AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.
Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication
Please cite the final published version:

Lundorff, M., Holmgren, H., Zachariae, R., Farver-Vestergaard, I., & O'Connor, M. (2017). Prevalence of prolonged grief disorder in adult bereavement: A systematic review and meta-analysis. *Journal of Affective Disorders*, *212*, 138-149. https://doi.org/10.1016/j.jad.2017.01.030

Publication metadata

Author's Accepted Manuscript

Prevalence of prolonged grief disorder in adult bereavement: A systematic review and meta-analysis

Marie Lundorff, Helle Holmgren, Robert Zachariae, Ingeborg Farver-Vestergaard, Maja O'Connor
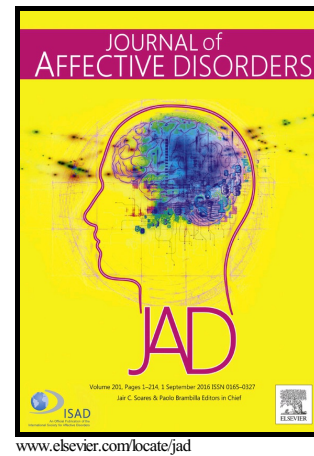
Cite this article as: Marie Lundorff, Helle Holmgren, Robert Zachariae, Ingeborg Farver-Vestergaard and Maja O'Connor, Prevalence of prolonged grief disorder in adult bereavement: A systematic review and meta-analysis, *Journal of Affective Disorders,* http://dx.doi.org/10.1016/j.jad.2017.01.030

Prevalence of prolonged grief disorder in adult bereavement: A systematic review and meta-analysis

Marie Lundorff[1], Helle Holmgren[1], Robert Zachariae[2], Ingeborg Farver-Vestergaard[2], & Maja O'Connor[1]

[1]Research Unit for Natural and Complicated Grief, Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus, Denmark

[2]Unit for Psycho-Oncology and Health Psychology, Department of Oncology, Aarhus University Hospital and Department of Psychology and Behavioural Sciences, Aarhus University, Aarhus, Denmark

Corresponding author: Marie Lundorff, MSc, PhD student.

Research Unit for Natural and Complicated Grief, Department of Psychology and Behavioural Sciences,

Aarhus University, Bartholins Allé 9, Bld. 1340, DK–8000 Aarhus C (Denmark)

E-mail: mlund@psy.au.dk

1. Introduction

The psychological and physiological reactions that follow the loss of a loved one are collectively known as grief. Most individuals go through a painful, but natural, grieving process where the intensity of grief-related distress typically decreases gradually over time (Jordan and Litz, 2014). Thus, although bereavement can be a highly stressful and significant life experience, most individuals have sufficient internal resources and external support to adequately cope with their grief and slowly readjust to a life without the deceased (Prigerson et al., 2009; Zisook & Shear, 2009). However, research has also shown that for a significant minority of bereaved individuals the grieving process is particularly complicated (e.g., Lichtenthal et al., 2011; Prigerson et al., 2009; Simon et al., 2007). Instead of a decreasing intensity of grief-related distress, these individuals experience severe grief reactions that become abnormally persistent and increasingly debilitating across time (Jordan and Litz, 2014; Maercker et al., 2013; Prigerson et al., 2009).

Prolonged grief disorder (PGD) is a proposed diagnostic category intended to classify bereaved individuals who experience notable dysfunction for atypically long periods of time following a significant loss (Prigerson et al., 2009). Core symptoms include a pervasive yearning for the deceased or persistent preoccupation with the deceased accompanied by intense emotional pain (World Health Organization, 2016). Furthermore, PGD is characterized by difficulties in engaging in social or enjoyable activities, a reduced ability to experience positive mood, and difficulties accepting the death of the loved one (World Health Organization, 2016). A duration criteria of six months is proposed to ensure that natural grief reactions in the acute state following bereavement are not confounded with the syndrome of PGD (Prigerson et al., 2009). Research have found symptoms of PGD to be associated with impairment of the bereaved person's familial, social, and occupational functioning to a similar extent as found for other mental disorders, e.g., depression and post-traumatic stress disorder (Jordan and Litz, 2014; Maercker et al., 2013; Prigerson et al., 2009; Shah and Meeks, 2012).

The identification of PGD fostered the issue of whether to include the condition as an official mental disorder in the diagnostic manuals. In relation to recent revisions of both the *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.; *DSM-5*; American Psychiatric Association, 2013) and the *International Classification of Diseases* (11th ed.; *ICD-11*; World Health Organization, 2016), working groups were established to investigate the validity, specificity, and treatability of PGD. While the group editing *DSM-5* initially

embraced the possibility of including PGD as a mental disorder, the proposal was finally rejected in 2013 (Bryant, 2014; Rosner, 2015). However, a code diagnostically corresponding to prolonged grief problems – Persistent Complex Bereavement Disorder (PCBD) – was added in the Appendix as a candidate disorder demanding further study (American Psychiatric Association, 2013; Maciejewski et al., 2016). Currently, PGD remains proposed for inclusion in the forthcoming *ICD-11*, scheduled for release in 2018 (World Health Organization, 2016). Inclusion of the diagnosis will initiate the application of the diagnosis in healthcare settings. However, a number of unanswered questions remain regarding the impact of PGD, including the prevalence of the disorder in the general population. To date, systematic reviews of PGD have addressed its predictors, its prevention, and the effect of interventions to reduce grief-related symptoms (e.g., Lobb et al., 2010; Rosner et al., 2010; Wittouck et al., 2011). So far, only a limited number of epidemiological studies have assessed the prevalence of PGD in general population samples, and, to the best of our knowledge, no systematic reviews combining and comparing the existing individual studies of PGD prevalence have yet been published.

With no clear-cut data regarding the prevalence of PGD, much published theory and research use rather arbitrary expressions, such as "a significant minority" (e.g., Jordan and Litz, 2014; Zisook et al., 2010), to describe the number of individuals experiencing severe complications following bereavement. However, healthcare services are in need of more precise estimates of the prevalence of PGD in the general population. Preferably, this information should be provided before the diagnosis is introduced to allocate economic and professional resources most effectively. We therefore conducted a systematic review and meta-analysis with the aim of providing an estimate of the prevalence of PGD in the general adult bereaved population.

2. Methods

The present study was protocol-based and conducted in accordance with recommendations from the Cochrane Collaboration (Higgins and Green, 2011) and the guidelines for Meta-Analysis Of Observational Studies in Epidemiology (MOOSE) (Stroup et al., 2000). Prior to the review, the protocol was submitted to PROSPERO – an international prospective register for review protocols – in May 2016 (Lundorff et al., 2016; registration number: CRD42016038416). The methods and results are documented according to the Preferred

Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Liberati et al., 2009).

2.1. Data sources and search strategy

A systematic keywords-based search was undertaken using the bibliographic databases of PubMed, PsycINFO, Embase, Web of Science, and CINAHL. The search string consisted of a combination of subject headings (MeSH terms) and free text (keywords) to reflect indexing differences between databases. Informed by the PICO approach (Sackett et al., 1997), keywords related to population ("prolonged grief disorder" OR "prolonged grief" OR "complicated grief" OR "traumatic grief" OR "pathological grief" OR "persistent complex bereavement-related disorder") were combined with words related to indicator (bereave* [MeSH] OR grief* [MeSH] OR grieving OR mourn* [MeSH] NOT "psychological trauma" [MeSH] NOT "violent death*") and outcome (prevalence [MeSH] OR frequency OR incidence [MeSH] OR proportion). Two authors independently performed the search throughout the period of May 2016.

2.2. Selection criteria

The review included epidemiological studies of individuals aged 18 years or older who suffered the loss of a loved one through mainly non-violent death causes. The outcome variable was prolonged grief, as assessed with a standardized, validated psychometric instrument. Studies had to discriminate between individuals with natural (below a stated cut-off) and prolonged grief reactions (above a stated cut-off) and provide a prevalence estimate of PGD in the sample. In alignment with the duration criteria for PGD, all included studies had to assess grief with a mean time of at least 6 months post-loss. The literature search was limited to papers published in English-language peer-reviewed journals. No limitation of publication date was applied.

Intervention and case studies were excluded, together with studies conducted on infants, children, adolescents (sample mean age <18 years), healthcare professionals, individuals with personality disorders, and psychiatric in-/outpatient samples. Studies exclusively investigating deaths by suicide, murder, natural disasters, or terrorist attacks were also excluded due to additional dimensions associated with such losses, e.g., severe comorbidity and risk of first hand exposure (Brent et al., 2009; Dyregrov et al., 2015; Piper et al., 2001; Tal Young

4

et al., 2012). Finally, studies focusing on other forms of loss unrelated to bereavement (e.g., divorce, major illness, migration, or unemployment) were excluded.

2.3. Screening procedures

Sources identified in the literature search were independently assessed for eligibility by two authors through a two-step process. The first screening was performed based on titles and abstracts, while full-text papers were retrieved and assessed in the second screening. After each screening, the two authors compared their results, discussed discrepancies, and negotiated a decision. Following each screening, interrater reliability was assessed using Cohen's Kappa statistic, κ (McHugh, 2012). Study screening was conducted using the online software Covidence (Veritas Health Innovation Ltd., 2016).

2.4. Risk of bias assessment

The included studies were systematically assessed for possible risk of bias (RoB) with a tool specifically developed for population-based prevalence studies (Hoy et al., 2012). This tool consists of ten items, including four items addressing external validity (i.e., representativeness of sample) and six items addressing internal validity (i.e., measurement reliability). The items are formulated as dichotomous "yes-no" questions, for example "Was the same mode of data collection used for all subjects?". Each item is scored with a value of 1 (yes) or 0 (no), yielding a RoB total score for each study with higher scores indicating lower RoB and higher methodological validity. In accordance with the guidelines, when a paper provided insufficient information to permit a judgment for a particular item, this item was rated high RoB (0 points) (Hoy et al., 2012). Studies with scores of 9 or 10 points (i.e., 9 or 10 'yes'-answers) were considered to have low RoB; studies with scores of 7 or 8 points were considered to have moderate RoB; and studies with scores of 6 or less points were considered to have high RoB.

2.5. Meta-analytical strategy

An inverse-variance weighted event rate (ER), was calculated for each study as summary statistic of prevalence. ER is defined as the proportion indicating how often the "event" of PGD occurs while taking the individual study's sample size into consideration (i.e., number of individuals with PGD, divided by the sample size), yielding an ER-range from 0 to 1. Comprehensive Meta-Analysis version 3 (CMA, Biostat, 2016) was used to convert raw data

into ERs and to calculate pooled inverse variance weighted ERs for each study or sub-group. For longitudinal studies with multiple assessment time points, the ER at the time closest to six months post-loss was included in the meta-analysis. A random effects model was chosen over a fixed effects model as ERs could be assumed to vary according to underlying study characteristics, e.g., between study populations of different ages, with different types of losses, and different times since loss. Publication bias, a widespread problem when conducting meta-analysis, was inspected visually with funnel plots (Higgins and Green, 2011) and analyzed with Egger's method (Egger et al., 1997). If our analyses indicated possible publication bias, the trim-and-fill procedure (Duval and Tweedie, 2000) was used to impute "missing studies" and calculate an adjusted ER accordingly.

Possible heterogeneity was assessed with the Cochran Q statistics for each analysis (Higgins et al., 2003), which is a Chi-squared statistic with higher values indicating larger heterogeneity of the prevalence rates. Since heterogeneity tests are often characterized by limited statistical power, a p-value ≤ 0.10 would indicate significant heterogeneity (Higgins and Green, 2011). To quantify the degree of heterogeneity, the I^2 was calculated to describe the percentage of the variability in prevalence estimates that is due to heterogeneity rather than sampling error (chance) (Higgins and Green, 2011).

Based on the prevalence estimate from the meta-analysis, we calculated two interval estimates. In addition to the confidence interval (i.e., precision of the estimate), we identified the prediction interval (i.e., how the true effects are distributed around the summary effect) based on *Tau* as an estimate of the standard deviation of the true effect sizes ($\mu \pm Z^{\alpha} \sqrt{\tau^2}$) (Borenstein et al., 2011).

If more than three studies reported relevant data, we conducted between-group analyses of ERs according to study design (cross-sectional vs. longitudinal), sampling method (probability vs. non-probability), assessor (self- vs. other-report), assessment instrument, level of RoB (high vs. moderate vs. low), and geographical study location (Western vs. Eastern). Finally, we conducted exploratory meta-regression analyses to investigate the role of potential continuous and categorical moderators (mean age of the sample, time since loss, percent women, response rate, and types of losses).

3. Results

3.1. Study selection and screening

The initial search for relevant papers yielded 686 records (191 in Pubmed; 159 in PsycINFO; 208 in Web of Science; 91 in Embase; 37 in CINAHL). Prior to the first screening, duplicates (n=299) were removed leaving 387 unique records for screening. Based on titles and abstracts, 252 records were judged irrelevant by two independent researchers (see Figure 1). The researchers initially disagreed on 18 (4.7%) papers yielding an interrater reliability of κ=0.89 (95% CI (confidence interval) 0.854-0.944), p<0.001. Applying common benchmarks for Cohen's Kappa on observer agreement for categorical data, this corresponds to a strong level of agreement (McHugh, 2012).

Full-text papers were obtained for the remaining 135 records and subjected to a second screening. After applying the a priori determined eligibility criteria to the full-text papers, an additional 106 records were excluded due to the following reasons: not a prevalence study (n=24); bereavement as result of violent death (n=23); not an empirical research paper, e.g., review or commentary (n=20); using non-relevant sample (n=12); no relevant grief outcome (n=11); not available in English (n=6), or irrelevant to the topic (n=1). Further nine studies were excluded due to use of results also analyzed in one of the remaining included studies. In this second round of screening, the researchers initially disagreed on 17 (12.6%) articles. The agreement between the two raters in this second round of screening was moderate (κ=0.70 (95% CI 0.569-0.831), p<.001) (McHugh, 2012). Keeping the numerous instruments to measure grief in mind, the disagreements were most often a result of different initial assessment of whether the studies differentiated between natural levels of grief and PGD symptomatic levels. After the second and final screening, a total of 29 full-text papers met the a priori stated inclusion criteria of this systematic review. The study selection process is described in Figure 1.

(Insert Figure 1 near here)

3.1.1. Post-hoc screening

Further examination of the 29 included studies identified through the two planned screening rounds revealed a highly diverse sample of studies differing in particular with respect to the sampling methods used. While some studies used population-based "stratified

7

two-stage random sampling" (e.g., Fujisawa et al., 2010; Kersting et al., 2011; Miyajima et al., 2014; Newson et al., 2011), a large proportion of the remaining studies used various types of convenience sampling, e.g., by recruiting family members of intensive care unit patients (Anderson et al., 2008; Kentish-Barnes et al., 2015; Siegel et al., 2008) or caregivers of terminally ill cancer patients (Chiu et al., 2011; McCarthy et al., 2010). Convenience samples tend to be unrepresentative of the population studied as a whole and are, furthermore, unlikely to satisfy the mathematical assumptions underlying many statistical analyses (Coolican, 2014). To maximize generalizability of our findings, prior to data extraction, we decided on excluding studies that investigated a particular population recruited via convenience sampling (e.g., cancer caregivers from one specific hospital) in a third post hoc screening.

Two authors independently rated the remaining studies and excluded an additional 19 studies using convenience sampling. Finally, following recommendations from the Cochrane Collaboration (Higgins and Green, 2011), we attempted to identify so-called grey literature by examining the reference lists of the included articles, which resulted in four additional records. This process resulted in a final total of 14 included studies (see Figure 1).

3.2. Study characteristics

Participant characteristics and study designs of the 14 unique studies included are summarized in Table 1. The included studies were published between 1994 and 2016 and were broadly geographically distributed with five studies carried out in Asian countries, three in Europe, two studies in Australia, and four in the United States. Nine of the studies employed a cross-sectional design, while the remaining five studies were longitudinal. Ten studies used surveys (self-report) to assess levels of grief reactions, while the remaining four studies were interview-based (other-report). Various recruitment procedures were employed in the included studies. Five studies used cluster sampling, another two studies used voluntary response sampling, and seven studies used random sampling methods. PGD was assessed with six different instruments: Bereavement Phenomenology Questionnaire (BPQ) (K=1), Core Bereavement Items (CBI) (K=1), Brief Grief Questionnaire (BGQ) (K=3), Inventory of Complicated Grief (ICG) (K=3), Inventory of Complicated Grief-Revised (ICG-R) (K=3), and Prolonged Grief-13 (PG-13) (K=3). All of these instruments were standardized and tested for reliability with Cronbach's alphas ranging from 0.82 to 0.94 (Burnett et al., 1997;

Byrne and Raphael, 1994; Prigerson et al., 2009, 1995b; Prigerson and Jacobs, 2001; Shear et al., 2006).

Sample sizes ranged from 57 (Byrne and Raphael, 1994) to 1402 (Kersting et al., 2011) (mean *N*=573.9, *SD*=439.3) yielding a total of 8035 participants in the included studies. Regarding gender, women comprised the majority of study participants, as an average of 62.1% was female. Single study mean age ranged from 21 (Varga et al., 2015) to 74.5 years (Byrne and Raphael, 1994) and the overall mean age was 52.2 years (*SD*=15.2). Participants' relation to the deceased varied with the most common types of losses being spousal (34.72%) and parental (27.65%) bereavement. The remaining participants had lost a grandparent (6.85%), a child (5.18%), a sibling (5.08%), a parent-in-law (4.00%), a friend (2.62%), or other loved-ones (13.88%).

(Insert Table 1 near here)

3.3. Risk of bias assessment

RoB is reported in Table 2. Four studies had a total score of 9 or 10 points (Kersting et al., 2011; Miyajima et al., 2014; Mizuno et al., 2012; Newson et al., 2011) which corresponds to a low RoB. Six studies scored 7 or 8 points (Fujisawa et al., 2010; He et al., 2014; Kim et al., 2015; Li and Prigerson, 2016; O'Connor et al., 2010; Prigerson et al., 2009), which corresponds to a moderate RoB. The remaining four studies scored 6 or less points (Byrne and Raphael, 1994; Goldsmith et al., 2008; Middleton et al., 1996; Varga et al., 2015) and were consequently labelled as high RoB. The studies' RoB scores ranged from 5 to 10 out of 10 possible points (*M*=7.5, *SD*=1.5). The assessment of bias was performed independently by two raters with an interrater reliability at a moderate level, and a Kappa value of 0.62 (95% CI 0.471-0.765), *p*<.001 (McHugh, 2012).

Additionally, RoB across the included studies was explored by dividing the RoB scale into the two categories addressing external (Items 1-4) and internal validity (Items 5-10), respectively. Table 2 displays how the studies rated low (marked grey) or high (not marked) RoB for each individual scale item. As illustrated in the table, the majority of studies show low external validity (high RoB scores on items 1-4). For example, only half of studies used some form of random sampling (item 3, 50.0%) and had high likelihood of non-response bias (item 4, 85.7%). Conversely, the majority of the included studies show high internal validity (low RoB scores on items 5-10). For example, all studies collected data directly

from the bereaved individuals as opposed to via a proxy (item 5, 100%), and most of the studies used a well-validated tool to measure PGD (item 7, 86%).

(Insert Table 2 near here)

3.4. Meta-analysis

Eight thousand and thirty-five bereaved individuals were included in the analysis and the pooled prevalence of PGD was 9.8% (95% CI 6.8-14.0; 95% PI (prediction interval) 2.0-36.3). The results are shown in Figure 2. The funnel plot for the main analysis appeared asymmetrical and Egger's test (p=0.011) indicated possible publication bias. Imputing missing studies with the trim-and-fill method resulted in an adjusted prevalence of 11.0%. Furthermore, analyses revealed the ERs to be highly and statistically significantly heterogeneous (I^2=96.9%, Q=416.1, p=0.001) (see Table 3).

(Insert Figure 2 near here)

3.4.1. Sub-group meta-analyses

Possible sources of heterogeneity were explored through a number of sub-group analyses (Table 3). The pooled ER of PGD tended to be smaller when studies used cross-sectional designs (9.6% compared to 10.4% for longitudinal designs), probability sampling methods (9.5% compared to 12.4% for non-probability sampling), and outcome assessment via other-reporting (9.6% compared to 9.8% among studies using self-report). When comparing studies conducted in Western countries (i.e., Australia, Denmark, Germany, the Netherlands, and the United States) and Eastern countries (i.e., China and Japan), the pooled ER of PGD was higher in Western studies (10.1%) than in Eastern studies (9.2%). The pooled prevalence of PGD in studies with a low risk of bias was higher (10.8%) compared to both studies with a moderate (9.8%) and a high (9.2%) risk of bias. However, none of these between-group analyses reached statistical significance.

Analysis of the instruments used to assess PGD revealed statistically significant between-groups differences ($p \leq 0.001$). The pooled prevalence of PGD was highest in the studies that had used ICG (20.5%), followed by studies using BGQ (12.3%), ICG-R (9.2%), and studies using PG-13 (3.2%). As seen in Table 3, heterogeneity analyses within each instrument continued to show significantly high levels.

(Insert Table 3 near here)

3.4.2. Moderators of PGD prevalence

Our meta-regression analysis indicated that a larger percentage of siblings lost in the study samples was associated with statistically significantly higher ERs (p=0.003; ß=0.15). Furthermore, study sample mean age emerged as a statistically near-significant (p=0.075) moderator with a positive regression slope (ß=0.03) indicating that older age was associated with higher ER (Table 4).

(Insert Table 4 near here)

4. Discussion

Distress following bereavement is a common reaction to the loss, and for a long time grief has been considered a natural human response. However, in some cases the grief reaction becomes persistent and interferes dramatically with the individual's function of life (Maercker and Lalor, 2012). This observation led to the operationalization of a grief-specific diagnosis – prolonged grief disorder (PGD). Although recent studies have investigated the prevalence of PGD, the disorder is often described anecdotally as a condition expected to affect "a considerable minority" of bereaved individuals (e.g., Boelen and Prigerson, 2007; Jordan and Litz, 2014; Prigerson et al., 1995a; Zisook et al., 2010). To identify a more precise definition of this "considerable minority" we conducted the, to the best of our knowledge, first systematic review and meta-analysis combining and comparing the results from a number of single studies investigating the prevalence of PGD in non-violent adult bereavement.

The pooled prevalence of PGD across the fourteen included studies was 9.8% and 11.0% when adjusted for potential publication bias. This result supports the assumption that the proposed disorder does in fact affect a significant minority of all bereaved individuals, more specifically one out of ten adults who experience a non-violent loss of a loved one. A number of meta-regression analyses were conducted to investigate whether different sample characteristics have an effect on the prevalence of PGD. As shown in Table 4, these analyses revealed an age-effect, indicating that older samples were associated with higher ERs. Similarly, a larger percentage of siblings lost in a sample was associated with significantly higher

prevalence even when adjusting for time since loss. However, as this analysis was only based on eight studies, the results are less robust and should be interpreted as exploratory.

In addition to the quantification of PGD prevalence, our systematic review and meta-analysis revealed a large degree of variability in prevalence estimates among the currently published studies. Our results indicate how a considerable amount of this variability (I^2=96.9%) is due to heterogeneity, suggesting systematic differences exist between study results. In an attempt to address this heterogeneity, we conducted a number of sub-group moderator analyses in relation to different study characteristics. While these between-group differences did not reach statistical significance, perhaps due to insufficient statistical power, prevalence estimates tended to be higher in studies a longitudinal design (than in cross-sectional studies), for studies conducted with non-probability sampling methods (compared to probability sampling), and in studies using self-reporting (compared to other-reporting). When comparing studies conducted in Western countries (i.e., Australia, Denmark, Germany, the Netherlands, and the United States) and Eastern countries (i.e., China and Japan), the pooled prevalence of PGD was higher in Western studies than in Eastern studies. If this difference holds out in future studies, a geographical variation could possibly reflect differences in the applied methodologies, but could also be caused by potential differences in the expression of grief across cultures. Although we found no significant differences between studies with a low, moderate, or high risk of bias, the four studies with a low risk of bias had a pooled prevalence of 10.8%, which is comparable to the overall prevalence rate found for all studies adjusted for publication bias (11.0%). It should be noted that all sub-groups continued to exhibit a high degree of heterogeneity (I^2 ranged from 51.6% to 98.9%) suggesting that the within-group differences in ERs stem from systematic differences in study characteristics rather than sampling error.

4.1. Heterogeneity and methodological diversity

Variability in study design may be described as methodological diversity (sometimes called methodological heterogeneity) and often reflect that the included studies suffer from various degrees of bias (Higgins and Green, 2011). Based on the risk of bias assessment conducted as part of this review, this appeared to be the case. Thus, four of the included studies had high RoB, six studies moderate RoB, and only four studies low RoB. Our assessment further revealed that while the identified studies generally were characterized by high levels of internal validity (e.g., measuring procedures suggested a correct estimate of

PGD among the sampled participants), most studies generally exhibited some methodological limitations regarding external validity (e.g., sampling procedures did not ensure representative samples) (Table 2).

In particular, we found the limited external validity mirrored in the sampling contexts of the included studies. Many studies were conducted in single towns or regions, used small sample sizes, or employed participant self-selection recruitment methods (e.g., Li and Prigerson, 2016; Middleton et al., 1996). Only a few studies used a population-based approach with random selection (Kersting et al., 2011; Miyajima et al., 2014; Mizuno et al., 2012; Newson et al., 2011). Another important external validity issue relates to the likelihood of non-response bias (Hoy et al., 2012), with only two studies taking efforts to minimize the impact of non-responders (Kim et al., 2015; Newson et al., 2011). As the group of individuals suffering the most following a loss might not have the sufficient resources to complete study participation, investigating and minimizing differences between responders and non-responders is crucial to avoid biases in bereavement studies. Taken together, the limited representativeness of some study-samples and the lack of non-responder analyses reduce the comparability between studies.

Another potential threat to the studies' comparability is the use of several different instruments to measure PGD. Although these assessment instruments all exhibited acceptable psychometric properties (Burnett et al., 1997; Byrne and Raphael, 1994; Prigerson et al., 2009, 1995b; Prigerson and Jacobs, 2001; Shear et al., 2006), possible differences between the scales in how the construct is measured could pose a challenge for comparison. This seems supported by the results of our sub-group analysis, which revealed significant between-instrument differences in ERs (Q=21.1, $p \leq 0.001$). However, we also found a large degree of within-instrument heterogeneity (I^2 ranged from 65.5% to 98.4%) when studies using the same scale were compared. One explanation for the within-instrument heterogeneity could be that some studies applied different cut-off scores than those stated in the original scale (e.g., Li and Prigerson, 2016). Furthermore, each instrument was used on very different study populations, which added to the large variation in prevalence rates obtained for each scale. For example, the ICG was used both in a study measuring PGD at an average of 6.7 months post-loss (Newson et al., 2011) resulting in a prevalence of 25.4% and in a study at 26 months post-loss (Li and Prigerson, 2016) showing a prevalence of PGD of 13.9%. This wide range was reflected in a high degree of within-instrument heterogeneity (I^2=95.6).

13

To sum up, the large between-study heterogeneity in our meta-analysis may be a result of (i) different degrees of bias among the studies; (ii) highly diverse sampling contexts and sampling procedures; (iii) a high risk of non-responder bias; and, perhaps most seriously, (iiii) the use of different assessment instruments with different translations and applied cut-off scores on highly diverse samples. All of this points to the necessity of aligning the undertaken research in future studies of PGD to reach firmer conclusions regarding the disorder's prevalence.

4.2. Clinical implications

Currently, PGD is a proposed candidate for inclusion as a mental disorder in the *ICD-11* (World Health Organization, 2016) just as the *DSM-5* acknowledged grief as a possible condition of clinical attention when Persistent Complex Bereavement Disorder (PCBD) was included in the Appendix (American Psychiatric Association, 2013). The results of a recently published study found that the differences between PGD and PCBD may be merely semantic as the two disorders show comparable predictive validity and similar high levels of diagnostic specificity (95.0-98.3%) (Maciejewski et al., 2016). This finding indicates that researchers and clinicians informed by both the *ICD-11* and the *DSM-5* may benefit from the results of the present review.

Somewhat surprisingly, our results revealed an age-effect suggesting that older age could be associated with higher prevalence of PGD (Table 4). In contrast with earlier claims that older people will be more prepared for coping with grief because they, due to biological reasons, are more likely to experience bereavement involving persons from their closest social network, our finding indicates how aging and bereavement may interact to create a vulnerability towards PGD. A reason for this could be that even though bereavement is more expected and more frequent, older individuals are likely to enter grief from a more problematic somatic health position (Hansson and Stroebe, 2007). Numerous studies have found loss in old age to be associated with a range of negative health-related outcomes such as nutritional problems and sleep disturbances as well as higher use of health services and higher mortality (e.g., Hansson and Stroebe, 2007; Richardson et al., 2003; Rosenbloom and Whittington, 1993). It might be that a lower general health position results in a range of negative consequences which in turn constrict the coping resources of older bereaved individuals and thus, increase the risk of PGD in this group. Furthermore, old age is often related to decreased social contact (e.g., Edmondson, 2013; Vanderwerker and Prigerson, 2004), which

for the bereaved individual might result in a disengagement from the surrounding world and a restricted attention on the deceased adding to develop PGD symptoms and maintain these over time. If the result from the present review is replicated in future studies, this will underline that older individuals might be more fragile grievers than expected which should raise special attention, for example regarding clinical training in therapy with older people.

As indicated by our results, one out of ten bereaved adults are expected to have clinically relevant levels of PGD symptoms. This result imply that the adverse outcomes associated with PGD are concentrated in a minority why the primary goal for clinicians is to identify those particularly vulnerable individuals. To ensure that specialist care is allocated to those in greatest need, clinicians may benefit from taking the duration criteria of PGD into account so as not to confound natural grief with prolonged grief. In a meta-analysis by Wittouck et al. (2011), preventive interventions did not appear effective in reducing the probability of developing PGD, and a growing body of research even show interventions aimed at addressing natural grief reactions to be contraindicated (e.g., Bonanno and Lilienfeld, 2008; Neimeyer, 2000). Thus, as preventive interventions for PGD may be both clinically ineffective and psychologically harmful, resources should be allocated for developing effective clinical screening strategies for PGD. Systematic screening procedures may improve identification of core issues and selection of optimal treatment strategies for people with high symptom levels, but it may also help identify those who are coping well with their grief and whom may not benefit or may even experience increased bereavement related distress following treatment. As shown by the present study, a number of different tools are currently used to screen for PGD (e.g., BGQ, ICG-R, and PG-13), which may result in different prevalence rates and differences in who is identified to be at risk. This clearly illustrates a need to align the field and develop a precise screening tool balancing between being theoretically sound, empirically well‑validated, and practically useful in clinical settings (Sealey et al., 2015).

4.3. Recommendations for future research

With a revealed prevalence of PGD of 9.8%, this meta-analysis provides a preliminary benchmark for future research in PGD to be compared against. However, the present study also found a large degree of heterogeneity among the currently published studies on prevalence of PGD. Although we aimed to identify quality studies, the relatively poor external validity and the high level of heterogeneity of the included studies suggest that the included

studies may not all be estimating the same phenomenon. While the high variability of prevalence rates across studies included in the meta-analysis is likely to be a result of methodological differences (e.g., sampling procedures, reduction of non-responder bias, time since loss, different assessment instruments), this does not rule out the existence of a true underlying prevalence rate of PGD. Until now, studies may not have been conducted with sufficient similarity in design and methods to detect this "true" prevalence of PGD.

Along with maintaining a high level of internal validity, our results indicate that special efforts should be made to secure high external validity when planning new research in PGD. This seems essential to enable generalizable conclusions beyond a specific sample, which a considerable proportion of the existing studies may be unable to do. Future original studies on PGD prevalence are recommended to take a population-based approach and increase representativeness, for example by recruitment using national registries. This will allow comparable analyses of different socio-demographic variables of responders and non-responders, thereby increasing the generalizability of results to the bereaved population in general. From a nosological perspective, expansions and alignment of this research seems crucial as PGD is included in the Appendix of *DSM-5* and is expected to be included as a mental disorder in *ICD-11* (Bryant, 2014; Maciejewski et al., 2016; Maercker et al., 2013; Rosner, 2015).

4.4. Strengths and limitations

A major strength of the present systematic review is the comprehensive selection process with two researchers independently conducting the literature search, study screening, and assessment of bias, reaching high levels of interrater reliability throughout the process. Additionally, the review is based on an a priori developed and publically registered research protocol (Lundorff et al., 2016). Moreover, a generally high level of internal validity was found for the included studies (e.g., studies used the same mode of data collection for all subjects and applied a clear case definition).

Despite these strengths, a number of limitations should be mentioned. To begin with, one limitation relates to the instrument applied for assessing risk of bias (RoB) (Hoy et al., 2012), which established how six studies had a low RoB, four studies had a moderate RoB, while the remaining four of the included studies had a high RoB. This tri-partition of studies into low, moderate, and high-risk categories was based on rather conservative benchmarks implying that only studies with a score of 9 or 10 out of 10 were considered low-risk studies. This resulted in some studies receiving a "moderate" categorization, even though they were

16

conducted and reported in a highly rigorous manner. In addition, the RoB assessment strictly addressed the risk of bias in relation to prevalence study quality parameters (e.g., reporting of numerator and denominator, analysis of non-responders). As some of the studies included in the analysis were not primarily designed as prevalence studies, these studies were rated as having high RoB although they were not necessarily low-quality studies as such. This implies that the overall methodological validity of the included studies might be higher than what the RoB assessment conducted as part of this review has exemplified.

Another issue relates to location bias, which is a common limitation for systematic reviews as some relevant studies may not have been included given the choice of search string and electronic databases. However, by manually examining the reference lists of the included full-text papers, actions were taken with the aim to ensure an exhaustive search and limit location bias (Papaioannou et al., 2010). This search resulted in inclusion of four additional records, not otherwise located by the electronic search. Furthermore, by applying English-language as an inclusion criterion, language bias might have affected the review. Yet, it should be noted that the present review only included epidemiological studies, which are less prone to suffer from language bias compared to effect studies (Higgins and Green, 2011). This adds to the argument that the identified studies are likely to be representative of the existing research on prevalence rates of PGD.

The methodological quality of the current literature, however, clearly varies and may be the most significant limitation of the generalizability of our results. First, only a small number of studies used random sampling procedures and analyzed and reported differences between responders and non-responders. Second, the field of PGD appears to be characterized by considerable methodological diversity across studies, including between-study differences in study design, instruments used, cut-off points applied (even within the same scale), assessment time since loss, assessor, and participant characteristics (e.g., relations to the deceased, age, gender). In addition, study sample sizes varied considerably from 57 to 1402 participants, which may challenge the comparability of our findings. These limitations in the literature may have led to the large variation of prevalence rates across studies (1.8-25.4%), significant statistical heterogeneity, and a wide prediction interval (2.2-36.4%), indicating the broad range of values that future observations may fall within. Further original, high-quality studies are needed to clarify PGD prevalence with less heterogenetic results.

5. Conclusion

The present systematic review and meta-analysis offers a first attempt at quantifying the prevalence of PGD, thereby improving our knowledge beyond what can be learned from earlier more narrative accounts. Our results indicated a prevalence of PGD among bereaved adults of 9.8% (95% CI 0.07-0.14) suggesting how one out of ten experiencing bereavement in adulthood will exhibit clinically significant levels of PGD symptoms. Our meta-regression analysis found age to be a near-significant moderator, suggesting that older age could be associated with higher prevalence of PGD. Considering that PGD is expected to be included in *ICD-11* and that a code corresponding to prolonged grief problems – Persistent Complex Bereavement Disorder (PCBD) – was included in the appendix of *DSM-5*, these results have clinical implications for international healthcare services.

Besides a numerical prevalence rate, our study further revealed how the presently available studies are characterized by methodological limitations, a large degree of between-study heterogeneity, and high risk for external validity bias reflected in a wide prediction interval (95% PI 0.022-0.364). Thus, generalizability may be challenged and the results should be interpreted with caution. Given that research results often inform important decisions on the allocation of economic and professional resources in public healthcare settings, the lack of consistency in the undertaken research of PGD is clearly of pertinent interest. To ensure that future decisions are informed by high-quality research, the registry-based studies included in this meta-analysis (Kersting et al., 2011; Miyajima et al., 2014; Mizuno et al., 2012; Newson et al., 2011) should be supplemented with other large-scale population-based studies using random sampling methods. This will increase the field's external validity by enabling comparative analyses between responders and non-responders. These future studies would be strengthened further if the applied assessment instruments are validated and combined with clinical diagnostic interviews to secure a correct classification of PGD and concurrently maintain a high level of internal validity.

References

American Psychiatric Association, 2013. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®), 5. ed. ed. American Psychiatric Publishing, Washington, D.C.

Anderson, W.G., Arnold, R.M., Angus, D.C., Bryce, C.L., 2008. Posttraumatic stress and complicated grief in family members of patients in the intensive care unit. J Gen Intern Med 23, 1871–1876. doi:10.1007/s11606-008-0770-2

Biostat, 2016. Comprehensive Meta‑Analysis Software (CMA).

Boelen, P.A., Prigerson, H.G., 2007. The influence of symptoms of prolonged grief disorder, depression, and anxiety on quality of life among bereaved adults. Eur. Arch. Psychiatry Clin. Neurosci. 257, 444–452. doi:10.1007/s00406-007-0744-0

Bonanno, G.A., Lilienfeld, S.O., 2008. Let's be realistic: When grief counseling is effective and when it's not. Prof. Psychol. Res. Pract. 39, 377–378. doi:http://dx.doi.org.ez.statsbiblioteket.dk:2048/10.1037/0735-7028.39.3.377

Borenstein, M., Hedges, L. V., Higgins, J.P.T., Rothstein, H.R., 2011. Introduction to Meta-Analysis. Wiley.

Brent, D., Melhem, N., Donohoe, M.B., Walker, M., 2009. The incidence and course of depression in bereaved youth 21 months after the loss of a parent to suicide, accident, or sudden natural death. Am. J. Psychiatry 166, 786–794 9p. doi:10.1176/appi.ajp.2009.08081244

Bryant, R.A., 2014. Prolonged grief: where to after Diagnostic and Statistical Manual of Mental Disorders, 5th Edition? Curr. Opin. Psychiatry 27, 21–26. doi:10.1097/yco.0000000000000031

Burnett, P., Middleton, W., Raphael, B., Martinek, N., 1997. Measuring core bereavement phenomena. Psychol. Med. 27, 49–57.

Byrne, G.J., Raphael, B., 1994. A longitudinal study of bereavement phenomena in recently widowed elderly men. Psychol. Med. 24, 411–421.

Chiu, Y.W., Yin, S.M., Hsieh, H.Y., Wu, W.C., Chuang, H.Y., Huang, C.T., 2011. Bereaved females are more likely to suffer from mood problems even if they do not meet the criteria for prolonged grief. Psychooncology 20, 1061–1068. doi:10.1002/pon.1811

Coolican, H., 2014. Research methods and statistics in psychology, Sixth edit. ed. Routledge, Hove.

Duval, S., Tweedie, R., 2000. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 56, 455–63.

Dyregrov, K., Dyregrov, A., Kristensen, P., 2015. Traumatic Bereavement and Terror: The Psychosocial Impact on Parents and Siblings 1.5 Years After the July 2011 Terror Killings in Norway. J. Loss Trauma 20, 556–576. doi:10.1080/15325024.2014.957603

Edmondson, R., 2013. The sociology of ageing, in: Oxford Textbook of Old Age Psychiatry. Oxford University Press, pp. 23–38. doi:10.1093/med/9780199644957.003.0003

Egger, M., Davey Smith, G., Schneider, M., Minder, C., 1997. Bias in meta-analysis detected by a simple, graphical test. BMJ 315, 629–34.

Fujisawa, D., Miyashita, M., Nakajima, S., Ito, M., Kato, M., Kim, Y., 2010. Prevalence and determinants of complicated grief in general population. J Affect Disord 127, 352–358. doi:10.1016/j.jad.2010.06.008

Goldsmith, B., Morrison, R.S., Vanderwerker, L.C., Prigerson, H.G., 2008. Elevated rates of prolonged grief disorder in African Americans. Death Stud 32, 352–365. doi:10.1080/07481180801929012

Hansson, R.O., Stroebe, M.S., 2007. Bereavement in late life : coping, adaptation, and developmental influences, 1. ed. ed. American Psychological Association, Washington, DC.

He, L., Tang, S.Q., Yu, W., Xu, W., Xie, Q.Y., Wang, J.P., 2014. The prevalence, comorbidity and risks of prolonged grief disorder among bereaved Chinese adults. Psychiatry Res 219, 347–352. doi:10.1016/j.psychres.2014.05.022

Higgins, J., Green, S., 2011. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0 [updated March 2011]. [WWW Document]. Cochrane Collab. URL www.cochrane-handbook.org

Higgins, J.P.T., Thompson, S.G., Deeks, J.J., Altman, D.G., 2003. Measuring inconsistency in meta-analyses. BMJ 327, 557–560. doi:10.1136/bmj.327.7414.557

Hoy, D., Brooks, P., Woolf, A., Blyth, F., March, L., Bain, C., Baker, P., Smith, E., Buchbinder, R., 2012. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. J. Clin. Epidemiol. 65, 934–939. doi:10.1016/j.jclinepi.2011.11.014

Jordan, A.H., Litz, B.T., 2014. Prolonged grief disorder: Diagnostic, assessment, and treatment considerations. Prof. Psychol. Res. Pract. 45, 180–187. doi:10.1037/a0036836

Kentish-Barnes, N., Chaize, M., Seegers, V., Legriel, S., Cariou, A., Jaber, S., Lefrant, J.Y., Floccard, B., Renault, A., Vinatier, I., Mathonnet, A., Reuter, D., Guisset, O., Cohen-Solal, Z., Cracco, C., Seguin, A., Durand-Gasselin, J., Eon, B., Thirion, M., Rigaud,

J.P., Philippon-Jouve, B., Argaud, L., Chouquer, R., Adda, M., Dedrie, C., Georges, H., Lebas, E., Rolin, N., Bollaert, P.E., Lecuyer, L., Viquesnel, G., Leone, M., Chalumeau-Lemoine, L., Garrouste, M., Schlemmer, B., Chevret, S., Falissard, B., Azoulay, E., 2015. Complicated grief after death of a relative in the intensive care unit. Eur. Respir. J. 45, 1341–1352. doi:10.1183/09031936.00160014

Kersting, A., Brähler, E., Glaesmer, H., Wagner, B., Brahler, E., Glaesmer, H., Wagner, B., Brähler, E., Glaesmer, H., Wagner, B., Brahler, E., Glaesmer, H., Wagner, B., 2011. Prevalence of complicated grief in a representative population-based sample. J Affect Disord 131, 339–343. doi:10.1016/j.jad.2010.11.032

Kim, Y., Carver, C.S., Spiegel, D., Mitchell, H.H.-R., Cannady, R.S., 2015. Role of family caregivers' self-perceived preparedness for the death of the cancer patient in long-term adjustment to bereavement. Psychooncology. doi:10.1002/pon.4042

Li, J., Prigerson, H.G., 2016. Assessment and associated features of prolonged grief disorder among Chinese bereaved individuals. Compr Psychiatry 66, 9–16. doi:10.1016/j.comppsych.2015.12.001

Liberati, A., Altman, D.G., Tetzlaff, J., Mulrow, C., Gøtzsche, P.C., Ioannidis, J.P.A., Clarke, M., Devereaux, P.J., Kleijnen, J., Moher, D., 2009. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. BMJ 339.

Lichtenthal, W.G., Nilsson, M., Kissane, D.W., Breitbart, W., Kacel, E., Jones, E.C., Prigerson, H.G., 2011. Underutilization of mental health services among bereaved caregivers with prolonged grief disorder. Psychiatr Serv 62, 1225–1229. doi:10.1176/appi.ps.62.10.1225

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 Stud. 34, 673–98. doi:10.1080/07481187.2010.496686

Lundorff, M., Holmgren, H., Farver-Vestergaard, I., Zachariae, R., O'Connor, M., 2016. Prevalence of prolonged grief disorder in middle and old age bereavement: a systematic review and meta-analysis [WWW Document]. PROSPERO Regist. CRD42016038416. URL http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016038416

Maciejewski, P.K., Maercker, A., Boelen, P.A., Prigerson, H.G., 2016. "Prolonged grief disorder" and "persistent complex bereavement disorder", but not "complicated grief",

are one and the same diagnostic entity: an analysis of data from the Yale Bereavement Study. World Psychiatry 15, 266–275. doi:10.1002/wps.20348

Maercker, A., Brewin, C.R., Bryant, R.A., Cloitre, M., van Ommeren, M., Jones, L.M., Humayan, A., Kagee, A., Llosa, A.E., Rousseau, C., Somasundaram, D.J., Souza, R., Suzuki, Y., Weissbecker, I., Wessely, S.C., First, M.B., Reed, G.M., 2013. Diagnosis and classification of disorders specifically associated with stress: proposals for ICD-11. World Psychiatry 12, 198–206. doi:10.1002/wps.20057

Maercker, A., Lalor, J., 2012. Diagnostic and clinical considerations in prolonged grief disorder. Dialogues Clin Neurosci 14, 167–176.

McCarthy, M.C., Clarke, N.E., Ting, C.L., Conroy, R., Anderson, V.A., Heath, J.A., 2010. Prevalence and predictors of parental grief and depression after the death of a child from cancer. J Palliat Med 13, 1321–1326. doi:10.1089/jpm.2010.0037

McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochem. Medica 22, 276–282.

Middleton, W., Burnett, P., Raphael, B., Martinek, N., 1996. The bereavement response: a cluster analysis. Br. J. Psychiatry J. Ment. Sci. 169, 167–171.

Miyajima, K., Fujisawa, D., Yoshimura, K., Ito, M., Nakajima, S., Shirahase, J., Mimura, M., Miyashita, M., 2014. Association between Quality of End-of-Life Care and Possible Complicated Grief among Bereaved Family Members. J Palliat Med 17, 1025–1031. doi:10.1089/jpm.2013.0552

Mizuno, Y., Kishimoto, J., Asukai, N., 2012. A nationwide random sampling survey of potential complicated grief in Japan. Death Stud 36, 447–461. doi:10.1080/07481187.2011.553323

Neimeyer, R.A., 2000. Searching for the meaning of meaning: Grief therapy and the process of reconstruction. Death Stud. 24, 541–558.

Newson, R.S., Boelen, P.A., Hek, K., Hofman, A., Tiemeier, H., 2011. The prevalence and characteristics of complicated grief in older adults. J Affect Disord 132, 231–238. doi:10.1016/j.jad.2011.02.021

O'Connor, M., Lasgaard, M., Shevlin, M., Guldin, M.-B., 2010. A confirmatory factor analysis of combined models of the Harvard Trauma Questionnaire and the Inventory of Complicated Grief-Revised: are we measuring complicated grief or posttraumatic stress? J. Anxiety Disord. 24, 672–679. doi:10.1016/j.janxdis.2010.04.009

Papaioannou, D., Sutton, A., Carroll, C., Booth, A., Wong, R., 2010. Literature searching for

social science systematic reviews: consideration of a range of search techniques. Health Info. Libr. J. 27, 114–122. doi:10.1111/j.1471-1842.2009.00863.x

Piper, W.E., Ogrodniczuk, J.S., Azim, H.F., Weideman, R., 2001. Prevalence of loss and complicated grief among psychiatric outpatients. Psychiatr Serv 52, 1069–1074. doi:10.1176/appi.ps.52.8.1069

Prigerson, H.G., Frank, E., Kasl, S. V, Reynolds, C.F., Anderson, B., Zubenko, G.S., Houck, P.R., George, C.J., Kupfer, D.J., 1995a. Complicated grief and bereavement-related depression as distinct disorders: preliminary empirical validation in elderly bereaved spouses. Am J Psychiatry 152, 22–30. doi:10.1176/ajp.152.1.22

Prigerson, H.G., Horowitz, M.J., Jacobs, S.C., Parkes, C.M., Aslan, M., Goodkin, K., Raphael, B., Marwit, S.J., Wortman, C., Neimeyer, R.A., Bonanno, G.A., Bonanno, G.A., Block, S.D., Kissane, D., Boelen, P., 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 Med. 6, 1–12. doi:10.1371/journal.pmed.1000121

Prigerson, H.G., Jacobs, S.C., 2001. Diagnostic criteria for traumatic grief: A rationale, consensus criteria, and preliminary empirical test, in: Stroebe, M.S., Hansson, R.O., Stroebe, W., Schut, H. (Eds.), Handbook of Bereavement Research: Consequences, Coping, and Care. American Psychological Association., Washington, D.C., pp. 614–646.

Prigerson, H.G., Maciejewski, P.K., Reynolds, C.F., Bierhals, A.J., Newsom, J.T., Fasiczka, A., Frank, E., Doman, J., Miller, M., 1995b. Inventory of Complicated Grief: a scale to measure maladaptive symptoms of loss. Psychiatry Res. 59, 65–79.

Richardson, S.J., Lund, D.A., Caserta, M.S., Dudley, W.N., Obray, S.J., 2003. Sleep patterns in older bereaved spouses. Omega-Journal of Death and Dying 47, 361–383. doi:10.2190/0d4f-epmw-3yul-c1jk

Rosenbloom, C.A., Whittington, F.J., 1993. The effects of bereavement on eating behaviors and nutrient intakes in elderly widowed persons. J. Gerontol. 48, S223-9.

Rosner, R., 2015. Prolonged grief: setting the research agenda. Eur. J. Psychotraumatol. 6. doi:10.3402/ejpt.v6.27303

Rosner, R., Kruse, J., Hagl, M., 2010. A meta-analysis of interventions for bereaved children and adolescents. Death Stud. 34, 99–136.

Sackett, D.L., Richardson, W.S., Rosenberg, W., Haynes, R.B., 1997. Evidence-based

medicine : how to practice and teach EBM. Churchill Livingstone, New York.

Sealey, M., Breen, L.J., O'Connor, M., Aoun, S.M., 2015. A scoping review of bereavement risk assessment measures: Implications for palliative care. Palliat. Med. 29, 577–589. doi:10.1177/0269216315576262

Shah, S.N., Meeks, S., 2012. Late-life bereavement and complicated grief: A proposed comprehensive framework. Aging Ment. Health 16, 39–56. doi:10.1080/13607863.2011.605054

Shear, K.M., Jackson, C.T., Essock, S.M., Donahue, S.A., Felton, C.J., 2006. Screening for Complicated Grief Among Project Liberty Service Recipients 18 Months After September 11, 2001. Psychiatr. Serv. 57, 1291–1297. doi:10.1176/appi.ps.57.9.1291

Siegel, M.D., Hayes, E., Vanderwerker, L.C., Loseth, D.B., Prigerson, H.G., 2008. Psychiatric illness in the next of kin of patients who die in the intensive care unit. Crit Care Med 36, 1722–1728. doi:10.1097/CCM.0b013e318174da72

Simon, N.M., Shear, K.M., Thompson, E.H., Zalta, A.K., Perlman, C., Reynolds, C.F., Frank, E., Melhem, N.M., Silowash, R., 2007. The prevalence and correlates of psychiatric comorbidity in individuals with complicated grief. Compr Psychiatry 48, 395–399. doi:10.1016/j.comppsych.2007.05.002

Stroup, D.F., Berlin, J.A., Morton, S.C., Olkin, I., Williamson, G.D., Rennie, D., Moher, D., Becker, B.J., Sipe, T.A., Thacker, S.B., 2000. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. JAMA 283, 2008–2012.

Tal Young, I., Iglewicz, A., Glorioso, D., Lanouette, N., Seay, K., Ilapakurti, M., Zisook, S., 2012. Suicide bereavement and complicated grief. Dialogues Clin Neurosci 14, 177–186.

Vanderwerker, L., Prigerson, H., 2004. Social support and technological connectedness as protective factors in bereavement. J. Loss Trauma.

Varga, M.A., McClam, T.M., Hassane, S., 2015. Grief Experiences Among Female American and Arab Undergraduate College Students. Omega-Journal of Death and Dying 72, 165–183. doi:10.1177/0030222815574834

Veritas Health Innovation Ltd., 2016. Covidence [WWW Document]. URL https://www.covidence.org/terms

Wittouck, C., Van Autreve, S., De Jaegere, E., Portzky, G., van Heeringen, K., 2011. The prevention and treatment of complicated grief: a meta-analysis. Clin. Psychol. Rev. 31,

69–78. doi:10.1016/j.cpr.2010.09.005

World Health Organization, 2016. ICD-11 Beta Draft - Joint Linearization for Mortality and Morbidity Statistics [WWW Document]. URL http://apps.who.int/classifications/icd11/browse/l-m/en#/http%253a%252f%252fid.who.int%252ficd%252fentity%252f1183832314

Zisook, S., Shear, K., 2009. Grief and bereavement: what psychiatrists need to know. World Psychiatry 8, 67–74.

Zisook, S., Simon, N.M., Reynolds, C.F., Pies, R., Lebowitz, B., Young, I.T., Madowitz, J., Shear, M.K., 2010. Bereavement, complicated grief, and DSM, part 2: complicated grief. J. Clin. Psychiatry 71, 1097–1098. doi:10.4088/JCP.10ac06391blu

Figure captions

Figure 1. Flow diagram of study selection.

Figure 2. Forest plot of random effects model meta-analysis of prevalence estimates (event rate) with 95% confidence intervals.

Table 1
Characteristics of included studies.

Author(s) (year)	Study setting	Study design	Diagnostic criteria (terminology; assessment instrument)	Sample recruitment and characteristics	Study population (sample size; % female; mean age)	Mean time post-loss (months)	Relation to the deceased
Byrne and Raphael (1994)	Australia	Longitudinal, other-report	Chronic grief; BPQ	Recently widowed elderly men. Cluster sample identified through repeated searches in State Registry.	57; 0% female; 74.52 years	13.0	100% spouse
Fujisawa et al. (2016)	Japan	Cross-sectional, self-report	Complicated grief; BGQ	Individuals aged 40–79 years who had experienced bereavement within the past ten years identified through stratified two-stage random sampling. Excluded parents who had lost a child.	969; 58.2% female; 56.32 years	72.0*	6.3% spouse; 48.3% parent; 25.5% parent-in law; 9.9% sibling; 10.0% other
Goldsmith et al. (2008)	USA	Longitudinal, self-report	Prolonged grief; ICG-R	Data from the Yale Bereavement Study (recruitment through a community-based outreach program) and the Coping with Cancer Study (multi-site recruitment of cancer caregivers). Cluster sampling.	538; 73% female; 59.00 years	6.0	77.5% spouse; 9.5% adult child; 5.6% parent; 7.4% other
He et al. (2014)	China	Cross-sectional, self-report	Prolonged grief; PG-13	Individuals who had experienced the loss of a significant other more than six months prior to study onset. Recruitment through multistage random sampling.	445; 78.4% female; 27.60 years	55.1*	23.9% parent; 1.8% child; 2.5% spouse; 2.9% sibling; 58.5% grandparents; 10.4% other
Kersting et al. (2011)	Germany	Cross-sectional, self-report	Complicated grief; ICG-R	Individuals who had experienced the loss of a significant other identified through random multi-stage sampling procedures.	1402; 55.6%* female; 43.87 years	117.1*	22.1% spouse; 2.6% child; 44.5% parent; 5.4% sibling; 25.4% other
Kim et al. (2015)	USA	Longitudinal, self-report	Prolonged grief; ICG	Adult bereaved cancer caregivers identified using multiple state cancer registries. Cluster sampling.	137; 82.5% female; 55.99 years	34.4*	57.7% spouse, 42.3% other
Li and Prigerson (2016)	China	Cross-sectional, self-report	Prolonged grief; ICG	Voluntary response sample recruited from two Chinese memorial websites. Inclusion criteria were Chinese adults, at least 18 years old who lost a first-degree relatives.	1099; 51% female; 41.85 years	26.0	40.1% father, 32.0% mother, 6% child, 11.3% spouse, 10.6% sibling
Middleton et al. (1996)	Australia	Longitudinal, other-report	Chronic grief; CBI	Participants were recently bereaved spouses under the age of 70, adult children who had lost a parent and parents who had lost a child. Recruitment through hospital notifying officers or through state registry. Voluntary response sampling.	120; 66.4% female; 44.65 years	12.0	36.7% spouse; 33.3% parent; 30% child
Miyajima et al. (2014)	Japan	Cross-sectional, self-report	Complicated grief; BGQ	Adult individuals who lost their loved one within the period between 6 months and 10 years were randomly sampled from census tracts.	641; 54.9% female; 58.93 years	53.6*	5.5% spouse, 51.0% parent, 26.5% parent-in-law, 7.8% sibling, 9.2% other
Mizuno et al. (2012)	Japan	Cross-sectional, self-report	Complicated grief; BGQ	Individuals aged 18 years or older who had any experience of significant loss were identified using stratified two-stage random sampling method.	850; 53.3% female; 51.00 years	142.8*	NR
Newson et al. (2011)	Netherlands	Cross-sectional, other-report	Complicated grief; ICG	Participants were identified through a cluster sampling of all inhabitants aged over 55 years living in the district of Rotterdam. Eligible participants self-reported to be currently experiencing grief.	1089; 73.74% female; 64.06 years	6.7	34.9% spouse/partner; 10.0% parent; 11.4% child; 15.8% sibling; 6.9% friend, 21% other
O'Connor et al. (2010)	Denmark	Cross-sectional, self-report	Complicated grief; ICG-R	A general sample of married elderly people living in Aarhus County was randomly selected through the Danish Central Person Register (CPR). Participants had experienced a significant interpersonal loss.	292; 50% female; 70.00 years	13.5	63% parent, 13% spouse, 12% sibling, 6% child, 5% friend; 1% other
Prigerson et al. (2009)	USA	Longitudinal, other-report	Prolonged grief; PG-13	Adult bereaved individuals recruited through a community-based outreach program in Connecticut and from pastoral care offices in the New Haven area. Cluster sampling.	215; 73.7% female; 61.8 years	9.0	83.9% spouse, 16.1% non-spouse
Varga et al. (2015)	USA	Cross-sectional, self-report	Prolonged grief; PG-13	Female undergraduate students randomly sampled from a university in the United Arab Emirates and from the United States.	181; 100% female; 21.00 years	9.6*	30.6% grandparent, 22.2% friends, 7.8% parents, 1.7% siblings, 37.7% other

Abbreviations: BGQ, Brief Grief Questionnaire (Shear et al., 2006); BPQ, Bereavement Phenomenology Questionnaire (Byrne and Raphael, 1994); CBI, Core Bereavement Items (Burnett et al., 1997); ICG, Inventory of Complicated Grief (Prigerson et al., 1995b); ICG-R, Inventory of Complicated Grief-Revised (Prigerson and Jacobs, 2001); PG-13, Prolonged Grief-13 (Prigerson et al., 2009); NR, information was not reported in the original study and could not be calculated
* numbers/percentages are calculated based on raw numbers, since they were not reported in the original study

Table 2
Results of risk of bias assessment.

Author(s) (year)	External validity				Internal validity						Total RoB [a]
	1. Representativeness	2. Sampling frame	3. Random selection	4. Non-response	5. Data collection	6. Case definition	7. Instrument	8. Same mode	9. Prevalence period	10. Numerator/ denom-inator	
Byrne et al. (1994)	0	1	0	0	1	0	0	1	1	1	High 5/10
Fujisawa et al. (2010)	1	1	1	0	1	1	1	1	1	0	Moderate 8/10
Goldsmith et al. (2010)	0	0	0	0	1	1	1	1	1	1	High 6/10
He et al. (2015)	1	0	0	0	1	1	1	1	1	1	Moderate 7/10
Kersting et al. (2011)	1	1	1	0	1	1	1	1	1	1	Low 9/10
Kim et al. (2015)	1	0	0	1	1	1	1	1	1	0	Moderate 7/10
Li et al. (2016)	0	1	0	0	1	1	1	1	1	1	Moderate 7/10
Middleton et al. (1996)	1	0	0	0	1	1	0	1	1	1	High 6/10
Miyajima et al. (2014)	1	1	1	0	1	1	1	1	1	1	Low 9/10
Mizuno et al. (2012)	1	1	1	0	1	1	1	1	1	1	Low 9/10
Newson et al. (2011)	1	1	1	1	1	1	1	1	1	1	Low 10/10
O'Connor et al. (2010)	1	1	1	0	1	1	1	1	1	0	Moderate 8/10
Prigerson et al. (2009)	1	1	0	0	1	1	1	1	1	1	Moderate 8/10
Varga et al. (2015)	0	0	1	0	1	1	1	1	0	1	High 6/10

Abbreviations: RoB, Risk of bias
[a] measured using the risk of bias tool developed by Hoy et al. (2012). Each item is scored with a value of 1 (low risk) or 0 (high risk), yielding a RoB total score for each study. A score of 9 or 10 was considered low risk of bias; a score of 7 or 8 was considered moderate risk of bias; a score of 6 or less was considered high risk of bias

Table 3

Pooled event rates of PGD grouped by study design, sampling method, assessor, assessment instrument, risk of bias, and region.

		K[a]	Total N	Event rate[b]	95% CI (95% PI)	Heterogeneity Q	p	I²[c]	Publication bias Egger's test, p	Adjusted Event rate[d]
All		14	8035	0.098	0.068-0.140 (0.020-0.363)	416.1	0.001	96.9	0.011	0.110
Study design	Cross-sectional	9	6968	0.096	0.060-0.151	374.8	0.001	97.9	0.018	0.201
	Longitudinal	5	975	0.104	0.058-0.179	31.7	0.001	87.4	NS	-
	Between-groups	*14*				*0.0*	*0.836*			
Sampling	Probability	12	6816	0.095	0.061-0.144	403.2	0.001	97.3	0.013	0.108
	Non-probability	2	1219	0.124	0.085-0.177	2.1	0.151	51.6	NS	-
	Between-groups	*14*				*0.8*	*0.364*			
Assessor	Self-report	10	6554	0.098	0.063-0.149	318.8	0.001	97.2	NS	-
	Other-report	4	1481	0.096	0.033-0.251	53.5	0.001	94.4	NS	-
	Between-groups	*14*				*0.00*	*0.974*			
Assessment instruments[e]	BGQ	3	2460	0.123	0.051-0.270	121.9	0.001	98.4	NS	-
	ICG	3	2325	0.205	0.129-0.310	45.9	0.001	95.6	NS	-
	ICG-R	3	2232	0.092	0.059-0.141	18.3	0.001	89.1	NS	-
	PG-13	3	841	0.032	0.017-0.062	5.8	0.055	65.5	NS	-
	Between-groups	*12*				*21.1*	*0.001*			
Risk of bias	High	4	896	0.092	0.061-0.138	7.8	0.050	61.6	NS	-
	Moderate	6	3157	0.098	0.057-0.165	122.8	0.001	95.9	NS	-
	Low	4	3982	0.108	0.045-0.238	263.0	0.001	98.9	NS	-
	Between-group	*14*				*0.12*	*0.944*			
Region	Western	9	4031	0.101	0.060-0.165	213.6	0.001	96.3	NS	
	Eastern	5	4004	0.092	0.048-0.170	195.0	0.001	98.0	NS	
	Between-groups	*14*				*0.05*	*0.826*			

Abbreviations: BGQ, Brief Grief Questionnaire (Shear et al., 2006); CI, confidence interval; ICG, Inventory of Complicated Grief (Prigerson et al., 1995b); ICG-R, Inventory of Complicated Grief-Revised (Prigerson and Jacobs, 2001); NS, non-significant; PG-13, Prolonged Grief-13 (Prigerson et al., 2009); PGD, Prolonged Grief Disorder; PI, prediction interval

[a] number of studies in analysis

[b] random effects model, weighted by inverse variance, calculated for K≥3

[c] heterogeneity; 0=none, 25=low, 50=moderate, 100=high

[d] if funnel plot and Egger's test (p < 0.05) suggest possible publication bias, event rate is adjusted with the Duval and Tweedie Fill-and-trim method

[e] instruments used in 3 or more studies included.

Table 4

Results of meta-regression analyses.

Moderator variable	Adjusted for [a]	K [b]	Beta [c]	P
Mean age of sample		14	0.03	0.075
Months since loss		14	-0.01	0.243
% women in sample		14	-0.003	0.831
Response rate		11	0.01	0.722
% spouse lost		12	0.00	0.693
% parent lost		10	0.01	0.575
% child lost		7	0.03	0.388
% sibling lost		8	0.15	0.003
% sibling lost	Mean age of sample	8	0.10	0.285
% sibling lost	Months since loss	8	0.17	0.002

[a] statistically significant moderators adjusted for variables with near-significant results ($p < 0.10$)
[b] number of studies in analysis
[c] mixed effects regression: unrestricted maximum likelihood