HHS Public Access

Author manuscript

Suicide Life Threat Behav. Author manuscript; available in PMC 2022 August 01.

Published in final edited form as:

Suicide Life Threat Behav. 2021 August ; 51(4): 673–683. doi:10.1111/sltb.12738.

Introducing the Psychological Autopsy Methodology Checklist

Kenneth R. Conner, PsyD, MPH[1,2], **Benjamin P. Chapman, PhD, MPH**[2,3], **Annette L. Beautrais, PhD**[4], **David A. Brent, MD**[5], **Jeffrey A. Bridge, PhD**[6], **Yeates Conwell, MD**[2], **Tyler Falter**[1], **Amanda Holbrook, BS**[1], **Barbara Schneider, MD, MSc**[7]

[1]Department of Emergency Medicine, University of Rochester Medical Center, Rochester, NY, USA

[2]Department of Psychiatry, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA

[3]Department of Public Health Sciences, University of Rochester School of Medicine and Dentistry, Rochester, NY, USA

[4]School of Health Sciences, University of Canterbury, Christchurch, New Zealand

[5]Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, USA

[6]Department of Pediatrics, The Ohio State University, Columbus, OH, USA

[7]Department of Psychiatry, Psychosomatics, and Psychotherapy, Centre of Psychiatry, Johann Wolfgang Goethe-University, Frankfurt/Main, Germany

Abstract

Objective: Case–control psychological autopsy studies are the research standard for the postmortem, quantitative study of ongoing or recent risk factors for suicide. We aimed to develop a reliable checklist of methodological quality of these studies.

Method: We adapted items from a validated checklist to address general methodological elements and created novel items to address the unique aspects of psychological autopsy research to generate a 16-item checklist assessing reporting, external validity, internal validity, and power. We used percent agreement and kappa to evaluate inter-rater reliability of the items and overall checklist based on independent ratings of 26 case–control psychological autopsy studies conducted internationally. We also summed the items to generate overall quality ratings, assessing internal consistency with coefficient alpha (α).

Results: Inter-rater reliability for the overall checklist was high (percent agreement, 86.5%) and that based conservatively on kappa was substantial (κ .71) whereas internal consistency was low (α = 0.56). The inter-rater reliability of the individual items showed acceptable to high agreement.

Conclusion: A novel checklist provides a reliable means to assess the methodological quality of specific elements of quantitative case–control psychological autopsy studies, providing detailed

Correspondence: Amanda Holbrook, Department of Emergency Medicine, University of Rochester Medical Center, Rochester, NY, USA. Amanda_Holbrook@URMC.Rochester.edu, apholbrook@yahoo.com.

guidance in planning such studies. Lower internal consistency may limit its utility as a summary measure of study quality.

INTRODUCTION

The psychological autopsy was developed to examine the psychological and contextual circumstances preceding suicide through use of interviews with proxy respondents (i.e., informants) of suicide decedents, along with other sources of data as available (e.g., records) (Litman et al., 1963; Robins et al., 1959). Subsequently, case–control methodology established for epidemiological research of rare outcomes (Rothman & Greenland, 1998; Vandenbroucke et al., 2007) was introduced and adapted for psychological autopsy investigations, providing a means to facilitate systematic comparisons between suicide decedents (cases) and a non-suicide reference group (controls). Case–control designs have remained the research standard for quantitative psychological autopsy research and have been used most widely in Asia, Australasia, Europe, and North America, along with a small number of reports from other regions. These studies have generated population specific risk estimates for a range of risk factors including mental disorders, stressful life events, physical illness burden, access to lethal means including firearms and pesticides, and local cultural contextual factors, with mental disorders being most widely reported including in systematic reviews and meta-analyses (Arsenault-Lapierre et al., 2004; Cavanagh et al., 2003; Conner et al., 2019; Yoshimasu et al., 2008).

Clear guidelines for conducting, reporting, and assessing the methodological quality of case–control epidemiological studies are available in the form of expert recommendations (Rothman & Greenland, 1998), consensus guidelines (Vandenbroucke et al., 2007), and checklists of methodological quality (Downs & Black, 1998). However, these general resources do not provide guidance for navigating the special circumstances of psychological autopsy research studies. This gap has been filled by several published papers and book chapters that offer recommendations for conducting psychological autopsy research studies (Beskow et al., 1990; Brent, 1989; Clark & Horton-Deutsch, 1992; Conner et al., 2011, 2012; Ebert, 1987; Hawton et al., 1998; Isometsä, 2001; Kõlves et al., 2021; Pouliot & De Leo, 2006; Schneidman, 1981; Shaffer et al., 1972). Limitations of this literature include that the range of recommendations and critiques can be overwhelming and are at times contradictory; there is not the same level of clarity or consensus as with the more general epidemiological literature on case–control study methodology; and we are aware of no reliable, published method for assessing the methodological quality of psychological autopsy research studies. The purpose of this study was to address the latter limitation by developing a reliable checklist of methodological quality of quantitative case–control psychological autopsy studies that may provide guidance in designing, reporting, and assessing these studies, and was in the service of increased transparency and standardization using this methodology.

METHOD

Scale development

The Psychological Autopsy Methodology Checklist (PAMC) was developed to address basic elements in case–control research methodology as well as the unique methods of psychological autopsy research. Standard elements are described in a validated checklist for studies using case–control designs and related methods developed by Downs and Black (1998) including reporting, external validity, and internal validity, with several items adapted from their report. Additional PAMC items were developed to consider specialized aspects of case–control psychological autopsy research guided by the authors collective experience and two reviews of case–control psychological autopsy research methodology (Conner et al., 2011, 2012). Studies contained in a meta-analysis of case–control psychological autopsy research of mood and substance use disorders (Conner et al., 2017) provided 35 reports that were used to refine the PAMC and establish its inter-rater reliability (Almasi et al., 2009; Appleby et al., 1999; Beautrais, 2001; Brent et al., 1999; Brent et al., 1993; Chan et al., 2009; Chen et al., 2006; Cheng, 1995; Cheng et al., 2000; Chiu et al., 2004; Conner et al., 2003; Conwell et al., 2010; De Leo et al., 2013a; De Leo et al., 2013b; Foster et al., 1999; Freuchen et al., 2012; Harwood et al., 2001; Khan et al., 2008; Kim et al., 2003; Kõlves et al., 2006; Kõlves et al., 2006; Manoranjitham et al., 2010; Page et al., 2014; Palacio et al., 2007; Préville et al., 2005; Renaud et al., 2008; Schneider et al., 2006; Shaffer et al., 1996; Shafii et al., 1988; Tong & Phillips, 2010; Vijayakumar & Rajkumar, 1999; Waern, 2003; Waern et al., 2002; Zhang et al., 2010; Zonda, 2006).

The PAMC was developed iteratively in five steps. Step 1 (draft of items): An initial list of items was generated that drew about equally from items adapted from Downs and Black's instrument and newly created items that were psychological autopsy specific. Step 2 (revision of items): The items were reviewed with co-authors with expertise in psychological research methodology revised until consensus was reached among the experts of a draft version of the instrument. Step 3 (training of raters): The first author trained two research associates who are new to suicide research and did not know any of the other study authors to score the PAMC. This training began with reading and discussion of the aforementioned psychological autopsy methodology reviews (Conner et al., 2011, 2012) and a review and discussion of each PAMC item. Next, over four rounds, the raters independently scored a series of case–control psychological autopsy research papers selected by the trainer, typically two per round, and came together with the trainer to discuss their ratings, resolve discrepancies, and generate a consensus final rating, with each round leading to minor modification of the instrument to improve reliability. The papers used in training were chosen to represent case–control psychological autopsy research internationally across a range of study quality and included eight reports contained in the meta-analysis. Step 4 (second revision): With the PAMC having been modified through the training process, the revised instrument was reviewed with the other authors, leading to additional minor changes and a final draft of the instrument. Step 5 (inter-rater reliability): Using the remaining 27 research reports contained in the meta-analysis (i.e., after excluding the eight reports used to train the raters), the trainer used an online program to order the studies at random before instructing the raters to independently rate these studies in order in batches of two

to four, to review the ratings in each batch at a consensus meeting, and to resolve any discrepancies and generate a consensus rating before going onto the next batch. The studies were reviewed in batches to prevent drift over the course of generating the ratings. During this process, the first paper the raters' reviewed generated additional questions that were resolved by the trainer, making it effectively a final training study. Therefore, this study was excluded from the inter-rater reliability analysis, and the raters independently scored the remaining 26 studies contained in the meta-analysis before arriving at consensus, per the procedure. These consensus ratings were made independently by the raters without input from the trainer or the other study authors. The 26 studies selected from the aforementioned meta-analysis (Conner et al., 2019) to generate the reliability analyses, along with the nine studies meta-analyzed that were used to train the raters, are described in Table 1.

One PAMC item that concerned the choice of controls showed low inter-rater reliability (percent agreement = 61.5%) and was dropped from the scale, yielding a final PAMC containing 16 items with scoring range from 0 to 18. The instrument with instructions for administration and scoring is provided in Appendix 1. PAMC items assess the Downs and Black's schema of quality of reporting (items 1–4), external validity (items 5–7), internal validity (items 8–15), along with an item on power (item 16). There are eight items that were adapted from Down and Black (items 1–5, 8–10) to provide coverage of fundamental features of case–control studies regardless of topic including reporting, internal validity, and external validity, along with eight items to address the unique needs of psychological autopsy research that warrant elaboration (items 6–7, 11–16). Two PAMC items assess the extent to which the cases (item 6) and controls (item 7) were representative of the study population. These items are credited in studies that achieve acquisition rates of 80%, deemed sufficient to mitigate concern about ascertainment bias, or when data are shown to suggest that missed subjects did not differ from those that were included and/or the source population. Item 11 ensures that determinations of suicide were determined or confirmed by researchers and/or were based on an official source (e.g., coroner report). Item 12 credits studies that interviewed more than one proxy respondent, assuming such interviews were performed systematically in both cases and controls and featured the main interview battery. Item 13 credits studies that systematically integrated information from records (e.g., primary care, school) or from interviews that served as a substitute for records (e.g., physicians, teachers). Item 14 credits studies that demonstrated the reliability of primary measures or that featured measures that are known to be reliable when used postmortem. Item 15 assesses the timeliness of postmortem interviews which optimally should be performed within 6 months of death to mitigate retrospective bias. Item 16 concerns statistical power, a key issue in psychological autopsy studies where accruing sufficient numbers of cases may be challenging.

Analysis

Our primary interest was inter-rater reliability, and therefore, percent agreement was computed (also referred to as observed agreement) and Kappa coefficients a) for the overall scale (16 items) and b) for subscales composed of the items adapted from Downs and Black (8 items) and the new psychological autopsy specific items (8 items). Kappa was interpreted based on the benchmarks (poor < 0.01, slight 0.01–0.20, fair 0.21–0.40, moderate 0.41–

0.60, substantial 0.61–0.80, almost perfect >0.80) provided by Landis and Koch (1977) and reprinted in Dunn (1989). Coefficient alpha was also calculated for the total scale and the subscales to provide a measure of internal consistency reliability. Secondary analysis examined agreement item-by-item within studies and across studies. Note that Kappa compares the observed agreement to that expected based on the base rate of category usage (Cohen, 1960). Thus, it is possible to have high agreement, but also high expected agreement due to frequent use of a particular category. In this case, overall agreement will be high, but Kappa low. As well, in instances of perfect observed and expected agreement, Kappa cannot even be computed. These instances occurred repeatedly at the item level and study level and, therefore, we report observed agreement for these variables. Finally, PAMC scores were generated based on summing all items after obtaining consensus agreement of the two independent raters on the scoring of any discrepant items, providing an overall measure of methodological quality.

RESULTS

Overall PAMC scores for the 35 research reports reviewed ranged from 8 to 17 (median = 12), with higher scores suggesting greater methodological quality. These PAMC results are shown in the fifth column of Table 1, categorized into lower scores (8–10), intermediate scores (11–13), and higher scores (14–17). Observed inter-rater reliability for the overall PAMC scale was high and that based more conservatively on Kappa was substantial (percent agreement 86.5%, expected agreement 53.4%, κ 0.71). For the subscale based on adapted Downs and Black's (1998) items, observed inter-rater reliability was high and that based on Kappa was moderate (percent agreement 80.9%, expected agreement 80.9%, κ 0.47). The high rate of expected agreement indicated that raters tended to use one particular item category frequently. For the subscale composed of new items specific to psychological autopsy research, observed inter-rater reliability was high and that based on Kappa was substantial (percent agreement 83.2%, expected agreement 46.4%, κ .69). Inter-rater agreement for the overall PAMC scale for each research report is presented in the final column in Table 1. Across studies, agreement between independent raters was acceptable to high, ranging from 70.6% to 100% (median = 82.4%).

Analyses of individual PAMC items are shown in Table 2. Agreement on items 3 and 4 was perfect (100%) across all 26 studies, and items 2 and 11 had only one instance of disagreement. In these instances of essentially perfect agreement between raters, expected agreement is also perfect or nearly perfect, because it is based on the marginal frequencies, and only a single category is used. In essence, these items showed no variation across studies and were always (or nearly always) rated the same by both raters. Therefore, Kappa cannot be calculated. There were also two items showing high observed agreement and expected agreement (items 5, 8). Kappa is thus not meaningful. Observed agreement between raters for the other items ranged from moderate (69.2%, item 7) to high (92.3%, item 15), and Kappa for the such items ranged from fair (κ 0.28, item 13) to almost perfect (κ 0.88, item 15).

Coefficient alpha provided an assessment of internal consistency of the PAMC scales. Alpha was low for the total scale (α = 0.56) as well as for the items derived from Downs

and Black (α = 0.31), and the newly created items (α = 0.48). These results indicate the aspects of study quality tapped by these items are modestly inter-related. To explore item heterogeneity, principal components analysis (conceptually treating items as formative, rather than reflective indicators) was used (Edwards & Bagozzi, 2000). This exploratory analysis did not suggest subcomponents with any obvious interpretation.

DISCUSSION

The PAMC is a hybrid instrument created from items that were adapted from an existing methodological quality rating checklist (Downs & Black, 1998) and new items that were designed to assess the special requirements of case–control psychological autopsy research. Accordingly, the PAMC provides a balance between assessment of standard observational research methods and those specific to psychological autopsy research, with an equal number of items assessing each domain. Results show that PAMC items have substantial inter-rater reliability for the overall scale and the subscales consisting of items from an established checklist as well as the new items. Inter-rater reliability was also acceptable to high across studies conducted internationally.

The instrument showed low internal consistency and may be tapping different components, suggesting the need for caution in interpreting the results when the items are summed. In other words, although summed PAMC scores provide a general benchmark for methodological quality, and we presume that studies scoring in the “higher” category have methodological advantages over those scoring in the “lower” category (see Table 1), the low internal consistency results suggest that it would be a mistake to make too much of more minor differences between scores. Indeed, the primary strength of the PAMC may be in the newly created items themselves which provide operational definitions for a range of methodological features unique to case–control psychological autopsy methodology that may inform the planning of studies and the identification of specific limitations of published reports. Accordingly, we have labeled the PAMC as a “checklist” and invite researchers to use it as such in designing their case–control psychological research studies. As a novel checklist, the PAMC addresses a gap in the field that contains several methodological reviews (Beskow et al., 1990; Brent, 1989; Clark & Horton-Deutsch, 1992; Conner et al., 2011, 2012; Ebert, 1987; Hawton et al., 1998; Isometsä, 2001; Pouliot & De Leo, 2006; Schneidman, 1981; Shaffer et al., 1972) but that has been lacking detailed, structured guidance.

There are other limitations of the PAMC. As an assessment of methodological quality, it does not examine other critical elements of studies such as their timeliness. The PAMC is designed to be applicable across the range of case–control psychological autopsy studies conducted internationally. Therefore, atypical but exciting features such as simultaneous recruitment and study of suicide attempt patients for comparison with suicides (Beautrais, 2001) or the gathering of uniform data in different countries to facilitate cross-national comparisons (Kõlves, Sisask, et al., 2006) are not reflected in PAMC scores. Designed for quantitative research, the PAMC does not assess the methodological rigor of psychological autopsy studies that use qualitative designs (Hjelmeland & Knizek, 2016) or the qualitative elements of studies that feature mixed quantitative and qualitative

designs (Zalsman et al., 2016). The instrument does not consider data from electronic communications (e.g., text messages, social media posts) prior to death because there is little precedence how this might be done in the context of case–control psychological autopsy research. Nonetheless, it seems essential that future studies build on recent efforts to capture electronic communications prior to suicide (Bryan et al., 2018) and develop strategies to gather comparable data from controls. Although the PAMC addresses several methodological features related to reporting, external validity, and internal validity, the coverage it provides in these areas is not exhaustive.

Case–control studies are considered a lower standard of evidence compared with research using prospective designs (Murad et al., 2016). That being said, case–control psychological autopsy studies seem irreplaceable for the purpose of examining proximal risk factors for suicide deaths, with the low incidence rate of suicide ruling out the routine use of longitudinal studies for the purpose of examining the final weeks of life, the focus of psychological autopsy research. Although the study of high-risk clinical populations, for example psychiatric patients, may provide samples with enriched risk to allow for the prospective study of proximal risk factors, this is not a replacement for psychological autopsy research of general population samples because most individuals who die by suicide are not in behavioral health care (Niederkrotenthaler et al., 2014; Stone et al., 2018). As well, even if prospective studies are able to examine proximal risk, the use of postmortem psychological autopsy methodology can fill in inevitable gaps when suicide deaths occur. In settings where there has been little study of suicide or in which the landscape of suicide has changed, case–control psychological autopsy studies can play a critical role in prioritizing targets for prevention, and we invite researchers to use the PAMC to inform study design and to assess specific features of methodological quality.

APPENDIX 1

The Psychological Autopsy Methodology Checklist (PAMC)

Research report: _____

Total Score: _____ Rater: _____

Instructions:

Rate all PAMC items and sum the total (range 0–18). Consider all information provided by the study authors in scoring difficult to rate items. Refer to prior published psychological autopsy reports of a dataset when needed. For items that remain unresolvable or unclear, rate the item down.

Reporting—1. Is the hypothesis/aim/objective of the study clearly described?

Yes—each hypothesis/aim/objective	2
Partial—some but not all hypotheses/aims/objectives	1

No	0

Notes. Adapted from Downs and Black (1998) item #1. Score 1 if the general focus or aim of the study is clear but specific questions or hypotheses are not provided.

2. Are the characteristics of the subjects included in the study clearly described?

Yes	1
No	0

Notes. Adapted from Downs and Black (1998) item #3.

3. Are the main findings of the study clearly described?

Yes	1
No	0

Notes. Downs and Black (1998) item #6.

4. Have actual probability values (e.g., 0.035 rather than <0.05), except where the probability value is less than 0.001, or confidence intervals been reported for the main outcomes?

Yes	1
No	0

Adapted from Downs and Black (1998) item #10.

External Validity—5. Are the cases representative of the population from which they were recruited?

Yes	1
No	0
Unable to determine	0

Notes. Adapted from Downs and Black (1998) item #11. Score the item based on the representativeness of cases only (comparability of controls is addressed in items 8 and 9). The labor intensity of psychological autopsy research may require sampling a small area (e.g., select city, county, or village) or combining data from different areas. Do not rate down unless the sample obtained is non-representative of the area from which it comes or it is unreasonable to combine the data from different areas if more than one area was sampled.

6. Is non-response bias adequately addressed in cases?

Yes	1
No	0
Unable to determine	0

Notes. Rate 1 if there is a high case acquisition rate (≥80%) or analyses of non-response are provided that do not suggest non-response bias (e.g., respondents of cases do not differ from non-respondents, respondents of cases do not differ from the source population). Rate 0 if the response rate of cases is lower than 80% and there is no information to assess non-response bias or such information indicates significant bias. Rate 0 if the response rate in cases is not provided.

7. Is non-response bias adequately addressed in controls?

Yes	1
No	0
Unable to determine	0

Notes. Rate 1 if there is a high control acquisition rate (≥80%) or analyses of non-response are provided that do not suggest non-response bias (e.g., respondents of controls do not differ from non-respondents, respondents of controls do not differ from the source population). Rate 0 if the response rate of controls is lower than 80% and there is no information to assess non-response bias or such information indicates significant bias. Rate 0 if the response rate in controls is not provided.

Internal Validity—8. Are the cases and controls recruited from the same population?

Yes	1
No	0
Unable to determine	0

Notes. Adapted from Downs and Black (1998) item #21.

9. Are the cases and controls recruited over the same period of time?

Yes	1
No	0
Unable to determine	0

Notes. Adapted from Downs and Black (1998) item #22. Score 1 if there is a reasonable lag in recruitment of controls that is attributable to a matching procedure. Score 1 if it may be

assumed that the cases and controls were recruited over the same general time period based on all information presented.

10. Is there adequate adjustment for confounding in analyses from which the central findings were drawn?

Yes	1
No	0
Unable to determine	0

Notes. Adapted from Downs and Black (1998) item #25. Adequate adjustment does not require exhaustive adjustment.

11. Determinations of suicide

Reported or confirmed by researchers	1
Determinations not provided	0
Unable to determine	0

Notes. Rate 1 if cause of death determinations of suicide is determined or confirmed by researchers either in all cases or in ambiguous cases (e.g., open verdicts, undetermined deaths). Rate 1 if the source of determination of suicide (e.g., coroner's report) or the method of determination (e.g., adjudication) is reported. Rate 0 if information is lacking about the source of determinations of suicide.

12. Are multiple interviews used and systematically integrated?

Yes	1
No	0
Unable to determine	0

Notes. Rate 1 if more than one main interview is conducted routinely with both cases and controls and the information from the interviews is combined in a planful manner (e.g., "rate up" for reports of symptoms, best estimate procedure, diagnostic consensus process, prioritizing one interview and using a second interview to assess reliability, seeking additional information to resolve discrepancies between interviews). Rate 0 if a single interview is used. Rate 0 if multiple interviews are sometimes used but are the exception. Rate 0 if multiple interviews are used with one group (e.g., cases) but not the other. Rate 0 if multiple interviews are used but the process for integrating the information is not mentioned or non-systematic. Rate 0 if additional interviews are limited to those that do not feature the main interview battery (e.g., discussions to obtain supplemental information from a provider).

13. Are records used and systematically integrated?

Yes or not applicable	1
No	0
Unable to determine	0

Notes. Rate 1 if one or more records (e.g., primary care, legal, school) are routinely obtained from cases and controls and integrated with other information in a planful manner. Rate 1 if interviews are substituted for records in a planful manner to obtain such information (e.g., interviews with primary care physicians). Rate 1, not applicable, if the setting rules out obtaining records (e.g., rural setting in a developing country where such records are widely unavailable). Rate 0 if records are used but they are poorly described or referred to broadly. Rate 0 if no information is provided how such records are integrated with other data or if such integration appears non-systematic. Rate 0 if records are obtained from cases only. Rate 0 if the only source of records is the death investigation.

14. Are key measure(s) reliable?

Yes	1
No	0
Unable to determine	0

Notes. Rate 1 if reliability of one or more key variable(s) is provided and is satisfactory. For example, a report on risk factors for suicide that includes a focus on mental disorders and reports satisfactory reliability of the measurement of mental disorders is rated 1. Rate 1 if such data are presented even if a secondary measure(s) has unsatisfactory reliability. Objective measures that are known to be reliable in postmortem research (e.g., blood alcohol concentration) may be assumed to be reliable. Rate 0 if a report focuses on a given variable (e.g., a report on risk associated with drug use) but does not report data on its reliability even if data on reliability of other measures is provided. Rate 0 if reliability data are absent. Rate 0 if reliability data of an instrument is referred to but information on its reliability in the sample is not provided.

15. Are there long delays in time to assessments from the date of suicide?

Time to assessment <6 months	2
Time to assessment 6 months to 1 year	1
Time to assessment >1 year	0
Unable to determine	0

Notes. Base the rating on the mean (or median) time to proxy respondent interview from the date of suicide in cases. The date of the control interview for control subjects is typically used as the reference date in psychological autopsy research and this is acceptable.

Power—16. Did the study provide an *a priori* sample size calculation?

Yes	1
No	0
Unable to determine	0

REFERENCES

Almasi K, Belso N, Kapur N, Webb R, Cooper J, Hadley S, Kerfoot M, Dunn G, Sotonyi P, Rihmer Z, & Appleby L (2009). Risk factors for suicide in Hungary: a case-control study. BMC Psychiatry, 9, 45. 10.1186/1471-244x-9-45. [PubMed: 19638202]

Appleby L, Cooper J, Amos T, & Faragher B (1999). Psychological autopsy study of suicides by people aged under 35. British Journal of Psychiatry, 175, 168–174. 10.1192/bjp.175.2.168.

Arsenault-Lapierre G, Kim C, & Turecki G (2004). Psychiatric diagnoses in 3275 suicides: A meta-analysis. BMC Psychiatry, 4, 37. 10.1186/1471-244x-4-37. [PubMed: 15527502]

Beautrais AL (2001). Suicides and serious suicide attempts: two populations or one? Psychological Medicine, 31(5), 837–845. 10.1017/s0033291701003889. [PubMed: 11459381]

Beskow J, Runeson B, & Asgård U (1990). Psychological autopsies: Methods and ethics. Suicide and Life-Threatening Behavior, 20(4), 307–323. [PubMed: 2087767]

Brent DA (1989). The psychological autopsy: methodological considerations for the study of adolescent suicide. Suicide and Life-Threatening Behavior, 19(1), 43–57. 10.1111/j.1943-278x.1989.tb00365.x. [PubMed: 2652387]

Brent DA, Baugher M, Bridge J, Chen T, & Chiappetta L (1999). Age-and sex-related risk factors for adolescent suicide. Journal of the American Academy of Child & Adolescent Psychiatry, 38(12), 1497–1505. 10.1097/00004583-199912000-00010. [PubMed: 10596249]

Brent DA, Perper JA, Moritz G, Allman C, Friend A, Roth C, Schweers J, Balach L, & Baugher M (1993). Psychiatric risk factors for adolescent suicide: A case-control study. Journal of the American Academy of Child & Adolescent Psychiatry, 32(3), 521–529. 10.1097/00004583-199305000-00006. [PubMed: 8496115]

Bryan CJ, Butner JE, Sinclair S, Bryan ABO, Hesse CM, & Rose AE (2018). Predictors of emerging suicide death among military personnel on social media networks. Suicide and Life-Threatening Behaviors, 48(4), 413–430. 10.1111/sltb.12370.

Cavanagh JT, Carson AJ, Sharpe M, & Lawrie SM (2003). Psychological autopsy studies of suicide: a systematic review. Psychological Medicine, 33(3), 395–405. 10.1017/s0033291702006943. [PubMed: 12701661]

Chan SS, Chiu HF, Chen EY, Chan WS, Wong PW, Chan CL, Law YW, & Yip PS (2009). Population-attributable risk of suicide conferred by axis I psychiatric diagnoses in a Hong Kong Chinese population. Psychiatric Services, 60(8), 1135–1138. 10.1176/ps.2009.60.8.1135. [PubMed: 19648205]

Chen EY, Chan WS, Wong PW, Chan SS, Chan CL, Law YW, Philips SL, Chan KK, Cheng JW, Liu KY, & Yip PS (2006). Suicide in Hong Kong: A case-control psychological autopsy study. Psychological Medicine, 36(6), 815–825. 10.1017/s0033291706007240. [PubMed: 16704748]

Cheng AT (1995). Mental illness and suicide. A case-control study in east Taiwan. Archives of General Psychiatry, 52(7), 594–603. 10.1001/archpsyc.1995.03950190076011. [PubMed: 7598636]

Cheng AT, Chen TH, Chen CC, & Jenkins R (2000). Psychosocial and psychiatric risk factors for suicide. Case-control psychological autopsy study. British Journal of Psychiatry, 177, 360–365. 10.1192/bjp.177.4.360.

Chiu HF, Yip PS, Chi I, Chan S, Tsoh J, Kwan CW, Li SF, Conwell Y, & Caine E (2004). Elderly suicide in Hong Kong–a case-controlled psychological autopsy study. Acta Psychiatrica Scandinavica, 109(4), 299–305. 10.1046/j.1600-0447.2003.00263.x. [PubMed: 15008804]

Clark DC, & Horton-Deutsch S (1992). Assessment in absentia: The value of the psychological autopsy method for studying antecedents of suicide and predicting future suicides. In Maris RW, Berman AL, Maltsberger JT, & Yufit RI (Eds.), Assessment and prediction of suicide (pp. 144–182)..

Cohen J (1960). A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 20, 37–46. 10.1177/001316446002000104.

Conner KR, Beautrais AL, Brent DA, Conwell Y, Phillips MR, & Schneider B (2011). The next generation of psychological autopsy studies. Part I. Interview content. Suicide and Life-Threatening Behavior, 41(6), 594–613. 10.1111/j.1943-278X.2011.00057.x. [PubMed: 22050639]

Conner KR, Beautrais AL, Brent DA, Conwell Y, Phillips MR, & Schneider B (2012). The next generation of psychological autopsy stuides. Part 2. Interview procedures. Suicide and Life-Threatening Behavior, 42(1), 86–103. 10.1111/j.1943-278X.2011.00073.x. [PubMed: 22288728]

Conner KR, Beautrais AL, & Conwell Y (2003). Moderators of the relationship between alcohol dependence and suicide and medically serious suicide attempts: Analyses of Canterbury Suicide Project data. Alcoholism Clinical and Experimental Research, 27(7), 1156–1161. 10.1097/01.Alc.0000075820.65197.Fd.

Conner KR, Bridge JA, Davidson DJ, Pilcher C, & Brent DA (2019). Metaanalysis of mood and substance use disorders in proximal risk for suicide deaths. Suicide and Life-Threatening Behavior, 49(1), 278–292. 10.1111/sltb.12422. [PubMed: 29193261]

Conner KR, Bridge JA, Davidson DJ, Plicher C, & Brent DA (2017). Meta-analysis of mood and substance use disorders in proximal risk for suicide deaths. Suicide and Life-Threatening Behavior, 49(1), 278–292. 10.1111/sltb.12422. [PubMed: 29193261]

Conwell Y, Duberstein PR, Hirsch JK, Conner KR, Eberly S, & Caine ED (2010). Health status and suicide in the second half of life. International Journal of Geriatric Psychiatry, 25(4), 371–379. 10.1002/gps.2348. [PubMed: 19582758]

De Leo D, Draper BM, Snowdon J, & Kõlves K (2013a). Contacts with health professionals before suicide: missed opportunities for prevention? Comprehensive Psychiatry, 54(7), 1117–1123. 10.1016/j.comppsych.2013.05.007. [PubMed: 23768696]

De Leo D, Draper BM, Snowdon J, & Kõlves K (2013b). Suicides in older adults: a case-control psychological autopsy study in Australia. Journal of Psychiatric Research, 47(7), 980–988. 10.1016/j.jpsychires.2013.02.009. [PubMed: 23522934]

Downs SH, & Black N (1998). The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. Journal of Epidemiology and Community Health, 52(6), 377–384. 10.1136/jech.52.6.377. [PubMed: 9764259]

Dunn G (1989). Design and analysis of reliability studies: The statistical evaluation of measurement errors. Oxford University Press.

Ebert BW (1987). Guide to conducting a psychological autopsy. Professional Psychology: Research and Practice, 18(1), 52–56. 10.1037/0735-7028.18.1.52.

Edwards JR, & Bagozzi RP (2000). On the nature and direction of relationships between constructs and measures. Psychological Methods, 5(2), 155–174. 10.1037/1082-989x.5.2.155. [PubMed: 10937327]

Foster T, Gillespie K, McClelland R, & Patterson C (1999). Risk factors for suicide independent of DSM-III-R Axis I disorder. Case-control psychological autopsy study in Northern Ireland. British Journal of Psychiatry, 175, 175–179. 10.1192/bjp.175.2.175.

Freuchen A, Kjelsberg E, Lundervold AJ, & Grøholt B (2012). Differences between children and adolescents who commit suicide and their peers: A psychological autopsy of suicide victims compared to accident victims and a community sample. Child and Adolescent Psychiatry and Mental Health, 6, 1. 10.1186/1753-2000-6-1. [PubMed: 22216948]

Harwood D, Hawton K, Hope T, & Jacoby R (2001). Psychiatric disorder and personality factors associated with suicide in older people: a descriptive and case-control study. International Journal of Geriatric Psychiatry, 16(2), 155–165. 10.1002/1099-1166(200102)16:2<155:aid-gps289>3.0.co;2-0. [PubMed: 11241720]

Hawton K, Appleby L, Platt S, Foster T, Cooper J, Malmberg A, & Simkin S (1998). The psychological autopsy approach to studying suicide: A review of methodological issues. Journal of Affective Disorders, 50(2–3), 269–276. 10.1016/s0165-0327(98)00033-0 [PubMed: 9858086]

Hjelmeland H, & Knizek BL (2016). Qualitative evidence in suicide: Findings from qualitative psychological autopsy studies. In Olson K, Young RA, & Schultz IZ (Eds.), Handbook of qualitative health research for evidence-based practice (pp. 355–371). Springer.

Isometsä ET (2001). Psychological autopsy studies–A review. European Psychiatry, 16(7), 379–385. 10.1016/s0924-9338(01)00594-6. [PubMed: 11728849]

Khan MM, Mahmud S, Karim MS, Zaman M, & Prince M (2008). Case-control study of suicide in Karachi, Pakistan. British Journal of Psychiatry, 193(5), 402–405. 10.1192/bjp.bp.107.042069.

Kim C, Lesage A, Seguin M, Chawky N, Vanier C, Lipp O, & Turecki G (2003). Patterns of co-morbidity in male suicide completers. Psychological Medicine, 33(7), 1299–1309. 10.1017/s0033291703008146. [PubMed: 14580083]

Kõlves K, Sisask M, Anion L, Samm A, & Värnik A (2006). Factors predicting suicide among Russians in Estonia in comparison with Estonians: case-control study. Croatian Medical Journal, 47(6), 869–877. [PubMed: 17171808]

Kõlves K, Tucholski E, Värnik A, & De Leo D (2021). Psychological autopsy: A tool in suicide research. In Kõlves K, Sisask M, Värnik P, Värnik A, & De Leo D (Eds.), Advancing suicide research (pp. 161–174). Hogrefe.

Kõlves K, Värnik A, Tooding LM, & Wasserman D (2006). The role of alcohol in suicide: A case-control psychological autopsy study. Psychological Medicine, 36(7), 923–930. 10.1017/s0033291706007707. [PubMed: 16650347]

Landis JR, & Koch GG (1977). The measurement of observer agreement for categorical data. Biometrics, 33(1), 159–174. [PubMed: 843571]

Litman RE, Curphey TJ, Shneidman ES, Farberow NL, & Tabachnick ND (1963). Investigations of equivocal suicides. Journal of the American Medical Association, 184, 924–929.

Manoranjitham SD, Rajkumar AP, Thangadurai P, Prasad J, Jayakaran R, & Jacob KS (2010). Risk factors for suicide in rural south India. British Journal of Psychiatry, 196(1), 26–30. 10.1192/bjp.bp.108.063347.

Murad MH, Asi N, Alsawas M, & Alahdab F (2016). New evidence pyramid. Evidence-based Medicine, 21(4), 125–127. 10.1136/ebmed-2016-110401. [PubMed: 27339128]

Niederkrotenthaler T, Logan JE, Karch DL, & Crosby A (2014). Characteristics of U.S. suicide decedents in 2005–2010 who had received mental health treatment. Psychiatric Services, 65(3), 387–390. 10.1176/appi.ps.201300124. [PubMed: 24584526]

Page A, Morrell S, Hobbs C, Carter G, Dudley M, Duflou J, & Taylor R (2014). Suicide in young adults: psychiatric and socio-economic factors from a case-control study. BMC Psychiatry, 14, 68. 10.1186/1471-244x-14-68. [PubMed: 24597482]

Palacio C, García J, Diago J, Zapata C, Lopez G, Ortiz J, & Lopez M (2007). Identification of suicide risk factors in Medellín, Colombia: a case-control study of psychological autopsy in a developing country. Archives of Suicide Research, 11(3), 297–308. 10.1080/13811110600894223. [PubMed: 17558615]

Pouliot L, & De Leo D (2006). Critical issues in psychological autopsy studies. Suicide and Life Threatening Behavior, 36(5), 491–510. 10.1521/suli.2006.36.5.491. [PubMed: 17087629]

Préville M, Hébert R, Boyer R, Bravo G, & Seguin M (2005). Physical health and mental disorder in elderly suicide: a case-control study. Aging & Mental Health, 9(6), 576–584. 10.1080/13607860500192973. [PubMed: 16214706]

Renaud J, Berlim MT, McGirr A, Tousignant M, & Turecki G (2008). Current psychiatric morbidity, aggression/impulsivity, and personality dimensions in child and adolescent suicide: A case-control study. Journal of Affective Disorders, 105(1–3), 221–228. 10.1016/j.jad.2007.05.013. [PubMed: 17568682]

Robins E, Murphy GE, Wilkinson RH Jr, Gassner S, & Kayes J (1959). Some clinical considerations in the prevention of suicide based on a study of 134 successful suicides. American Journal of Public Health, 49(7), 888–899. 10.2105/ajph.49.7.888. [PubMed: 13661481]

Rothman KJ, & Greenland S (1998). Case-control studies. In Rothman KJ, & Greenland S (Eds.), Modern epidemiology (pp. 33–114). Lippincott-Raven.

Schneider B, Wetterling T, Sargk D, Schneider F, Schnabel A, Maurer K, & Fritze J (2006). Axis I disorders and personality disorders as risk factors for suicide. European Archives of Psychiatry and Clinical Neuroscience, 256(1), 17–27. 10.1007/s00406-005-0593-7. [PubMed: 16133739]

Schneidman ES (1981). The psychological autopsy. Suicide and Life-Threatening Behavior, 11(4), 325–340. 10.1111/j.1943-278X.1981.tb01009.x.

Shaffer D, Gould MS, Fisher P, Trautman P, Moreau D, Kleinman M, & Flory M (1996). Psychiatric diagnosis in child and adolescent suicide. Archives of General Psychiatry, 53(4), 339–348. 10.1001/archpsyc.1996.01830040075012. [PubMed: 8634012]

Shaffer JW, Perlin S, Schmidt CW Jr, & Himelfarb M (1972). Assessment in absentia: new directions in the psychological autopsy. Johns Hopkins Medical Journal, 130(5), 308–316.

Shafii M, Steltz-Lenarsky J, Derrick AM, Beckner C, & Whittinghill JR (1988). Comorbidity of mental disorders in the post-mortem diagnosis of completed suicide in children and adolescents. Journal of Affective Disorders, 15(3), 227–233. 10.1016/0165-0327(88)90020-1. [PubMed: 2975295]

Stone DM, Simon TR, Fowler KA, Kegler SR, Yuan K, Holland KM, Ivey-Stephenson AZ, & Crosby AE (2018). Vital signs: Trends in state suicide rates - United States, 1999–2016 and Circumstances Contributing to Suicide - 27 States, 2015. Morbidity and Mortality Weekly Report, 67(22), 617–624. 10.15585/mmwr.mm6722a1. [PubMed: 29879094]

Tong Y, & Phillips MR (2010). Cohort-specific risk of suicide for different mental disorders in China. British Journal of Psychiatry, 196(6), 467–473. 10.1192/bjp.bp.109.070409.

Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, Pocock SJ, Poole C, Schlesselman JJ, & Egger M (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and elaboration. PLoS Med, 4(10), e297. 10.1371/journal.pmed.0040297. [PubMed: 17941715]

Vijayakumar L, & Rajkumar S (1999). Are risk factors for suicide universal? A case-control study in India. Acta Psychiatrica Scandinavica, 99(6), 407–411. 10.1111/j.1600-0447.1999.tb00985.x. [PubMed: 10408261]

Waern M (2003). Alcohol dependence and misuse in elderly suicides. Alcohol and Alcoholism, 38(3), 249–254. 10.1093/alcalc/agg060. [PubMed: 12711660]

Waern M, Runeson BS, Allebeck P, Beskow J, Rubenowitz E, Skoog I, & Wilhelmsson K (2002). Mental disorder in elderly suicides: A case-control study. American Journal of Psychiatry, 159(3), 450–455. 10.1176/appi.ajp.159.3.450.

Yoshimasu K, Kiyohara C, & Miyashita K (2008). Suicidal risk factors and completed suicide: meta-analyses based on psychological autopsy studies. Environmental Health and Preventive Medicine, 13(5), 243–256. 10.1007/s12199-008-0037-x. [PubMed: 19568911]

Zalsman G, Siman Tov Y, Tzuriel D, Shoval G, Barzilay R, Tiech Fire N, Sherf M, & John Mann J (2016). Psychological autopsy of seventy high school suicides: Combined qualitative/quantitative approach. European Psychiatry, 38, 8–14. 10.1016/j.eurpsy.2016.05.005. [PubMed: 27611329]

Zhang J, Xiao S, & Zhou L (2010). Mental disorders and suicide among young rural Chinese: A case-control psychological autopsy study. American Journal of Psychiatry, 167(7), 773–781. 10.1176/appi.ajp.2010.09101476.

Zonda T (2006). One-hundred cases of suicide in Budapest: a case-controlled psychological autopsy study. Crisis, 27(3), 125–129. 10.1027/0227-5910.27.3.125. [PubMed: 17091822]

TABLE 1

Reports used in training and inter-rater reliability analysis with consensus PAMC results

Report	*N*	Country	Age	PAMC	Agreement
Almasi et al. (2009)	194	Hungary	Median = 43	Higher	Training
Appleby et al. (1999)	84	England	Under 35	Intermed.	Training
Beautrais (2001) [1]	202	New Zealand	14-plus	Intermed.	100%
Brent et al. (1999) [2]	140	United States	13–19	Intermed.	88.2%
Brent et al. (1993) [2]	67	United States	Under 20	Higher	Training
Chan et al. (2009) [3]	150	China (Hong Kong)	15–59	Intermed.	76.5%
Chen et al. (2006) [3]	150	China (Hong Kong)	15–59	Intermed.	82.4%
Cheng (1995) [4]	117	Taiwan	15-plus	Higher	82.4%
Cheng et al. (2000) [4]	117	Taiwan	15-plus	Higher	Training
Chiu et al. (2004)	70	China (Hong Kong)	60-plus	Lower	82.4%
Conner et al. (2003) [1]	193	New Zealand	18-plus	Intermed.	88.2%
Conwell et al. (2010)	86	United States	50-plus	Intermed.	82.4%
De Leo et al. (2013a)[5]	261	Australia	35-plus	Intermed.	94.1%
De Leo et al. (2013b)[5]	261	Australia	35-plus	Lower	88.2%
Foster et al. (1999)	117	Northern Ireland	14-plus	Intermed.	64.7%
Freuchen et al. (2012)	41	Norway	Under 16	Lower	76.5%
Harwood et al. (2001)	54	England	60-plus	Intermed.	82.4%
Khan et al. (2008)	100	Pakistan	Not given	Intermed.	82.4%
Kim et al. (2003)	115	Canada	18–65	Lower	Training
Kõlves, Sisask, et al., 2006; Kõlves, Värnik, et al., 2006[6]	427	Russia, Estonia	Mean=48	Intermed.	82.4%
Kõlves, Sisask, et al., 2006; Kõlves, Värnik, et al., 2006[6]	419	Russia, Estonia	Mean=48	Lower	94.1%
Manoranjitham et al. (2010)	100	India	Mean=42	Higher	88.2%
Page et al. (2014)	84	Australia	18–34	Lower	82.4%
Palacio et al. (2007)	108	Colombia	Median=29	Intermed.	Training
Préville et al. (2005)	95	Canada	60-plus	Intermed.	100%
Renaud et al. (2008)	55	Canada	11–18	Intermed.	94.1%
Schneider et al. (2006)	163	Germany	Mean=50	Higher	76.5%
Shaffer et al. (1996)	120	United States	Under 20	Higher	100%

Report	*N*	Country	Age	PAMC	Agreement
Shafii et al. (1988)	21	United States	11–19	Intermed.	88.2%
Tong & Phillips. (2010)	895	China	11-plus	Higher	70.6%
Vijayakumar & Rajkumar (1999)	100	India	15-plus	Higher	Training
Waern et al. (2002) [7]	85	Sweden	65-plus	Intermed.	Training
Waern (2003) [7]	88	Sweden	65-plus	Lower	88.2%
Zhang et al. (2010)	392	China	15–34	Higher	76.5%
Zonda (2006)	100	Hungary	Mean=52	Lower	Training

Notes: Reports (n = 35) are from a meta-analysis of case–control psychological autopsy studies reporting results for suicide risk associated with mood or substance use disorders (Conner et al., 2017). Studies with common numerators have overlapping samples. Sample sizes (N) and age are based on cases. Reports were used in the training of the raters (n = 9, training) or in the inter-rater reliability analyses (n = 26, reliability). The reports used in training were rescored and brought to consensus following completion of the reliability analysis to promote accuracy. The fifth column shows PAMC results (range 8–17, median =12) based on consensus agreement of two raters and categorized into Lower scores (8–10), Intermediate scores (11–13), and Higher scores (14–17). The final column shows the results of the reliability analysis (percent agreement) for the PAMC between the two raters, made independently (prior to consensus) for non-training reports.

TABLE 2

Description of PAMC items and observed inter-rater agreement

Item	Description	Domain	Source	Agreement (%)
1	Hypotheses clear	Reporting	D&B	73.1
2	Characteristics of subjects clear	Reporting	D&B	100
3	Main findings clear	Reporting	D&B	100
4	Exact p-values or 95% CI's reported	Reporting	D&B	100
5	Cases representative of population	External validity	D&B	88.5
6	Non-response bias mitigated in cases	External validity	New	80.8
7	Non-response bias mitigated in controls	External validity	New	69.2
8	Cases and controls from same population	Internal validity	D&B	88.5
9	Cases and controls recruited same time period	Internal validity	D&B	80.8
10	Adequate adjustment for confounding	Internal validity	D&B	92.3
11	Determinations of suicide reported	Internal validity	New	100
12	Multiple interviews used and integrated	Internal validity	New	96.2
13	Records used and integrated	Internal validity	New	69.2
14	Reliability of key measures acceptable	Internal validity	New	76.9
15	Delays to assessments after suicide minimized	Internal validity	New	92.3
16	*A priori* power calculation provided	Power	New	84.6

Notes: New = item specific to psychological autopsy methodology created for the PAMC. CI's = confidence intervals. D&B = item adapted from a general checklist of methodological quality: Downs and Black (1998).