
Article:
Postlethwaite, A., Kellett, S. and Mataix-Cols, D. (2019) Prevalence of hoarding disorder: A systematic review and meta-analysis. Journal of Affective Disorders, 256. pp. 309-316. ISSN 0165-0327

https://doi.org/10.1016/j.jad.2019.06.004

1.0 Introduction

Hoarding disorder (HD) is defined as a persistent difficulty discarding possessions, resulting in an accumulation of belongings causing severe clutter and the obstruction/congestion of living areas which creates significant distress and impairment in functioning (APA, 2013). Mean age of onset of hoarding symptoms has been estimated to be 13.4 years, with 60% of patients reporting that the onset of symptoms occurred by age 12, increasing to 80% by age 18 (Grisham, Frost, Steketee, Kim, & Hood, 2006). Levels of clutter in the home can range from moderate to extreme levels, which then create associated and increasing levels of impairment (Timpano et al., 2013). When severe hoarding creates and maintains significant clutter in the home, this creates serious risks to personal safety from falls, food contamination, infestation, fire and impeded escape routes (Steketee & Frost, 2014). This array of threats to personal safety are particularly evident within the older adult HD population (Kim, Steketee, & Frost, 2001). HD presents a burden in terms of increased occupational impairment (Neave et al., 2017; Tolin et al., 2008). HD also impacts on others, with family members and carers experiencing it as problematic (Drury et al., 2014; Frost & Gross, 1993). In more severe cases, hoarding threatens the health and safety of neighbours. Complaints are addressed by multiple community services creating associated costs through social service involvement (Tolin et al., 2008) and an associated risk of social shunning in the community (Frost, Steketee, & Williams, 2000).

Research into this newly recognised disorder is still in its infancy (Mataix-Cols and Fernández de la Cruz, 2018). One of the many areas of considerable uncertainty is the actual prevalence of HD. Several frequently cited studies have previously attempted to estimate the point prevalence of HD in adults in sufficiently sized samples, with estimates widely varying from 1.5% to 6% of the general population (e.g. Iervolino et

al., 2009; Mueller et al., 2009; Nordsletten, Reichenberg, et al., 2013; Samuels et al., 2008), with rates tending towards the lower level. From a public health perspective, these disparate estimates are too wide to be useful in guiding the allocation of resources for HD. The commonly cited studies of HD prevalence also possess significant methodological limitations, such as the use of single items included in other instruments not initially designed to detect HD, use of definitions that do not match the current DSM-5 criteria, samples not being representative of the general population due to self-selection, small samples, low response rates and an over reliance on self-report measures. The methodological design of any individual prevalence study can result in systematic error or bias, then leading to overestimation or underestimation of the true prevalence of a disease or disorder (Higgins & Green, 2011). It is therefore inappropriate to denote any one study as being the most accurate or representative of the general population (Barendregt, Doi, Lee, Norman & Vos, 2003). By pooling multiple prevalence studies, it is possible to then estimate an overall HD prevalence rate with greater precision. Also, by combining estimates from different regions of the world that have similar characteristics (e.g. emerging versus developed nations) then also identify otherwise hidden associations (Fiest, Pringham, Pattern, Svenson & Jette, 2014). Consequently, it is important to assess the methodological quality of studies included in any prevalence review (Hoy et al., 2012). This can be achieved by assessing risk of bias, with the selection and synthesis of only the most rigorous and well controlled studies being likely to then reveal a trustworthy prevalence base rate (Higgins & Green, 2011).

Whilst various studies have reported the prevalence of HD in differing populations, there has been no previous attempt to consolidate these studies in order to derive a robust prevalence estimate of HD or to assess how rates reported are affected by methodological factors in the original studies. Ascertaining the population prevalence of HD has important healthcare implications, as it is difficult to design and

justify interventions for HD when the community burden is unreliable or unspecified (Mansfield, Sim, Jordan & Jordan, 2016). The current systematic review therefore had three objectives. The first objective was to conduct a comprehensive, systematic literature search to identify all relevant studies that have reported prevalence data for HD in the general population. The second objective was to conduct a meta-analysis in order to provide a more precise estimate of the prevalence of HD in working age adults. The decision to limit the review to working age adults was based on the fact that hoarding symptoms are typically to assess accurately in children and adolescents (Tolin, Meunier, Frost & Steketee, 2010), and so the prevalence estimate is likely to be inaccurate when including child samples. Hoarding symptoms are typically mild during childhood due to parents typically preventing clutter accumulation, lack of space, and children typically lacking the financial means to consistently acquire possessions (Storch, Rahman, Park, Reid, Murphy & Lewin, 2011). The third objective was to assess whether the exhibited variation in the HD prevalence estimate was associated with the following factors (a) prevalence type (e.g. point vs lifetime prevalence), (b) method of assessment (e.g. interview, self-report) and (c) study quality.

2.0 Method

2.1 Registration and search strategy

This review follows the recommendations regarding the reporting of meta-analyses of observational studies as outlined by Stroup et al. (2000). The study protocol was registered with the PROSPERO international prospective register of systematic reviews (http://www.crd.york.ac.uk/prospero), registration ID: CRD42018093809. An electronic search of three academic databases (PsycINFO, Medline, and Web of Science) was conducted in March 2018. The search specified that within the title, abstract, or topic the article must contain the term “hoard*” (using the asterisk wildcard function to

ensure that all variations were included e.g. “hoarding”, “hoarder”). In addition, the search specified that the article must contain either the term “prevalence” or “incidence.” Search results were limited to human studies, adult populations (18+ years of age) and journal articles. Only English language articles were included in the review. Within the Web of Science search “Medline” and “Zoological Records” were excluded, to avoid duplication (as Medline was searched independently) and to avoid returning animal studies. Further limitations were placed on the Web of Science search by excluding irrelevant areas such as toxicology, architecture, energy fields, optics etc. Searches of the three databases returned 267, 73, and 16 results respectively. After the removal of duplicates, 288 papers were retained for further evaluation. References quoted in the identified papers were hand-searched for any further eligible papers, with one additional paper being identified.

2.2 Eligibility criteria

Papers were relevant if they reported hoarding prevalence data. The minimum required sample size for selection was calculated using the conventional formula (Daniel, 1999; Lwanga & Lemeshow, 1991; Naing, Winn, & Rusli, 2006):

$$n = \frac{Z^2 P(1-P)}{d^2}$$

Where n = sample size,
Z = Z statistic for level of confidence,
P = expected prevalence,
d = precision

The expected prevalence was set to 1.5% (or P = 0.015), with this value taken from a recent and commonly cited HD prevalence estimate (Nordsletten, Reichenberg, et al., 2013). This study was chosen as the reference, because it is the only study to have employed DSM-5 criteria and in-home assessments of clutter. As the expected prevalence was less than 10%, the precision was set to half of P, or 0.0075, as per recommendations (Naing et al., 2006). The confidence interval value was set to 95% (Z = 1.96). Consequently, only those studies with a community sample of greater than

1,009 participants met eligibility criteria and this created an appropriately conservative sampling method.

Articles were excluded if they did not relate to hoarding, did not report original study data (e.g. reviews, book chapters), considered clinical samples only, were comparative studies (e.g. comparing a clinical group with a control group), focused solely on relatives of hoarders, focussed on the clinicians delivering treatment, reported qualitative data only, evaluated child/adolescent population prevalence, evaluated older adult population prevalence or did not report sufficient data. The process of paper selection is presented as a PRISMA diagram (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009) in Figure 1. Initially titles of all 289 non-duplicate papers were scrutinised; 224 articles were excluded based on their title or abstract. Full texts of the remaining 36 papers were examined and 25 were excluded. A total of 11 papers were deemed eligible and were included in the review.

2.3 Data extraction

A data extraction form was used to extract equivalent details of methods and results from each study. Information extracted included: country, sample size, sample age range, sample mean age, response rate, percentage females in sample, hoarding assessment tool, method of collection/assessment, type of prevalence assessed, and reported HD prevalence. The data extraction form also included aspects data relevant to the risk of bias.

2.4 Assessing risk of bias

Risk of bias was assessed using a validated tool developed to assess the methodological quality of prevalence studies (Hoy et al., 2012). The tool consists of 10 items that assess both internal validity (measurement bias) and external validity (selection and non-response bias). Having excluded one item from the tool, Thomas, Sanders, Doust, Beller, and Glasziou (2015) considered studies to be at high risk of bias if they met the

criteria for low risk of bias on 3 items or less. Studies that met criteria for 4 or 5 items were classified as being at moderate risk of bias, and those that met criteria for 6 or more were considered to be at low risk of bias. The current study adopted the categories as reported by Taylor et al. (2014): low (0-3 high-risk items), moderate (4-5 high-risk items), high (6 or more high-risk items). If the information related to an item was unclear in the original study, high risk of bias was recorded for that item.

All of the studies were rated by a second rater. Three of the studies were selected at random and rated by rater 2, a trainee clinical psychologist and the remaining nine studies were second rated by rater 3, a consultant clinical psychologist. To evaluate inter-rater reliability, the intraclass correlation co-efficient (ICC) estimates were calculated using a two-way mixed effects model. Results indicated a moderate degree (Koo & Li, 2016) of reliability between both rater 1 and rater 2: ICC = 0.704, 95% CI: [0.386, 0.858], with good agreement between rater 1 and rater 3 ICC = 0.761, 95% CI: [0.611, 0.836]. Disagreements between the raters were discussed until consensus was reached.

2.5 Meta-analysis

The statistical software package Comprehensive Meta-Analysis version 3 (Borenstein, Hedges, Higgins, & Rothstein, 2018) was used for this prevalence meta-analysis. The unit of data analysed was the estimated prevalence of HD. A random-effects model was used, as it could not be assumed that the studies were functionally identical. Studies were weighted by the inverse of their variance. Therefore, studies with larger samples yielded more precise estimates of the population effect size and so had greater weight towards the estimated mean (Borenstein, Hedges, Higgins, & Rothstein, 2010).

Publication bias was assessed by examining a funnel plot depicting the estimates of each of the studies, following guidelines by Sterne et al. (2011). It is expected that 95% of studies will fall within the funnel plot lines that represent 1.96 standard errors, if

no bias is present. Reliance on visual inspection of funnel plots has been criticised as being unreliable (Terrin, Schmid, & Lau, 2005) and lacking in statistical power (Sterne et al., 2011). Therefore, publication bias was also evaluated statistically using Egger's regression intercept, whereby P values of less than 0.1 indicate statistically significant asymmetry (Egger, Smith, Schneider, & Minder, 1997). Heterogeneity was calculated using Cochran's Q statistic, where a significant P value ($P < 0.05$) indicates statistically significant differences between the studies, and Higgins' I^2, where it has been suggested that a value of 0.25 indicates low heterogeneity, 0.50 indicates medium heterogeneity, and 0.75 equals high heterogeneity (Higgins, Thompson, Deeks, & Altman, 2003).

Moderator analysis was used to assess the association between prevalence and the categorical variables "prevalence type", "method of data collection", and "study quality" (i.e. overall risk of bias rating). Large variation in where studies were conducted made the categorical variable of "location" inappropriate for moderator analysis. As heterogeneity was detected, meta-regression was used to assess the association between prevalence and the following continuous variables: year of publication, proportion of females (gender) and response rate (Thompson & Higgins, 2002). Sample mean age was not analysed as only $k = 5$ studies reported this information.

3.0 Results

A total of $k = 11$ studies, with $n = 53{,}378$ participants were included in the meta-analysis. One of these studies, (Ivanov et al., 2017), reported two different samples based on age, therefore these were treated as separate samples for the analysis. An overview of the study characteristics is presented in Table 1.

3.1 Study characteristics

The majority of the samples included in the analysis were sourced from Europe. Two samples originated in Sweden (as part of the same study), two from the Netherlands, two from Germany and two from the United Kingdom. The remaining samples were sourced from Italy, Australia, Singapore, with a final sample consisting of participants across six differing countries (Belgium, France, Germany, Italy, Netherlands and Spain). Seven of the samples assessed presence of HD using self-report measures, whilst a further two had participants complete self-report measures in the presence of a researcher. The self-report measure used most often was the Hoarding Rating Scale (Tolin, Frost, & Steketee, 2010) and this was used in half of the samples (i.e. 6/12). Three studies assessed participants by interview: Fullana et al. (2010) and Subramaniam et al. (2014) reported using a single item from the OCD symptom checklist of the Composite International Diagnostic interview (Wittchen, 1994), whereas Nordsletten et al. (2013) used the Structured Interview for Hoarding Disorder (Nordsletten, Fernández de la Cruz, et al., 2013). Response rates ranged from 35.9% to 75.9%. The proportion of females ranged from 54.9% to 89.3%. Publication dates ranged from 2009 to 2017.

3.2 HD prevalence

Ten point prevalence estimates (N = 43,958) and two lifetime HD prevalence estimates (N = 9,420) were identified and included in the meta-analysis, with a collective total of N = 53,387 participants. Point prevalence estimates ranged from 0.8-6.03%, and the two lifetime prevalence estimates were 0.8% and 3.5% respectively. The pooled point prevalence estimate for the studies was 2.6%, 95% confidence interval: [1.7 - 3.7%], and the pooled lifetime prevalence estimate was 1.7%, 95% confidence interval: [0.4-6.8%]. There was no significant difference between the pooled lifetime and pooled point prevalence estimates (see covariate analysis). Under the random effects model the overall pooled prevalence estimate for the studies was 2.5%, with a 95% confidence

interval of 1.7-3.6% (see Table 2). Visual inspection of the funnel plot (see Figure 2) suggests an asymmetrical distribution. Egger's regression intercept did not indicate statistically significant asymmetry ($p = 0.114$). However, there was high heterogeneity between the prevalence studies ($Q = 466.521$, $df = 11$, $p < 0.01$, $I^2 = 97.642$).

3.3 Risk of bias

Overall the risk of bias across the studies was low (see Table 3). Of the 12 samples, across 11 studies, 11 were deemed to be at low risk of bias, with a single study (López-Solà et al., 2014) rated as being at moderate risk of bias. No single study was rated as at high risk of bias. Ten of the eleven studies used widely accepted definitions of HD and nine of the eleven studies employed valid case detection methods. Both Fullana et al. (2010) and Subramaniam et al.(2014) used the Composite International Diagnostic Interview which is limited in its assessment of HD (i.e. a single question in the OCD symptom checklist). All numerators and denominators were appropriate and no errors in reporting were detected. The largest possible source of bias related to response rates (see Figure 3). Hoy et al. (2012) stipulated that any prevalence study is at high risk of bias if the response rate is less than 75%, with risk of bias increasing when studies do not statistically compare responders and non-responders. Only two studies were deemed to be at low risk of response rate related bias: Subramaniam et al., (2014) achieved a response rate of over 75% and Cath et al., (2017) compared responders and non-responders to show no differences. Two studies (Bulli et al., 2014; Zilhão et al., 2016) failed to report response rates, and did not report sufficient detail for the response rate to be calculated. The mean response rate was 53.25%. Another significant potential source of bias was how representative the study participants were of the population. Half of the studies (6/12) were at high risk of bias with regards to this concern (e.g. female participants in the studies ranged from 54.9-89.3%, suggesting a bias towards majority female samples).

3.4 Covariate analysis

Moderator analysis indicated no effect for prevalence type (lifetime, point), $Q_{between} = 0.285$, df = 1, p = 0.593. Moderator analysis for study quality (overall risk of bias score, 2 levels: low, moderate) was non-significant, $Q_{between} = 0.113$, df = 1, p = 0.736, as was the moderator analysis for "method of data collection" (3 levels: self-report survey, self-report with assistance and clinical interview), $Q_{between} = 4.524$, df = 2, p = 0.104. Meta-regression indicated non-significant effects for response rate (coefficient = 0.5973, Q = 0.10, p = 0.7516, $Tau^2 = 0.4585$), gender (coefficient = -0.4805, Q = 0.05, p = 0.8179, $Tau^2 = 0.3837$) and year of publication (coefficient = -0.1164, Q = 3.04, p = 0.0811, $Tau^2 = 0.4440$).

4.0 Discussion

The aim of this review was to conduct a comprehensive, systematic literature review to identify relevant studies that have reported prevalence data for adult HD, to summarise the characteristics of these studies and then calculate a pooled estimate of the prevalence of HD using meta-analytic techniques. Through the systematic review process, eleven studies were identified, reporting ten point HD prevalence estimates and two lifetime HD prevalence estimates. The pooled point prevalence estimate found for HD was 2.6%, 95% confidence interval: [1.7-3.7%], and the pooled lifetime prevalence HD estimate found was 1.7%, 95% confidence interval: [0.4-6.8%]. There was no significant difference between the pooled lifetime and pooled point prevalence estimates. The overall pooled prevalence estimate was therefore 2.5%, 95% confidence interval: [1.7-3.6%]. The potential for publication bias influencing this result was identified via an asymmetrical funnel plot. There are several other causes of plot asymmetry such as differences between the methodologies employed (Terrin et al., 2005), chance and the selection of assessment measures (Tang & Liu, 2000). Evidence of heterogeneity is a primary concern in any prevalence meta analyses, because it may be

highlighting that the health condition of interest may actually vary between studies (Barendregt et al. 2003).

The quality of the studies included in the current review was however relatively high, with only one study (López-Solà et al., 2014) scoring moderate on overall risk of bias, and the remaining studies scoring low overall risk of bias. The studies analysed were generally methodologically rigorous in their use of appropriate sampling frames, collecting data directly from participants and using appropriate prevalence periods and case definition methods. Inclusion criteria also demanded a large minimal sample size in the original studies based on an expected HD prevalence of 1.5% and this ensured the meta analysis of relatively large prevalence studies. Nine studies did not meet criteria for inclusion because their sample sizes were inadequate. However, response rates within the studies analysed were an issue, as it is recommended that in prevalence studies the response rate should be at least 75% (Hoy et al., 2012). Inspection of the relevant risk of bias item revealed a large proportion (11/12) of samples failed to meet recommended response rate criteria.

No statistically significant effects were found for the method of HD assessment, study quality, response rate, gender and year of publication. However, although Cochran's Q often has high power for detecting statistical tests of main effects, it is often underpowered when used for moderator analyses (Hedges & Pigott, 2004). Thompson and Higgins (2002) concluded that when the number of studies included in meta analytic reviews is low, the potential for robust conclusions based on meta-regression become limited. Given the small number of prevalence studies in the current review, it is therefore difficult to determine whether there truly was no moderation effect, or whether the non-significant findings were due to the meta-regression being underpowered. With this caveat in mind, the lack of any gender differences in HD prevalence is an interesting finding, considering the previous marked uncertainty around this issue in the literature. In epidemiological samples, studies are

divided on whether hoarding symptoms are more common in one sex versus another (Pertusa et al 2010). However, our results suggest no gender differences in prevalence.

Studies differed in terms of the assessment methods used to detect HD. Six of the twelve prevalence estimates were based on case detection using self-report methods and three used clinical interviewing methods (Fullana et al., 2010; Nordsletten, Reichenberg, et al., 2013; Subramaniam et al., 2014). Whilst studies based on interviewing do provide diagnostic accuracy, they are not without limitations in an HD context. For example, door-to-door recruitment is a standard recruitment strategy in epidemiology. The suitability of this method for accessing the HD population is questionable when considering the possible concerns (e.g., eviction) and practical constraints (e.g., ability to move through the clutter) which may limit willingness (or actual ability) to opening doors to researchers. Thus, while self-report based studies may overestimate HD prevalence, door-to-door studies may underestimate HD prevalence. In terms of the reliability and validity of the measures used to assess for HD, then any small error applied over the datasets used may have produced a relevant and non-negligible number of cases that were falsely classified in the original studies. Therefore, the overall prevalence rate produced may have been affected by the number of undetected "false positives" in the original studies. The current meta-analysis identified a low prevalence base rate for HD. Any mental health disorder with a relatively low base rate is also prone to yielding a high "false positive" rate that can often exceed the false negative rate (Baldessarini, Finklestein, & Arana, 1983).

Self-criticism and shame have also been shown to be associated with HD (Chou et al., 2018) and it has been suggested that this may be in response to feeling "personally defective" due to the high levels of clutter in the home (Weingarden & Renshaw, 2015). High levels of shame may discourage people with HD from participating in research due to social desirability response bias (Huang, Liao, & Chang, 1998). Additionally, insight can vary

greatly amongst people with HD, with a large proportion judged to have poor insight (Kim et al., 2001; Tolin, Fitch, Frost, & Steketee, 2010). Individuals with poor insight may be unreliable sources of HD hoarding prevalence. It is also worth considering that the sample of studies in the current review were conducted predominantly in developed Western countries and this precluded combining and comparing estimates from different regions of the world (Fiest et al. 2014). Therefore, this study has generated a pooled prevalence estimate of HD which is specific (and therefore limited to) developed nations. Research into HD in non-Western countries has begun, with results indicating that the core features of HD appear stable across cultures (Nordsletten et al., 2018).

The current review focused on the adult prevalence and did not consider the impact of age on HD prevalence. A single study of HD prevalence in older adults was excluded and the study also did not contain a large enough sample. The decision to exclude older adult studies in the systematic review process was based on the motivation to make the meta-analysis as specific as possible to working age adults. It also would subsequently not have been possible to complete any moderator analyses due to the small number of older adult studies (Thompson & Higgins, 2002). Research suggests that hoarding symptoms may begin in childhood and adolescence (Grisham et al., 2006), with severity of symptoms potentially increasing with age (Ayers, Saxena, Golshan, & Wetherell, 2010). Cath et al. (2017) found that in a sample of 15,194 participants, hoarding severity increased reliably with age, beginning at around age 30-35, with the highest prevalence rates being amongst individuals aged over 65.

The findings carry implications for the design of future HD prevalence studies. Often the type of prevalence being assessed (i.e. point, period, or lifetime) was not explicitly stated and so future studies should explicitly state the type of prevalence assessed. Consistent attempts should be made to maximise response rates. Strategies that have been shown to be

effective in improving self-report response rates include providing monetary incentives, personalising questionnaires and letters, using colour ink, and sending surveys by recorded delivery (Edwards et al., 2002). Studies should always conduct comparisons of responders and non-responders and consistently report the findings of these analyses (Hoy et al., 2012). Future prevalence studies should seek to recruit a sufficient number of participants based on a priori calculations, such as the one conducted in this review. Given the pooled HD prevalence estimate of 2.5% (95% confidence interval: [1.7-3.6%]) reported in this study, studies should seek to recruit at least N = 599 participants. Ideally, studies should seek to recruit N = 889 participants (calculation based on the lower limit of the 95% confidence interval, i.e. 1.7% prevalence). Had this minimal sample size been used as the inclusion criteria for this review, an additional paper would have been eligible. Samuels et al. (2008) reported an unweighted prevalence of 3.7% in a sample of N = 735 participants. This study was however similarly methodologically limited to the studies included in the review, as a response rate was only 59% and the sample was biased towards females (63.3%). Finally, it is recommended that future prevalence studies collate and analyse HD prevalence data by participant age bands, and again this would demand planning to collect large samples.

5.0 Conclusions

The results of this review indicate that the prevalence of HD appears relatively low and consistent across a range of Western/developed countries and there was no difference between point and lifetime prevalence HD rates. The pooled prevalence estimate of HD in the populations studied was 2.5%. There was however significant variation between studies in terms of response rates, location, gender proportions and assessment methods used. The analysis of the influence of these differences was potentially underpowered, due to the small number of studies analysed. Although this review suggests more than 2 in 100 people in the community might meet diagnostic criteria for HD, people with HD may not participate in

epidemiological research due to lack of insight or shame. Future HD prevalence studies need to plan for large samples (N > 889), clearly define the type of prevalence being assessed, triangulate assessment methods (i.e. use diagnostic interviewing, valid and reliable self-report measures and assessments of the home environment) and report comparisons of responders and non-responders. The need for further research into the prevalence of HD in developing countries and across different age groups is now indicated.

References

American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: American Psychiatric Association.

Andrews, G., Henderson, S., & Hall, W. (2001). Prevalence, comorbidity, disability and service utilisation: Overview of the Australian National Mental Health Survey. The British Journal of Psychiatry, 178, 145-153. doi:10.1192/bjp.178.2.145

Ayers, C. R., Dozier, M. E., & Mayes, T. L. (2017). Psychometric evaluation of the Saving Inventory-Revised in older adults. Clinical Gerontologist, 40, 191-196. doi:10.1080/07317115.2016.1267056

Ayers, C. R., Iqbal, Y., & Stricklanda, K. (2014). Medical conditions in geriatric hoarding disorder patients. Aging & Mental Health, 18, 148-151. doi:10.1080/13607863.2013.814105

Ayers, C. R., Saxena, S., Golshan, S., & Wetherell, J. L. (2010). Age at onset and clinical features of late life compulsive hoarding. International Journal of Geriatric Psychiatry, 25, 142-149. doi:10.1002/gps.2310

Baldessarini, R. J., Finklestein, S., & Arana, G. W. (1983). The predictive power of diagnostic tests and the effect of prevalence of illness. Archives of General Psychiatry, 40, 569-573. doi:10.1001/archpsyc.1983.01790050095011

Barendregt, J.J., Doi, S.A., Lee, Y.Y., Norman, R.E. & Vos, T. (2003). Meta-analyis of prevalence. Epidemiological Community Health, 67, 974-978. doi: 10.1136/jech-2013-203104

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2018). Comprehensive Meta Analysis Version 3.0. www.meta-analysis.com

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. Research Synthesis Methods, 1, 97-111. doi:10.1002/jrsm.12

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). When Does it Make Sense to Perform a Meta-Analysis? In M. Borenstein, L. V. Hedges, J. P. T. Higgins, & H. R. Rothstein (Eds.) Introduction to Meta-Analysis. New York, NY: John Wiley & Sons Ltd.

*Bulli, F., Melli, G., Carraresi, C., Stopani, E., Pertusa, A., & Frost, R. O. (2014). Hoarding behaviour in an Italian non-clinical sample. Behavioural and Cognitive Psychotherapy, 42, 297-311. doi:10.1017/S1352465812001105

Carneiro, I., & Natasha, H. (2011). Introduction To Epidemiology (2nd ed.). New York, NY: Cambridge University Press

*Cath, D. C., Nizar, K., Boomsma, D., & Mathews, C. A. (2017). Age-specific prevalence of hoarding and obsessive compulsive disorder: a population-based study. American Journal of Geriatric Psychiatry, 25, 245-255. doi:10.1016/j.jagp.2016.11.006

Chou, C. Y., Tsoh, J., Vigil, O., Bain, D., Uhm, S. Y., Howell, G., . . . Mathews, C. A. (2018). Contributions of self-criticism and shame to hoarding. Psychiatry Research, 262, 488-493. doi:10.1016/j.psychres.2017.09.030

Daniel, W. W. (1999). Biostatistics: A Foundation for Analysis in the Health Sciences (7th ed.). New York: John Wiley & Sons.

Drury, H., Ajmi, S., Fernández de la Cruz, L., Nordsletten, A.E., & Mataix-Cols, D. (2014). Caregiver burden, family accommodation, health and well-being in relatives of individuals with hoarding disorder. Journal of Affective Disorders, 159, 7-14. doi: 10.1016/j.jad.2014.01.023.

Edwards, P., Roberts, I., Clarke, M., DiGuiseppi, C., Pratap, S., Wentz, R., & Kwan, I. (2002). Increasing response rates to postal questionnaires: systematic review. BMJ, 324, 1183. doi:10.1136/bmj.324.7347.1183

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. BMJ, 315, 629-634. doi:10.1136/bmj.315.7109.629

Fox, D. M. (2005). Evidence of evidence-based health policy: the politics of systematic reviews in coverage decisions. Health Affairs, 24, 114-122. doi:10.1377/hlthaff.24.1.114

Fiest, K.M., Pringham, T., Pattern, S.B., Svenson, S.B., & Jette, N. (2014). The role of systematic reviews and meta-analysis of incidence and prevalance studies in neuroepidemiology. Neuroepidemiology, 42, 16-24. doi: 10.1159/0003555533

Frost, R. O., & Gross, R. C. (1993). The hoarding of possessions. Behaviour Research and Therapy, 31, 367-381. doi:10.1016/0005-7967(93)90094-B

Frost, R. O., Steketee, G., & Grisham, J. (2004). Measurement of compulsive hoarding: Saving Inventory-Revised. Behaviour Research and Therapy, 42, 1163-1182. doi:10.1016/j.brat.2003.07.006

Frost, R. O., Steketee, G., & Tolin, D. F. (2011). Comorbidity in hoarding disorder. Depression and Anxiety, 28, 876-884. doi:10.1002/da.20861

Frost, R. O., Steketee, G., Tolin, D. F., & Renaud, S. (2008). Development and validation of the Clutter Image Rating. Journal of Psychopathology and Behavioral Assessment, 30, 193-203. doi:10.1007/s10862-007-9068-7

Frost, R. O., Steketee, G., & Williams, L. (2000). Hoarding: a community health problem. Health and Social Care in the Community, 8, 229-234. doi:j.1365-2524.2000.00245.x

Frost, R. O., Steketee, G., Williams, L. F., & Warren, R. (2000). Mood, personality disorder symptoms and disability in obsessive compulsive hoarders: a comparison with clinical

and nonclinical controls. Behaviour Research and Therapy, 38, 1071-1081. doi:10.1016/s0005-7967(99)00137-0

*Fullana, M. A., Vilagut, G., Rojas-Farreras, S., Mataix-Cols, D., de Graaf, R., Demyttenaere, K., . . . investigators, E. M. (2010). Obsessive-compulsive symptom dimensions in the general population: results from an epidemiological study in six European countries. Journal of Affective Disorders, 124, 291-299. doi:10.1016/j.jad.2009.11.020

Grisham, J. R., Frost, R. O., Steketee, G., Kim, H.-J., & Hood, S. (2006). Age of onset of compulsive hoarding. Journal of Anxiety Disorders, 20, 675-686. doi:10.1016/j.janxdis.2005.07.004

Hayward, L. C., & Coles, M. E. (2009). lucidating the relation of hoarding to obsessive compulsive disorder and impulse control disorders. Journal of Psychopathology and Behavioral Assessment, 31, 220-227. doi:10.1007/s10862-008-9106-0

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. Psychological Methods, 9, 426-445. doi:10.1037/1082-989X.9.4.426

Higgins, J. P., & Green, S. (2011). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. Retrieved from handbook.cochrane.org

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

Hoy, D., Brooks, P., Woolf, A., Blyth, F., March, L., Bain, C., . . . Buchbinder, R. (2012). Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. Journal of Clinical Epidemiology, 65, 934-939. doi:10.1016/j.jclinepi.2011.11.014

Huang, C., Liao, H., & Chang, S. (1998). Social desirability and the clinical self-report inventory: Methodological reconsideration. Journal of Clinical Psychology, 54, 517-528. doi:10.1002/(sici)1097-4679(199806)54:4<517::aid-jclp13>3.0.co;2-i

*Iervolino, A. C., Perroud, N., Fullana, M. A., Guipponi, M., Cherkas, L., Collier, D. A., & Mataix-Cols, D. (2009). Prevalence and heritability of compulsive hoarding: a twin study. The American Journal of Psychiatry, 166, 1156-1161. doi:10.1176/appi.ajp.2009.08121789

*Ivanov, V. Z., Nordsletten, A., Mataix-Cols, D., Serlachius, E., Lichtenstein, P., Lundström, S., . . . Rück, C. (2017). Heritability of hoarding symptoms across adolescence and young adulthood: A longitudinal twin study. PLoS One, 12, e0179541. doi:10.1371/journal.pone.0179541

Kim, H.-J., Steketee, G., & Frost, R. O. (2001). Hoarding by elderly people. Health & Social Work, 26, 176-184. doi:10.1093/hsw/26.3.176

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15, 155-163. doi:10.1016/j.jcm.2016.02.012

Loranger, A. W., Sartorius, N., Andreoli, A., Berger, P., Buchheim, P., Channabasavanna, S. M., . . . Regier, D. A. (1994). The International Personality Disorder Examination: The World Health Organization/Alcohol, Drug Abuse, and Mental Health Administration International Pilot Study of Personality Disorders. Archives of General Psychiatry, 51, 215-224. doi:10.1001/archpsyc.1994.03950030051005

Lwanga, S. K., & Lemeshow, S. (1991). Sample Size Determination in Health Studies. Geneva: World Health Organization.

*López-Solà, C., Fontenelle, L. F., Alonso, P., Cuadras, D., Foley, D. L., Pantelis, C., . . . Harrison, B. J. (2014). Prevalence and heritability of obsessive-compulsive spectrum

and anxiety disorder symptoms: A survey of the Australian Twin Registry. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 165B, 314-325. doi:10.1002/ajmg.b.32233

Mansfield, K.E., Sim, J., Jordan, J.L., & Jordan, K.P. (2016). A systematic review and meta-analysis of the prevalence of chronic widespread pain in the general population. Pain, 157, 55-64. doi: org/10.1097/jpain.00000000000314

Maraz, A., Griffiths, M. D., & Demetrovics, Z. (2016). The prevalence of compulsive buying: a meta-analysis. Addiction, 111, 408-419. doi:10.1111/add.13223

Mataix-Cols, D. & Fernández de la Cruz, L. (2018). Hoarding Disorder has finally arrived, but many challenges lie ahead. World Psychiatry, 17(2), 224-225.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: the PRISMA statement. PLoS Medicine(6), e1000097. doi:10.1371/journal.pmed1000097

*Mueller, A., Mitchell, J. E., Crosby, R. D., Glaesmer, H., & de Zwaan, M. (2009). The prevalence of compulsive hoarding and its association with compulsive buying in a German population-based sample. Behaviour Research and Therapy, 47, 705-709. doi:10.1016/j.brat.2009.04.005

Naing, L., Winn, T., & Rusli, B. N. (2006). Practical Issues in Calculating the Sample Size for Prevalence Studies. Archives of Orofacial Sciences, 1, 9-14.

National Institute of Mental Health. (n.d.). What is Prevalence? Retrieved from www.nimh.nih.gov/health/statistics

Neave, N., Caiazza, R., Hamilton, C., McInnes, L., Saxton, T. K., Deary, V., & Wood, M. (2017). The economic costs of hoarding behaviours in local authority/housing association tenants and private home owners in the north-east of England. Public Health, 148, 137-139. doi:10.1016/j.puhe.2017.04.010

Nordsletten, A. E., Fernández de la Cruz, L., Pertusa, A., Reichenberg, A., Hatch, S. L., & Mataix-Cols, D. (2013). The Structured Interview for Hoarding Disorder (SIHD): development, usage and further validation. Journal of Obsessive-Compulsive and Related Disorders, 2, 346-350. doi:10.1016/j.jocrd.2013.06.003

Nordsletten, A. E., Fernández de la Cruz, L., Aluco, E., Alonso, P., López-Solà, C., Menchón, J. M., . . . Mataix-Cols, D. (2018). A transcultural study of hoarding disorder: Insights from the United Kingdom, Spain, Japan, and Brazil. Transcultural Psychiatry, 55, 261-285. doi:10.1177/1363461518759203

*Nordsletten, A. E., Reichenberg, A., Hatch, S. L., Fernández de la Cruz, L., Pertusa, A., Hotopf, M., & Mataix-Cols, D. (2013). Epidemiology of hoarding disorder. The British Journal of Psychiatry, 203, 445-452. doi:10.1192/bjp.bp.113.130195

Pertusa, A., Frost, R.O., Fullana, M.A., Samuels, J., Steketee, G., Tolin, D., Saxena, S., Leckman, J.F. & Mataix-Cols, D. (2010). Refining the diagnostic boundaries of compulsive hoarding: A critical review. Clinical Psychology Review, 30(4), 371-386.

Samuels, J. F., Bienvenu, O. J., Grados, M. A., Cullen, B., Riddle, M. A., Liang, K. Y., . . . Nestadt, G. (2008). Prevalence and correlates of hoarding behavior in a community-based sample. Behaviour Research and Therapy, 46, 836-844. doi:10.1016/j.brat.2008.04.004

Steketee, G., & Frost, R. O. (2014). Phenomenology of hoarding. In R. O. Frost & G. Steketee (Eds.), The Oxford Handbook of Hoarding and Acquiring. New York, NY: Oxford University Press.

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., . . . Higgins, J. P. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ, 343, d4002. doi:10.1136/bmj.d4002

Storch, E. A., Rahman, O., Park, J. M., Reid, J., Murphy, T. K., & L ewin, A. B. (2011). Compulsive hoarding in children. Journal of Clinical Psychology, 67, 507-516. doi:10.1002/jclp.20794

Stroup, D. F., Berline, J. A., Olkin, I., Williamson, G. D., Rennie, D., Moher, D., . . . Thacker, S. B. (2000). Meta-analysis of observational studies in epidemiology: a proposal for reporting. JAMA, 283, 2008-2012. doi:10.1001/jama.283.15.2008

*Subramaniam, M., Abdin, E., Vaingankar, J. A., Picco, L., & Chong, S. A. (2014). Hoarding in an Asian population: prevalence, correlates, disability and quality of life. ANNALS Academy of Medicine Singapore, 43, 535-543. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/25523857

Tang, J. L., & Liu, J. L. (2000). Misleading funnel plot for detection of bias in meta-analysis. Journal of Clinical Epidemiology, 53, 477-484. doi:10.1016/s0895-4356(99)00204-8

Taylor, J. B., Goode, A. P., George, S. Z., & Cook, C. E. (2014). Incidence and risk factors for first-time incident low back pain: a systematic review and meta-analysis. The Spine Journal, 14, 2299-2319. doi:10.1016/j.spinee.2014.01.026

Terrin, N., Schmid, C. H., & Lau, J. (2005). In an empirical evaluation of the funnel plot, researchers could not visually identify publication bias. Journal of Clinical Epidemiology, 58, 894-901. doi:10.1016/j.jclinepi.2005.01.006

Thomas, R., Sanders, S., Doust, J., Beller, E., & Glasziou, P. (2015). Prevalence of Attention-Deficit/Hyperactivity Disorder: a systematic review and meta-analysis. Pediatrics, 135. doi:10.1542/peds.2014-3482

Thompson, S. G., & Higgins, J. P. (2002). How should meta-regression analyses be undertaken and interpreted? Statistics in Medicine, 21, 1559-1573. doi:10.1002/sim.1187

Timpano, K. R., Broman-Fulks, J. J., Glaesmer, H., Exner, C., Rief, W., Olatunji, B. O., . . . Schmidt, N. B. (2013). A taxometric exploration of the latent structure of hoarding. Psychological Assessment, 25, 194-203. doi:10.1037/a0029966

*Timpano, K. R., Exner, C., Glaesmer, H., Rief, W., Keshaviah, A., Brähler, E., & Wilhelm, S. (2011). The epidemiology of the proposed DSM-5 hoarding disorder: exploration of the acquisition specifier, associated features, and distress. The Journal of Clinical Psychiatry, 72, 780-786. doi:10.4088/JCP.10m06380

Tolin, D. F., Fitch, K. E., Frost, R. O., & Steketee, G. (2010). Family informants' perceptions of insight in compulsive hoarding. Cognitive Therapy and Research, 34, 69-81. doi:10.1007/s10608-008-9217-7

Tolin, D. F., Frost, R. O., & Steketee, G. (2010). A brief interview for assessing compulsive hoarding: the Hoarding Rating Scale-Interview. Psychiatry Res, 178, 147-152. doi:10.1016/j.psychres.2009.05.001

Tolin, D. F., Frost, R. O., Steketee, G., Gray, K. D., & Fitch, K. E. (2008). The economic and social burden of compulsive hoarding. Psychiartry Research, 160, 200-211. doi:10.1016/j.psychres.2007.08.008

Tolin, D. F., Meunier, S. A., Frost, R. O., & Steketee, G. (2010). Course of compulsive hoarding and its relationship to life events. Depression and Anxiety, 27, 829-838. doi:10.1002/da.20684

Webb, P., & Bain, C. (2011). Essential Epidemiology An Introduction for Students and Health Professionals. New York, NY: Cambridge University Press.

Weingarden, H., & Renshaw, K. D. (2015). Shame in the obsessive compulsive related disorders: a conceptual review. Journal of Affective Disorders, 171, 74-84. doi:10.1016/j.jad.2014.09.010

Wittchen, H. U. (1994). Reliability and validity studies of the WHO-Composite International Diagnostic Interview (CIDI): a critical review. Journal of Psychiatric Research, 28, 57-84. doi:10.1016/0022-3956(94)90036-1

Worden, B. L., DiLoreto, J., & Tolin, D. F. (2014). Insight and Motivation. In R. O. Frost & G. Steketee (Eds.), The Oxford Handbook of Hoarding and Acquiring. New York, NY: Oxford University Press.

*Zilhão, N. R., Smit, D. J., Boomsma, D. I., & Cath, D. C. (2016). Cross-disorder genetic analysis of tic disorders, obsessive-compulsive, and hoarding symptoms. Frontiers in Psychiatry, 7, 120. doi:10.3389/fpsyt.2016.00120

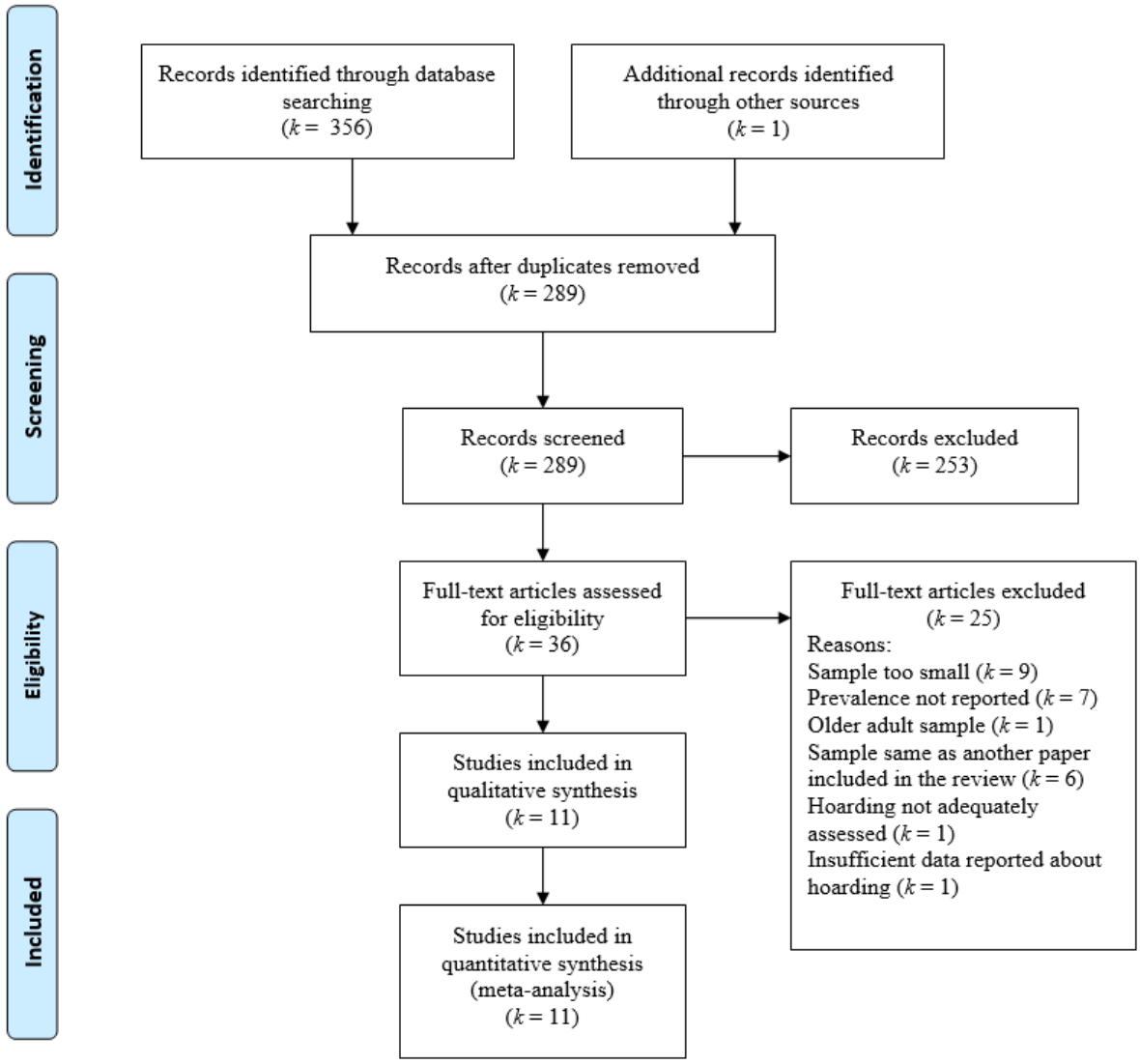

Figure 1. PRISMA diagram

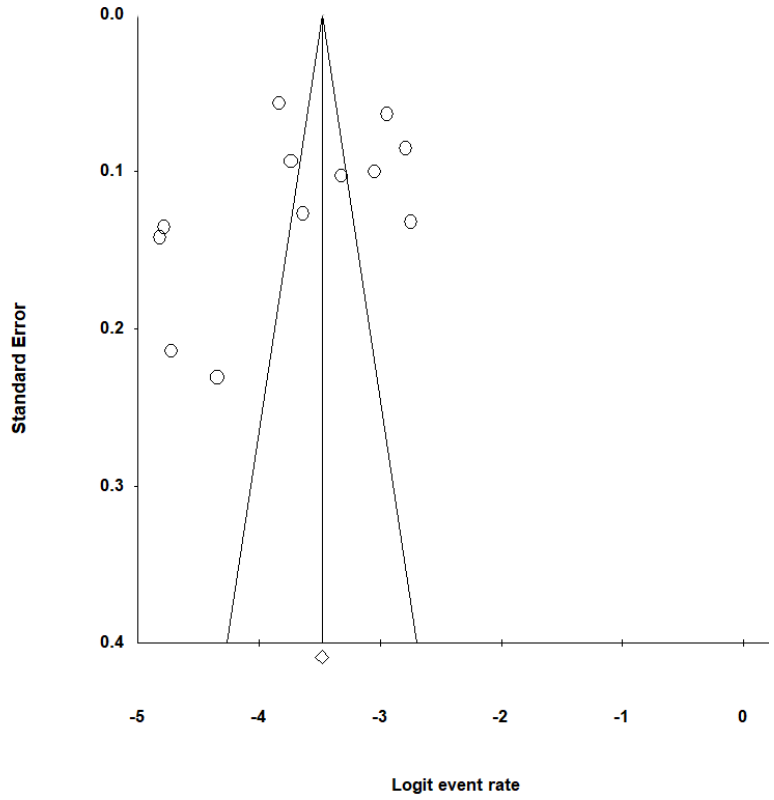

Figure 2. Funnel plot distribution of standard error

Table 1

Study characteristics

Reference	Authors (Year)	Country	Data collection method	N	Response rate (%)	Sample mean age (range)	Female (%)	Hoarding assessment instrument	Prevalence type	Hoarding prevalence (%)
1	Bulli, Melli, Carraresi, & Stopani (2014)	Italy	Self-report survey	1012	NR	36.6 (18-84)	62.7	SI-R	Point	6.03
2	Cath, Nizar, Boomsma, & Mathews (2017)	Netherlands	Self-report survey	15194	45	NR	64	HRS-SR	Point	2.12
3	Iervolino et al. (2009)	UK	Self-report survey	5022	60.41	55.5 (17-86)	89.3	HRS-SR	Point	2.3
4	López-Solà et al. (2014)	Australia	Self-report survey	2495	35.9	NR	58.8	HRS-SR	Point	2.57
5	Mueller et al. (2009)	Germany	Self-report with assistance	2307	61.9	NR	54.9	German Compulsive Hoarding Inventory (adapted SI-R)	Point	4.55
6	Nordsletten et al. (2013)	UK	Interview	1482	51.9	NR	56.5	SIHD, MINI, HRS-SR	Point	1.3
7	Subramaniam et al. (2014)	Singapore	Interview	6616	75.9	NR	NR	CIDI	Lifetime	0.8
8	Timpano et al. (2011)	Germany	Self-report with assistance	2512	54.25	48.8 (14-94)	55.8	German Hoarding Rating Scale and DSM criteria	Point	5.8
9	Zilhão, Smit, Boomsma, & Cath (2016)	Netherlands	Self-report survey	5221	NR	33.61	NR	HRS-SR	Point	5
10	Fullana et al. (2010)	Belgium, France, Germany, Italy, Netherlands, Spain	Interview	2804	61.2	NR	58.9	Single question in the CIDI	Lifetime	3.5
11	Ivanov et al. (2017)a	Sweden	Self-report survey	2495	48	18	58	HRS-SR	Point	0.9
12	Ivanov et al. (2017)b	Sweden	Self-report survey	6218	38	23.8 (20-28)	61	HRS-SR	Point	0.8

CIDI = Composite International Diagnostic Interview; HRS-SR = Hoarding Rating Scale - Self Report; MINI = Mini International Neuropsychiatric Interview;

SIHD = Structured Interview for Hoarding Disorder; SI-R = Saving Inventory Revised. NR = Not reported.

Table 2

Forest plot of HD prevalence estimates

Study	Event rate	Lower Limit	Upper Limit
Subramaniam et al. (2014)	0.008	0.006	0.011
Fullana et al. (2010)	0.035	0.029	0.042
Lifetime	**0.017**	**0.004**	**0.068**
Bulli, Melli, Carraresi, & Stopani (2014)	0.060	0.047	0.077
Cath, Nizar, Boomsma, & Mathews (2017)	0.021	0.019	0.024
Iervolino, et al. (2009)	0.023	0.019	0.028
López-Solà et al. (2014)	0.026	0.020	0.033
Mueller et al. (2009)	0.046	0.038	0.055
Nordsletten et al. (2013)	0.013	0.008	0.020
Timpano et al. (2011)	0.058	0.050	0.068
Zilhão, Smit, Boomsma, & Cath (2016)	0.050	0.044	0.056
Ivanov et al. (2017)a	0.009	0.006	0.013
Ivanov et al. (2017)b	0.008	0.006	0.011
Point	**0.026**	**0.017**	**0.037**
Overall	**0.025**	**0.017**	**0.036**

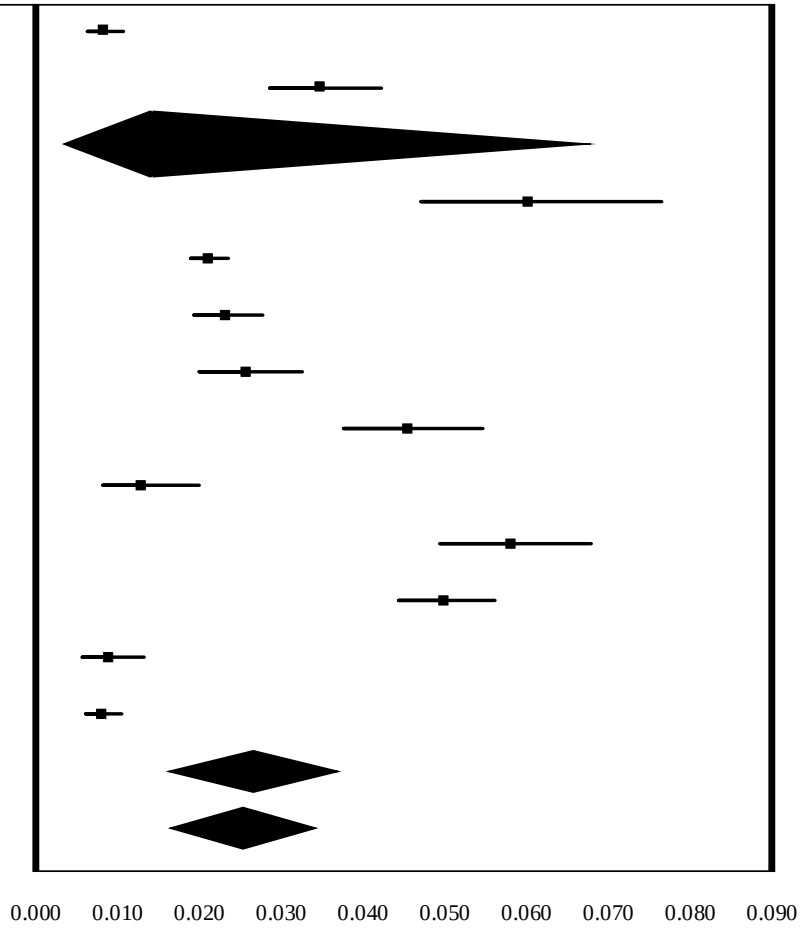

Table 3

Risk of bias ratings for each HD prevalence study

Study	Target population	Sampling frame	Sample selection	Response rate	Information collected direct from subject	Case definition	Valid instrument	Consistent mode of collection	Prevalence period	Errors in reporting	Overall rating
Bulli et al. (2014)	✓	✓	✗	✗	✓	✓	✓	✗	✓	✓	Low
Cath et al. (2017)	✓	✓	✗	✓	✓	✓	✓	✓	✓	✓	Low
Iervolino et al. (2009)	✗	✓	✓	✗	✓	✓	✓	✓	✓	✓	Low
López-Solà et al. (2014)	✗	✓	✗	✗	✓	✓	✓	✗	✓	✓	Moderate
Mueller et al. (2009)	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	Low
Nordsletten et al. (2013)	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	Low
Subramaniam et al. (2014)	✓	✓	✓	✓	✓	✓	✗	✓	✓	✓	Low
Timpano et al. (2011)	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	Low
Zilhão et al. (2016)	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	Low
Fullana et al. (2010)	✓	✓	✓	✗	✓	✗	✗	✓	✓	✓	Low
Ivanov et al. (2017)a	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	Low
Ivanov et al. (2017)b	✓	✓	✓	✗	✓	✓	✓	✓	✓	✓	Low

Tick indicates risk of bias criteria met therefore low risk of bias; cross indicates risk of bias criteria not met therefore high risk of bias

Figure 3. Summary of risk of bias across all included HD prevalence study samples