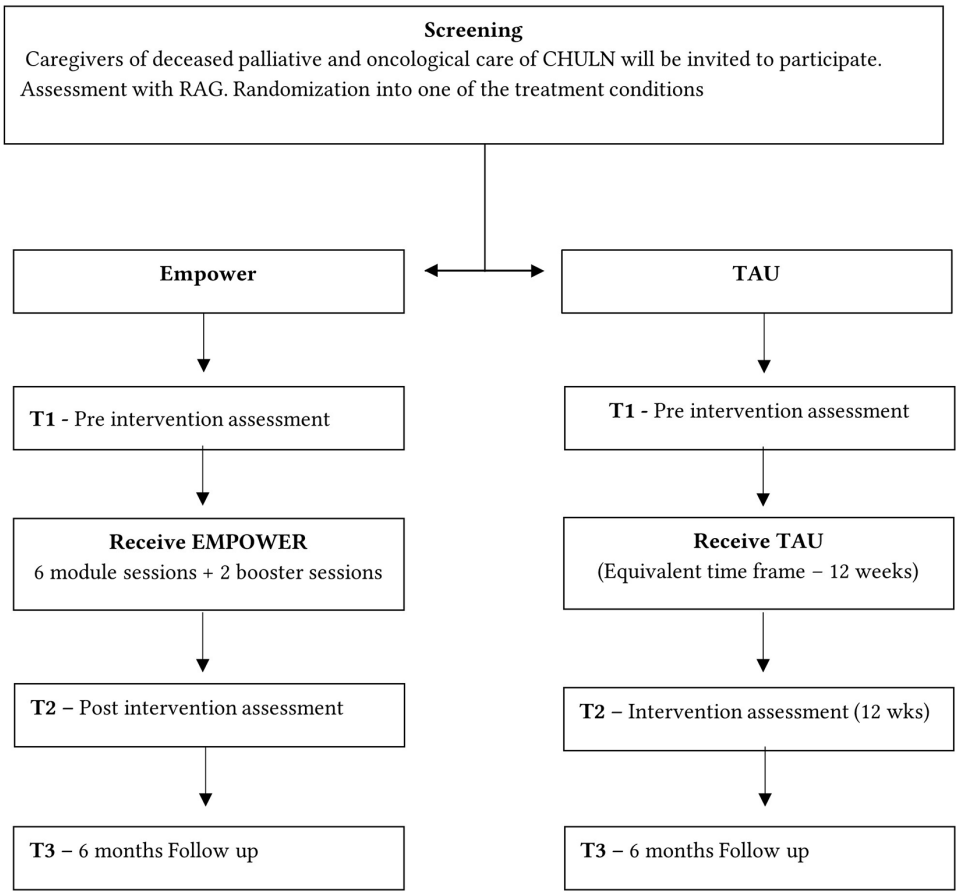
The implementation team of TAU will include three licensed psychotherapists or MA-level psychologists in advanced postgraduate clinical training. All therapists will have at least five years of experience and at least two years of experience working with complicated grief.

During the trial, Empower-Grief psychologists will have weekly group supervision, while TAU psychologists will have their usual practices that include team reunions to discuss cases. Ongoing supervision and monitoring of psychologists in the experimental group will include regular check-ins, feedback sessions, and opportunities for providers to seek guidance or clarification. All providers will have access to the same materials, resources, and tools required for their sessions.

Fidelity checks will be conducted to ensure quality and consistency across psychologists. Purposely constructed fidelity assessment tools will include a checklist with intended tasks, content, format, and psychologists' behaviour, completed after each session and discussed in supervision. Records will be maintained for each fidelity check, including dates, psychologist names, and fidelity assessment results, along with any corrective actions taken. The primary goal of these procedures is to ensure treatment adherence, maintain low non-participation rates, and control differential non-participation rates, particularly between follow-up times and types of termination. The expected outcomes include overall low non-participation rates, specifically in the considered conditions.

Figure 1

Overview of the Study Design



Assessments and Instruments

Three measurement points will be considered (Table 1). Considering the context of data gathering, care has been taken to avoid overburdening participants with excessive questionnaires. The screening will include sociodemographic data and RAG. The first assessment, before the first session, includes some additional sociodemographics, PG13-R, HADS, Brief COPE, ECR-RS, AAQ-II, and MSPSS. The second assessment will include PG13-R, HADS and the working alliance measure (WAI-S). The follow-up assessment period (3rd assessment) will occur six months after the last assessment. These will include all outcomes (PG13-R, HADS, PTGI). The interventions will share the same assessment

periods and instruments. The follow-up assessment ensures that all participants are evaluated one year after the death, enabling a PGD diagnosis.

Table 1

Time Points for Measurement

Instrument	Screening	T1	T2	6 mths
Sociodemographics	x	x		
Risk Assessment for Grief (RAG)	x			
PG-13-R		x	x	x
The Hospital Anxiety and Depression Scale (HADS)		x	x	x
The Posttraumatic Growth Inventory (PTGI)				x
Brief COPE		x		
Experiences in Close Relationships (ECR-RS)		x		
Acceptance and Action Questionnaire II (AAQ-II)		x		
The Multidimensional Scale of Perceived Social Support (MSPSS)		x		
The Working Alliance Inventory-short form (WAI-S)			x	

Note. T1 = Pre-treatment; T2 = 12 weeks/Post-treatment; 6 months = follow up.

Participants will be contacted sensitively to inquire about the reasons for discontinuing the intervention. Clinical teams will receive supervision for each intervention, and adherence to treatment measures will be monitored.

Screening

Sociodemographic data, such as participants' gender, age, marital status, education, employment, and mental health and treatment history, are also collected.

Risk Assessment for Grief (RAG) — RAG is a hetero-assessment scale of PGD risk developed by the New Zealand Ministry of Health (Ministry of Health, 2017) and adapted for the DGS norm 003/2019 (DGS, 2019). It includes four items (anger, accusation/guilt, current relationships and general coping with grief) evaluated on a 5-point Likert scale. This instrument classifies the degree of PGD risk as low (<7 points), moderate (7-10 points) and high (≥ 10 points).

Primary Outcome

Prolonged Grief Scale – Revised (PG13-R) — The PG13-R is a self-report scale developed by Prigerson and colleagues (Prigerson et al., 2021) based on DSM-5-TR criteria for PGD. PG-13-R includes 13 items, evaluated on a 5-point Likert scale. PG-13-R grief symptoms represent a unidimensional construct with high degrees of internal consistency in research conducted at Yale (Cronbach's α = .83), Utrecht (Cronbach's α = .90), and Oxford (Cronbach's α = .93) universities. PGD diagnosis is attributed when the person obtains a total value greater than 30 and meets the temporal (12 months) and impairment criteria.

Secondary Outcome

The Hospital Anxiety and Depression Scale (HADS) — The HADS, originally developed for individuals with somatic diseases in hospitals (Zigmond & Snaith, 1983), measures anxiety and depression symptoms. The scale was developed for non-psychiatric populations and includes 14 items evaluated on a 4-point (0-3) Likert scale. Values higher than eight in each subscale suggest clinically relevant symptomatology. In the Portuguese adaptation (Pais-Ribeiro et al., 2007), Cronbach alpha values were .76 in anxiety and .81 in depression.

Predictors and Other Relevant Variables

Posttraumatic Growth Inventory (PTGI) — The PTGI (Tedeschi & Calhoun, 1996) evaluates the extent of perceived positive transformations following a challenging event. Comprising 21 items, it encompasses five subscales: personal strengths, appreciation of life, new possibilities, spiritual changes, and relation to others. Participants were instructed to assess each item on a 6-point Likert scale, ranging from 1 to 6. The overall PTGI score represents the sum of all items and higher scores indicated higher levels of posttraumatic growth. The original version of the PTGI has shown excellent internal reliability (Cronbach's $\alpha = 0.90$) and acceptable test-retest reliability ($r = .71$). The PTGI has also shown good psychometric properties in breast cancer Portuguese samples (Silva et al., 2009).

Brief COPE — The Brief COPE (Carver, 1997) is composed of 28 items organized into 14 coping strategies. Each coping strategy is measured using two items (humour, positive reframing, emotional and social support, acceptance, religion, instrumental support, planning, active coping, behavioural disengagement, self-blaming, substance use, venting, self-distraction, and denial). Participants were asked to answer each item with a Likert-type response scale ranging from 0 (I never do this) to 3 (I always do this). Mean scores were calculated for each coping factor. The study of internal consistency for each factor using Cronbach's alpha shows adequate values (in general $\alpha \leq .60$), taking into account there are only two items per factor for the original (Carver, 1997) and the Portuguese version (Ribeiro & Rodrigues, 2004).

Experiences in Close Relationships (ECR-RS) — The ECR-RS (Fraley et al., 2011) is a self-report instrument that measures adult attachment to relevant persons – in this case, the deceased – rated on a seven-point Likert scale that ranges from 1 (strongly disagree) to 7 (strongly agree). Its nine items are grouped into two dimensions: attachment-related anxiety (Items 1-6) and avoidance (Items 7-9). The total subscale score consists of the mean of the items and ranges from 1 to 7, with higher scores indicating higher attachment avoidance or anxiety. The Portuguese version showed adequate reliability (α ranged from .72 to .91) and construct validity (Moreira et al., 2015).

Acceptance and Action Questionnaire II (AAQ-II) — The AAQ (Bond et al., 2011) is a 7-item questionnaire with a single-factor structure measuring acceptance, experiential avoidance, and psychological inflexibility. In the Portuguese validation (Pinto-Gouveia et al., 2012), results from a Confirmatory Factor Analysis showed the goodness of fit of the model, and this version also demonstrated an excellent level of internal consistency (α = .90) and good convergent and discriminant validity. Individuals are asked to rate each statement on a 7-point Likert scale ranging from 1 (never true) to 7 (always true). This scale reflects the single domain of psychological inflexibility, with higher scores indicating greater psychological inflexibility.

The Multidimensional Scale of Perceived Social Support (MSPSS) — The MSPSS (Zimet et al., 1988) measures perceived social support from family, friends, and others. The 12 items are grouped into three factors (family, friends and significant others), each with four items, using a 7-point Likert scale (0 = strongly disagree, 7 = strongly agree). Examples of items include “There is a special person who is close by when I need him/her"; "I can talk about my problems with my family”. The Portuguese version of the MSPSS (Carvalho et al., 2011) demonstrated good psychometric qualities (α ranged from .85 and .95).

The Working Alliance Inventory-Short Form (WAI-S) — The WAI-S (Tracey & Kokotovic, 1989) is a widely studied 12-item self-report that measures how the client perceives the therapeutic alliance with the therapist, with a 7-point Likert scale ranging from 1 (never) to 7 (always). In addition to the overall alliance score, the WAI-S includes three specific subscales: bond, tasks, and goals (four items, respectively). Internal consistency for the Portuguese version is .89 for the patient version and .85 for the therapist version (Machado & Horvath, 1999).

Statistical Analysis

To evaluate the impact of Empower-Grief relative to TAU on the primary outcome (prolonged grief symptoms) and secondary outcome (anxiety and depression), we will use a mixed-effects model. A mixed-effects model for longitudinal data (pre-post and a 6-month follow-up) accounts for the hierarchical structure of the data, with participants nested within treatment groups and the correlation between repeated measurements within subjects. The intention-to-treat principle will be followed in the analysis, ensuring that all participants randomized to each treatment condition are included in the analysis, regardless of whether they completed all assessments or adhered to the treatment protocol. To address missing data, multiple imputations will be employed, a statistical method that generates multiple plausible values for missing data points based on the observed data. This method preserves the relationships between variables and reduces the potential for bias due to missing data.

Before drawing inferences, a preliminary examination of model assumptions, including checks for the normality of residuals and homogeneity of variances, will be conducted. The primary outcome, prolonged grief severity, will be modelled using a linear mixed-effects model with fixed effects for treatment condition (Empower-Grief vs. TAU), time point (pre-treatment, post-treatment, and 6-month follow-up), and their interaction, and random effects for participants and treatment groups. Secondary outcomes, anxiety and depression symptoms, will be analyzed using similar mixed-effects models.

The sample size for this research project was estimated using G*Power software, employing a repeated measures analysis with a mixed-effects model. Assuming a low effect size of $d = 0.25$, a significance level of $\alpha = 0.05$, and a power of $1\text{-}\beta = 0.80$. With two treatment groups, Empower-Grief and treatment as usual, and measurements taken at three moments (baseline, post-treatment, and 6-month follow-up), the resulting estimated sample size for the study was 44 participants for both treatments, representing a robust size to detect the anticipated effect with high statistical power.

To assess the predictive value of the considered predictors (i.e., posttraumatic growth, coping, attachment to the deceased, psychological flexibility, social support, and therapeutic alliance) on treatment outcomes, separate regression analyses will be conducted for each outcome measure (prolonged grief, anxiety, and depression). For each outcome, we will fit a mixed-effects model with the baseline variable as a fixed effect, adjusting for treatment condition, time, and their interaction. The random effects' structure will remain the same as in the primary analysis.

Results

The RCT is ongoing. We have contacted 294 potential participants and recruited 44 to this moment. Recruitment will close in April 2024. We plan to complete data analyses by December 2024.

Discussion

This study responds to a lack of research in grief interventions, which is relevant considering the high prevalence of PGD and the general delay in providing care. Assessing the differentiation of intervention leads to increases and equality in access to empirically based interventions. By empowering caregivers to respond to a challenging grief process, the intervention may contribute to preventing pathological long-term psychological reactions. The present research project aims to assess the efficacy of a low-intensity selective intervention compared with TAU in terms of prolonged grief symptoms and distress. The project's second goal is to identify potential predictors of the outcome and adherence to these treatments – considering the matching conditions.

This research has several strengths. First, the study of low-intensity interventions – recommended by international and national bodies (Shear et al., 2017) – has a good cost-benefit ratio and increases access to psychological interventions in this social context. Second, Empower-Grief is flexibly delivered so that it can be adapted to caregivers' dynamic needs and post-mortem context. Third, it was developed and refined using stakeholder feedback to optimize its acceptability and fit for its recipients' specific needs. Also, identifying predictors of adherence and change serves to increase sensitivity in aligning interventions to the needs of clients.

The fact that the current study is conducted in a practice setting leads to some constraints – namely, the lack of manualization of TAU or the need to keep the research questionnaire relatively small. However, the practice-based nature of the context of data gathering provides a demonstration of how a structured intervention such as Empower-Grief can be useful. Results from this study will, therefore, provide a pathway for improving clinical decision-making and tailoring treatments to an individual's specific needs, which is crucial to meeting citizens' needs in the context of the rational provision of services.

This research is situated in the field of intervention personalization and psychological services research. It can inform and enrich the predicting process by considering other dimensions of patient characteristics, such as psychological flexibility or attachment, thus fostering new research and informing bereavement support practice.

Funding: This study was supported by funding from APPsyCI–Applied Psychology Research Center Capabilities and Inclusion (FCT/UIDB/05299/2020).

Acknowledgments: We would like to acknowledge the clinicians involved in the present research: Andrea Mendes, Cristina Pereira, Daniela Nogueira, Inês Minhós, Joana Semedo dos Santos, Mário Nascimento, and Tatiana Tavares.

Competing Interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics Statement: Ethics approval for the study was provided by the ethics committee of the ISPA – Instituto Universitário (Portugal), approval no. I-138-2-24. All participants provided written informed consent.

Trial Registration: Trial registered at ClinicalTrials.gov, Identifier: NCT06270381 (registered February 21, 2024).

Reporting Guidelines: The present report was written in accordance with the SPIRIT 2013 Statement (Standard Protocol Items: Recommendations for Interventional Trials).

Social Media Accounts: *David D. Neto:* LinkedIn, ResearchGate

References

American Psychiatric Association. (2022). *Diagnostic and statistical manual of mental disorders* (5th ed., text rev.). https://doi.org/10.1176/appi.books.9780890425787

Bond, F. W., Hayes, S. C., Baer, R. A., Carpenter, K. M., Guenole, N., Orcutt, H. K., Waltz, T., & Zettle, R. D. (2011). Preliminary psychometric properties of the Acceptance and Action Questionnaire–II: A revised measure of psychological inflexibility and experiential avoidance. *Behavior Therapy, 42*(4), 676–688. https://doi.org/10.1016/j.beth.2011.03.007

Carvalho, S., Pinto-Gouveia, J., Pimentel, P., Maia, D., & Mota-Pereira, J. (2011). Características psicométricas da versão portuguesa da escala multidimensional de suporte social percebido [Psychometric properties of the Portuguese version of the Multidimensional Perceived Social Support]. *Psychologica, 54*, 331–357. https://doi.org/10.14195/1647-8606_54_13

Carver, C. S. (1997). You want to measure coping but your protocol's too long: Consider the Brief COPE. *International Journal of Behavioral Medicine, 4*(1), 92–100. https://doi.org/10.1207/s15327558ijbm0401_6

Chan, A.-W., Tetzlaff, J. M., Gøtzsche, P. C., Altman, D. G., Mann, H., Berlin, J. A., Dickersin, K., Hróbjartsson, A., Schulz, K. F., Parulekar, W. R., Krleža-Jerić, K., Laupacis, A., & Moher, D. (2013). SPIRIT 2013 explanation and elaboration: Guidance for protocols of clinical trials. *BMJ, 346*, Article e7586. https://doi.org/10.1136/bmj.e7586

Coelho, A., Albuquerque, S., Neto, D. D., Silva, A. N., Resende, A., Nogueira, D., Tavares, T., & Minhós, I. (2024). *EMPOWER-Grief: Reforçar e mobilizar o potencial para o bem-estar e resiliência emocional prevenção no luto* [EMPOWER-Grief: Enhancing and Mobilizing the POtential for Wellness and Emotional Resilience in grief] (Technical report; ISPA – Instituto Universitário). http://hdl.handle.net/10400.12/9584

Coelho, A., Roberto, M., Barros, L., & Barbosa, A. (2022). Family caregiver grief and post-loss adjustment: A longitudinal cohort study. *Palliative & Supportive Care, 20*(3), 348–356. https://doi.org/10.1017/S147895152100095X

Davis, E. L., Deane, F. P., Lyons, G. C. B., & Barclay, G. D. (2017). Is higher acceptance associated with less anticipatory grief among patients in palliative care? *Journal of Pain and Symptom Management, 54*(1), 120–125. https://doi.org/10.1016/j.jpainsymman.2017.03.012

DGS – Direção Geral de Saúde. (2019). *Normas para a intervenção diferenciada no luto prolongado de adultos* [Guidelines for differential intervention in prolonged grief in adults]. https://normas.dgs.min-saude.pt/2019/04/23/modelo-de-intervencao-diferenciada-no-luto-prolongado-em-adultos/

Eisma, M. C., & Stroebe, M. S. (2021). Emotion regulatory strategies in complicated grief: A systematic review. *Behavior Therapy, 52*(1), 234–249. https://doi.org/10.1016/j.beth.2020.04.004

Fraley, R. C., Heffernan, M. E., Vicary, A. M., & Brumbaugh, C. C. (2011). The experiences in close relationships – Relationship Structures Questionnaire: A method for assessing attachment orientations across relationships. *Psychological Assessment, 23*(3), 615–625. https://doi.org/10.1037/a0022898

Guldin, M.-B., Vedsted, P., Zachariae, R., Olesen, F., & Jensen, A. B. (2012). Complicated grief and need for professional support in family caregivers of cancer patients in palliative care: A longitudinal cohort study. *Supportive Care in Cancer, 20*(8), 1679–1685. https://doi.org/10.1007/s00520-011-1260-3

Lichtenthal, W. G., Viola, M., Rogers, M., Roberts, K. E., Lief, L., Cox, C. E., Brewin, C. R., Xu, J. C., Maciejewski, P. K., Pan, C. X., Coats, T., Ouyang, D. J., Rabin, S., Vaughan, S. C., Breitbart, W., Marenberg, M. E., & Prigerson, H. G. (2022). Development and preliminary evaluation of EMPOWER for surrogate decision-makers of critically ill patients. *Palliative & Supportive Care, 20*(2), 167–177. https://doi.org/10.1017/S1478951521000626

Machado, P. P., & Horvath, A. (1999). Inventário da aliança terapêutica (WAI) [Working alliance inventory]. In M. R. Simões, M. M. Gonçalves, & L. S. Almeida (Eds.), *Testes e provas psicológicas em Portugal* [Psychological tests and instruments in Portugal] (Vol. 2, pp. 87-94). Braga, APPORT/SHO.

Ministry of Health. (2017). *Te Ara Whakapiri: Principles and guidance for the last days of life* (2nd ed). Wellington, New Zealand: Ministry of Health. https://www.tewhatuora.govt.nz/assets/Publications/Palliative/Te-Ara-Whakapiri/te-ara-whakapiri-principles-guidance-last-days-of-life-apr17.pdf

Moreira, H., Martins, T., Gouveia, M. J., & Canavarro, M. C. (2015). Assessing adult attachment across different contexts: Validation of the Portuguese version of the Experiences in Close Relationships–Relationship Structures questionnaire. *Journal of Personality Assessment, 97*(1), 22–30. https://doi.org/10.1080/00223891.2014.950377

Pais-Ribeiro, J., Silva, I., Ferreira, T., Martins, A., Meneses, R., & Baltar, M. (2007). Validation study of a Portuguese version of the Hospital Anxiety and Depression Scale. *Psychology, Health and Medicine, 12*(2), 225–237. https://doi.org/10.1080/13548500500524088

Pinto-Gouveia, J., Gregório, S., Dinis, A., & Xavier, A. (2012). Experiential avoidance in clinical and non-clinical samples: AAQ-II Portuguese version. *International Journal of Psychology & Psychological Therapy, 12*(2), 139–156. https://www.ijpsy.com/volumen12/num2/323/experiential-avoidance-in-clinical-and-non-EN.pdf

Prigerson, H. G., Boelen, P. A., Xu, J., Smith, K. V., & Maciejewski, P. K. (2021). Validation of the new DSM-5-TR criteria for prolonged grief disorder and the PG-13-Revised (PG-13-R) scale. *World Psychiatry, 20*(1), 96–106. https://doi.org/10.1002/wps.20823

Ribeiro, J. L. P., & Rodrigues, A. P. (2004). Questões acerca do coping: A propósito do estudo de adaptação do Brief COPE [Questions about coping: Regarding the adaptation of the Brief COPE]. *Psicologia, Saúde & Doenças, 1*, 3–15. https://www.redalyc.org/articulo.oa?id=36250101

Shear, M. K., Muldberg, S., & Periyakoil, V. (2017). Supporting patients who are bereaved. *British Medical Journal, 358*, Article j2854. https://doi.org/10.1136/bmj.j2854

Silva, S., Moreira, H., Pinto, S., & Canavarro, M. C. (2009). Cancro da mama e desenvolvimento pessoal e relacional: Estudo das características psicométricas do Inventário de Desenvolvimento Pós-Traumático numa amostra de mulheres da população portuguesa [Breast cancer and

personal and relational development: Study of the psychometric characteristics of the Post-Traumatic Development Inventory in a sample of women from the Portuguese population]. *Revista Iberoamericana de Diagnóstico e Avaliação Psicológica, 28*(2), 105–133. https://www.redalyc.org/pdf/4596/459645444007.pdf

Tedeschi, R. G., & Calhoun, L. G. (1996). The posttraumatic growth inventory: Measuring the positive legacy of trauma. *Journal of Traumatic Stress, 9*(3), 455–471. https://doi.org/10.1002/jts.2490090305

Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment, 1*(3), 207–210. https://doi.org/10.1037/1040-3590.1.3.207

Winston, A., & Lujack, A. M. (2015). Supportive psychotherapy. In A. Tasman, J. Kay, J. A. Lieberman, M. B. First, & M. B. Riba (Eds.), *Psychiatry* (pp. 1871-1881). Wiley. https://doi.org/10.1002/9781118753378.ch92

World Health Organization. (2019). *International statistical classification of diseases and related health problems* (11th ed.). https://icd.who.int/

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica, 67*(6), 361–370. https://doi.org/10.1111/j.1600-0447.1983.tb09716.x

Zimet, G. D., Dahlem, N. W., Zimet, S. G., & Farley, G. K. (1988). The Multidimensional Scale of Perceived Social Support. *Journal of Personality Assessment, 52*(1), 30–41. https://doi.org/10.1207/s15327752jpa5201_2

Clinical Psychology in Europe

PsychOpen GOLD

Letter to the Editor, Commentary

Check for updates

Can a 1-Item Scale for Psychotherapy Outcomes Be Psychometrically Robust?

Scott T. Meier [1]

[1] *Counseling, School and Educational Psychology, University at Buffalo, Buffalo, NY, USA.*

Clinical Psychology in Europe, 2025, Vol. 7(1), Article e15207, https://doi.org/10.32872/cpe.15207

Published (VoR): 2025-02-28

Corresponding Author: Scott T. Meier, 80 Londonderry Ln, Getzville, NY 14068, USA. Phone: +1 716 906-9420. E-mail: stmeier@buffalo.edu

Gonçalves et al. (2024) recently described the selection of a 1-item outcome scale for the European Psychotherapy Consortium. The field has been trending toward brief scales because of research indicating greater patient compliance with fewer items (Miller et al., 2003; Miller et al., 2005). From a psychometric perspective, however, the 1-item emotional and psychological outcomes (EPO-1) measure is likely to produce low reliability and validity estimates. This results from measurement principles indicating that (a) reliability estimates increase with the number of items, and (b) validity estimates depend upon reliability. Measurement error decreases with an increasing number of item responses because random error sources tend to balance or cancel (Meier, 2013). If a patient misunderstands a question, for example, this error becomes a major influence on data in a single item self-report. Because multiple factors typically influence responses to any psychological item, scores on 1-item scales are less likely to be sensitive to change resulting from psychotherapy than a multi-item scale that aggregates change-relevant variance (Meier, 1997).

Other research suggests that many patients will not interpret the EPO-1 as test developers intended (Schwarz, 1999). Labeling this problem as *intracategory variability*, Dohrenwend (2006) observed that test-takers respond to item content on a self-report measure based on a wide range of personal experiences. When asked to report on a recent serious illness, for example, respondents will describe episodes that vary from simple flu to heart attacks. As a result, the basis on which individuals respond to health-related categories on self-report measures can range "from *the catastrophic* to *the trivial*" (Dohrenwend, 2006, p. 479). The EPO-1's content is "At this moment, how well do you feel you are getting along emotionally and psychologically?" Patients respond on a 5-point scale ranging from 0 ("Very poorly; I can barely manage to deal with things") to 4

("Very well; I have no important complaints"). For many individuals, these are cognitively complex tasks likely to lead to heterogeneous response processes and ratings.

Given that single item measures are inappropriate with ambiguous constructs (Allen et al., 2022), future research should evaluate reliability and validity estimates for the EPO-1. At a minimum, EPO-1 scores should evidence (a) stability over time in the absence of any intervention, (b) change over time when the patient participates in a psychosocial intervention, and (c) moderate to high correlations with existing measures of outcome. If EPO-1 scores fail to meet these standards, possible next steps include (a) augmenting EPO-1 data with one or more item(s) related to common factors that have been shown to influence outcome and (b) developing a system that minimizes respondent burden. Regarding (a), working alliance would appear to be a strong candidate given that psychotherapy researchers consistently find a modest positive effect of the client/therapist alliance on outcomes (Flückiger et al., 2018). Regarding (b), recent studies suggest that AI could produce outcome information through analysis of text produced by client discourse recorded during therapy sessions as well as clinicians' unstructured progress notes (Chu et al., 2024).

Funding: The author has no funding to report.

Acknowledgments: The author has no additional (i.e., non-financial) support to report.

Competing Interests: The author has declared that no competing interests exist.

References

Allen, M. S., Iliescu, D., & Greiff, S. (2022). Single item measures in psychological science: A call to action [Editorial]. *European Journal of Psychological Assessment, 38*(1), 1–5. https://doi.org/10.1027/1015-5759/a000699

Chu, C., Sun, T., Zhang, B., & Rounds, R. (2024). *Assessing vocational interests through chat: Development and validation of the Career Guidance Chatbot (CGC-bot)* [Unpublished manuscript]. University of Illinois at Urbana-Champaign.

Dohrenwend, B. P. (2006). Inventorying stressful life events as risk factors for psychopathology: Toward resolution of the problem of intracategory variability. *Psychological Bulletin, 132*(3), 477–495. https://doi.org/10.1037/0033-2909.132.3.477

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy, 55*(4), 316–340. https://doi.org/10.1037/pst0000172

Gonçalves, M. M., Lutz, W., Schwartz, B., Oliveira, J. T., Saarni, S. E., Tishby, O., Rubel, J. A., Boehnke, J. R., Montesano, A., Paiva, D., Ceridono, D., Zech, E., Willemsen, J., Saarni, S. I., Kompan Erzar, K., Janeiro, L., Gelo, O. C. G., Errázuriz, P., Holas, P., . . . Barkham, M. (2024).

Developing a European Psychotherapy Consortium (EPoC): Towards adopting a single-item self-report outcome measure across European countries. *Clinical Psychology in Europe, 6*(3), Article e13827. https://doi.org/10.32872/cpe.13827

Meier, S. (1997). Nomothetic item selection rules for tests of psychological interventions. *Psychotherapy Research, 7*(4), 419–427. https://doi.org/10.1080/10503309712331332113

Meier, S. T. (2013). *Measuring change in counseling and psychotherapy*. Guilford Press.

Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2003). The Outcome Rating Scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2*, 91–100. https://www.researchgate.net/publication/242159752

Miller, S. D., Duncan, B. L., Sorrell, R., & Brown, G. S. (2005). The Partners for Change Outcome Management System. *Journal of Clinical Psychology, 61*(2), 199–208. https://doi.org/10.1002/jclp.20111

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *The American Psychologist, 54*(2), 93–105. https://doi.org/10.1037/0003-066X.54.2.93

Clinical Psychology in Europe

PsychOpen GOLD

Letter to the Editor, Commentary

Response to the Commentary „Can a 1-Item Scale for Psychotherapy Outcomes Be Psychometrically Robust?"

Brian Schwartz[1], Miguel M. Gonçalves[2], Wolfgang Lutz[1], João Tiago Oliveira[2], Suoma E. Saarni[3,4,5], Orya Tishby[6], Michael Barkham[7], European Psychotherapy Consortium (EPoC)

[1] *Department of Psychology, Trier University, Trier, Germany.* **[2]** *CIPsi – Psychology Research Center, School of Psychology, University of Minho, Braga, Portugal.* **[3]** *Department of Psychiatry, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland.* **[4]** *Department of Psychiatry, Helsinki University Hospital HUS and Helsinki University, Helsinki, Finland.* **[5]** *Department of Psychiatry, Wellbeing Services County of Päijät-Häme, Lahti, Finland.* **[6]** *Department of Psychology, Hebrew University, Jerusalem, Israel.* **[7]** *Clinical and Applied Psychology Unit, School of Psychology, University of Sheffield, Sheffield, United Kingdom.*

Clinical Psychology in Europe, 2025, Vol. 7(1), Article e16921, https://doi.org/10.32872/cpe.16921

Published (VoR): 2025-02-28

Corresponding Author: Brian Schwartz, Clinical Psychology and Psychotherapy, Department of Psychology, Trier University, D-54286 Trier, Germany. Phone: +49 651 201 4321. E-mail: schwartzb@uni-trier.de

We thank the author(s) for their commentary in response to our article on the European Psychotherapy Consortium (EPoC) of the European Chapter of the Society for Psychotherapy Research (SPR) published in *Clinical Psychology in Europe* (Gonçalves et al., 2024) and welcome the opportunity to respond. Whilst we appreciate the detailed comments on our article made by the author(s), we would like to take this opportunity to primarily address the specific criticisms of the EPO-1 single-item scale.

Overall, the commentary criticizes our article on a combined theoretical and psychometric basis as a first, but so far, successful, attempt to coordinate the administration of patient outcome data collected during the course of treatment and implemented across Europe. Such criticism fails to recognize the practical impact of our strategic approach, which aims to significantly advance the paradigm of patient-focused research. Our objective has been to advance the field of psychological therapies and pave the way for the first steps towards coordinated data collection, data-based quality assurance, and practice-based evidence into the therapeutic process across national borders. The single-item Emotional and Psychological Outcome-1 measure (EPO-1) addresses the reality of identifying a common measurement language across different clinics in different

European countries with varying structures of assessment and other healthcare systems, some with already existing measurement systems and some with none. Hence, we are attempting what we consider to be a unique program of research implementation that is both feasible in most clinical settings (i.e., minimal demands on patients, therapists, and healthcare systems) and yet has a grand vision in transcending national boundaries that are, so often, limitations to research collaboration. Against this background, the commentary fails to fully appreciate the objectives and value of the EPO-1 item's introduction.

In addition to presenting our general view of the project, which is broader than that of the commentary's author(s), we would also like to respond to some specific points of criticism below.

The author(s) assumes that the EPO-1 item has limited reliability and validity. In response, we direct the author(s) to the high correlations of the instrument with other established outcome instruments, such as the Outcome Questionnaire-30 (OQ-30; r = .601), the Questionnaire for the Evaluation of Psychotherapy (FEP-2; r = .626), and the Patient Health Questionnaire-9 (PHQ-9; r = .630), which can be found in detail in Chapter 4 (Appendix) of *Bergin and Garfield's handbook of psychotherapy and behavior change* (Lutz et al., 2021). The empirical data (n = 521) also show that the pre-post effect sizes, measured with the EPO-1, are as strongly related to the above instruments and that the individual effect sizes are comparable to those of the other measures. The EPO-1 pre-post-effect sizes refute the assumption that the single-item measure EPO-1 is less sensitive to change than multi-item scales (pre-post-effect sizes, e.g., EPO-1(Likert): d = 1.086; EPO-1(analogue): d = 1.469; BSI: d = 0.879; OQ-30: d = 1.320; Lutz et al., 2021).

For future research, the author(s) recommends demonstrating the stability of the item over time without intervention, its change in response to intervention, and its correlations with established measures. The last two points (change sensitivity and convergent validity) have already been empirically demonstrated (see above; Lutz et al., 2021), leaving only the first point (reliability in the absence of intervention) to be addressed in future studies. Many points of criticism in the commentary refer to the general disadvantages of self-report questionnaires. They are not specifically related to single-item scales (e.g., the varying interpretation of questions by individual patients or the cognitive complexity of evaluating the item).

Further, the author(s) suggests extending the EPO-1 item with an alliance item and using large language models (LLMs), which are important and topical issues in recent psychotherapy research. However, the EPO-1 is intentionally designed as a *single-item measure* to ensure easy implementation in clinical practice. Moreover, this outcome measure assesses psychological well-being, not therapeutic alliance. While LLMs have become valuable tools in psychotherapy research, the EPO-1 item is a low-burden measure for patients, collected via self-report. Therefore, it should not be replaced by video or text analyses, which could capture a different perspective.

Our response addresses the concerns expressed based on theoretical considerations with empirical evidence. Furthermore, we would argue that the benefits and potential of these efforts to introduce a standardized outcome measure across Europe outweigh the theoretical (and, as demonstrated, not necessarily valid) criticisms. We are aware of the common problems of single-item scales, which is why the EPoC does not only focus on the EPO-1 item but also on developing and implementing crosswalks to create a standard measure structure between different clinics across Europe.

In summary, EPoC is a project that aims to evolve and create large, heterogeneous data sets from different countries that will facilitate practice-based evidence and data-informed psychological therapy. EPO-1 is currently translated into 13 languages, and two more are expected to be added soon. We hope that more institutions will join our initiative in the future and adopt the item in their assessments. Ultimately, whether it is taken up in the field is an empirical question.

Funding: This work was supported by the German Research Foundation (DFG) under project numbers 493169211, 504507043, and 525286173 granted to Wolfgang Lutz and partially conducted at the Psychology Research Centre (PSI/01662), School of Psychology, University of Minho, and supported by the Portuguese Foundation for Science and Technology and the Portuguese Ministry of Science, Technology and Higher Education (UID/PSI/01662/2019), through national funds (PIDDAC).

Acknowledgments: The authors have no additional (i.e., non-financial) support to report.

Competing Interests: Michael Barkham declares that he is a co-developer of the CORE-OM and CORE-10.

Social Media Accounts: *Brian Schwartz:* Bluesky: @schwartz-psyres.bsky.social

References

Gonçalves, M. M., Lutz, W., Schwartz, B., Oliveira, J. T., Saarni, S. E., Tishby, O., Rubel, J. A., Boehnke, J. R., Montesano, A., Paiva, D., Ceridono, D., Zech, E., Willemsen, J., Saarni, S. I., Kompan Erzar, K., Janeiro, L., Gelo, O. C. G., Errázuriz, P., Holas, P., … Barkham, M. (2024). Developing a European Psychotherapy Consortium (EPoC): Towards adopting a single-item self-report outcome measure across European countries. *Clinical Psychology in Europe, 6*(3), Article 13827. https://doi.org/10.32872/cpe.13827

Lutz, W., De Jong, K., Rubel, J. A., & Delgadillo, J. (2021). Measuring, predicting and tracking change in psychotherapy. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (7th ed., pp. 89–133). Wiley.

EACLIPT

Clinical Psychology in Europe (CPE) is the official journal of the European Association of Clinical Psychology and Psychological Treatment (EACLIPT).

leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.