



Long-term outcomes of childhood sexual abuse: An umbrella review

Helen P. Hailes

Balliol College

**A thesis submitted in fulfilment of the requirements for the degree:
Master of Science by Research in Psychiatry**

Trinity Term, 2017

Word Count: 20,830

Acknowledgments

I would like to thank everyone who has contributed to and supported my work on this thesis. In particular, I am grateful for the invaluable guidance and feedback from my supervisor, Professor Seena Fazel. I would like to thank Dr. Rongqin Yu as well, for her frequent advice and assistance with data collection. I also owe my gratitude to other members of the Oxford Forensic Psychiatry research team, Achim Wolf, Taanvi Ramesh, and Monique Ewen for answering many questions, helping me prepare for assessment, and organizing logistics, and to Daan Wesselink for proof reading and general support.

Finally, I would like to express a personal thank you to my family in the US and UK and to my friends in Oxford for their support throughout the thesis writing process and for making this year such an enjoyable experience.

I would like to disclose no conflict of interest or external funding sources for this thesis.

Table of Contents

Acknowledgments.....	2
Abstract.....	5
Chapter 1: Introduction	6
1.1 Why Study Childhood Sexual Abuse?.....	6
1.2 Childhood Sexual Abuse and Later Outcomes	7
1.3 Meta-Analyses and Umbrella Reviews.....	9
1.4 Previous Umbrella Reviews on Outcomes of Childhood Sexual Abuse.....	11
1.5 Methodological Challenges	13
1.6 Aims of This Study.....	16
Chapter 2: Methods	16
2.1 Data Sources	17
2.2 Study Selection and Inclusion Criteria	19
2.3 Data Extraction	23
2.4 Data Analysis	27
2.5 Population Attributable Risk.....	28
2.6 Subgroup Analysis by Gender.....	30
2.7 Analysis of Evidence Quality	31
Chapter 3: Results	35
3.1 Study Characteristics	35
3.2 Meta-Analysis	38
3.3 Population Attributable Risk.....	42
3.4 Subgroup Analysis by Gender.....	44
3.5 AMSTAR.....	45
3.6 Heterogeneity	48
3.7 Excess Statistical Significance	48
3.8 Small Study Effects.....	50
3.9 Prediction Intervals.....	51
3.10 Overall Quality Analysis.....	53
Chapter 4: Discussion	56
4.1 Effect Size Findings	56
4.2 Quality Analysis Findings.....	57
4.3 Potential Outcome Mechanisms.....	59
4.4 Analysis of Gender Differences	63

4.5 Mediating, Moderating, and Confounding Factors	66
4.6 Gaps in the Literature.....	68
4.7 Umbrella Review Limitations	70
4.8 Research Implications.....	71
4.9 Clinical Implications	72
4.10 Conclusion.....	75
References.....	77
Appendix 1: AMSTAR Checklist.....	84
Appendix 2: PRISMA Checklist	87

Abstract

TITLE: Long-term outcomes of childhood sexual abuse: An umbrella review

NAME: Helen P. Hailes

COLLEGE: Balliol College

DEGREE: Master of Science by Research in Psychiatry

TERM SUBMITTED: Trinity Term, 2017

INTRODUCTION: With a global prevalence of up to 10% in boys and 20% in girls, childhood sexual abuse is a serious public health concern. Primary studies and meta-analyses have investigated the long-term implications of childhood sexual abuse on a range of health and psychosocial outcomes. The aim of this umbrella review is to systematically review previous meta-analyses on the long-term outcomes of childhood sexual abuse, comparing findings across outcomes, assessing quality, and identifying gaps in the literature.

METHODS: Four databases (PsycINFO, PubMed, CINAHL, and Global Health) were searched, and 28 childhood sexual abuse outcomes were identified across 15 studies. Effect size data were extracted by two independent reviewers, and statistics were compared and pooled across outcomes. Population attributable fractions were calculated, and a subgroup analysis by gender was conducted. Quality analysis was performed using the AMSTAR (Assessing the Methodological Quality of Systematic Reviews) checklist, and heterogeneity, excess statistical significance, small study effects, and prediction intervals were assessed.

RESULTS: Odds ratios were small to moderate for all outcomes, ranging from OR=1.2 (95% CI=1.1-1.4) for unprotected sexual intercourse to OR=3.4 (95% CI=2.3-4.8) for sex offending versus non-sex offending. Odds ratios were pooled for psychiatric diagnoses (OR=2.6; 95% CI=2.4-2.8), psychosocial outcomes (OR=2.0; 95% CI=1.6-2.4), and other health diagnoses (OR=1.6; 95% CI=1.4-1.7). Unadjusted population attributable fractions ranged from PAF=1.7% (95% CI=0.7%-3.3%) to PAF=14.7% (95% CI=9.6%-19.9%). Of the five outcomes included in the gender subgroup analysis, two (depression and anxiety) had significantly higher odds ratios for women than men. There was variable quality of evidence and methodology across meta-analyses, indicated by AMSTAR scores, heterogeneity, excess statistical significance, small study effects, and prediction interval. The four outcomes with the highest overall quality scores were eating disorders, posttraumatic stress disorder, schizophrenia, and sex offending versus non-sex offending.

DISCUSSION: This umbrella review found significant effects of childhood sexual abuse for most examined outcomes. However, only three outcomes had significant findings and met rigorous quality standards: eating disorders, posttraumatic stress disorder, and sex offending versus non-sex offending. All other outcomes were either insignificant or had low to moderate quality evidence. Childhood sexual abuse was also found to have a significantly greater effect on the outcomes of anxiety and depression for women than for men. Possible mechanisms for the effects of childhood sexual abuse on the primary outcome findings, such as the Hypothalamic-Pituitary-Adrenal axis and the victim to perpetrator cycle, could be targets for intervention.

Chapter 1: Introduction

1.1 Why Study Childhood Sexual Abuse?

Childhood sexual abuse is a globally prevalent, high-profile, and controversial phenomenon. The World Health Organization defines childhood sexual abuse as, “the involvement of a child in sexual activity that he or she does not fully comprehend and is unable to give informed consent to, or for which the child is not developmentally prepared, or else that violate the laws or social taboos of society” (World Health Organization (2003). However, childhood sexual abuse is not consistently defined, and various other definitions specify cut-off ages, such as 15 or 18 years of age, physical contact during the abuse, or a certain age difference between the victim and perpetrator (Devries et al., 2014; Lloyd & Operario, 2012).

Prevalence estimates for childhood sexual abuse vary widely, depending on the definition, as well as country, gender, methodology, and study quality. In a 2012 report from the International Society for Prevention of Child Abuse and Neglect (ISPCAN), the World Health Organization (WHO) estimated a global prevalence of 5-10% for boys and 20% for girls (von Ins et al., 2012). These estimates are also supported by a 2009 review of 65 studies from 22 countries that found global childhood sexual abuse estimates of 8% for boys and 20% for girls (Pereda, Guilera, Forns, & Gomez-Benito, 2009). According to a report by the University of Minnesota and the U.S. Department of Justice’s Office on Violence Against Women, the Centers for Disease Control and Prevention (CDC) found that 4% of boys and 11% of girls in the United States report having been forced to have sex (University of Minnesota & U.S. Department of Justice Office on Violence Against Women, 2009). While these numbers may be lower than the above-cited global estimates due to nationality, the

definition is also significantly narrower, excluding non-contact forms of abuse, such as exhibitionism and forced pornography viewing as well as certain forms of contact abuse, such as fondling.

Regardless of differences in definitions and prevalence estimates, childhood sexual abuse remains a frequent media headline and cause of public concern and outrage. From scandals involving religious institutions, to sports clubs, to celebrities and prominent public figures, childhood sexual abuse and its potential harmful ramifications are familiar topics to anyone who follows the news. According to these high profile scandals, the harms caused by childhood sexual abuse are numerous, severe, and lasting, with survivors describing posttraumatic symptoms and lifelong physical and mental damage (Benhold, 2016; Blow, 2015; Walsh, 2016).

Given both the prevalence of childhood sexual abuse and its reported severe harm, scientific research and clinical understanding of the mechanisms and effects of childhood sexual abuse are of immediate importance. Thorough scientific investigation into the long-term implications of childhood sexual abuse is critical for ensuring that those who have been abused receive effective treatment and for preventing and mitigating harm to future victims.

1.2 Childhood Sexual Abuse and Later Outcomes

The earliest review on the long-term implications of childhood sexual abuse identified for the current study was published in the late 1980s (Browne & Finkelhor, 1986). While recognizing that it was surveying a relatively young body of literature that suffered from methodological deficiencies, the review found that childhood sexual abuse had been linked to a wide variety of negative outcomes across early primary studies from the 1960s to

the 1980s. This narrative review, focusing exclusively on outcomes for women, found that adults who had experienced childhood sexual abuse were more likely than the general population to experience psychological and related problems such as depression, anxiety, self-destructive behaviour, feelings of low self-esteem, stigma, social isolation, sexual revictimisation, and substance abuse. The review also found some, though less consistent, evidence for various forms of sexual dysfunction.

Researchers since the 1970s have debated whether and to what extent childhood sexual abuse causes suffering or lasting harm to victims (Finkelhor, 1979; Summit & Kryso, 1978), and the most controversial paper in this vein even garnered the attention of the United States Congress. The authors of this 1988 meta-analysis did not find that childhood sexual abuse caused pervasive, intense harm, and they found that those negative effects of childhood sexual abuse that did exist affected girls more than boys. While these findings were controversial, they were perhaps less controversial than the authors' conclusion that argued for a redefinition of childhood sexual abuse to distinguished harmful sexual experiences in childhood from non-harmful experiences, which could be recast with the more neutral term "adult-child sex" (Rind, Tromovitch, & Bauserman, 1998)

In 1999, the article came to the attention of the United States Congress, which passed a resolution denouncing any suggestion made by the article that sex between adults and willing children might be non-harmful and potentially even beneficial (Civic Impulse, 2017). The resolution also formally opposed any legislation or policy to normalize sex between children and adults or to lower the age of consent, and it urged researchers to continue studying childhood sexual abuse with sound methodologies, to provide more scientific data for policy makers. The response from the scientific community was substantial and varied,

with some researchers pointing to methodological weaknesses in the controversial meta-analysis (e.g. the samples and outcomes included were skewed to minimise apparent effect) (Ondersma, Berliner, Chaffin, & Cordon, 2001; Spiegel, 2000; Whittenburg, Tice, Baker, & Lemmey, 2001), others decrying it as an advocacy paper (Dallam, 2001), and yet others defending it as scientifically sound but simply unpalatable (Oellerich, 2000). Public debate surrounding the article has slowly diminished over the years, but as recently as 2007, the authors were still publishing rebuttals to ongoing criticisms of the paper (Rind & Tromovitch, 2007).

Since this surge in debate about the nature and extent of harm that can be attributed to childhood sexual abuse, scientists have continued to publish a steady stream of primary studies and systematic reviews investigating the associations between childhood sexual abuse and a host of health and psychosocial outcomes. Meta-analyses have found significant effects of childhood sexual abuse on depression and anxiety (Amado, Arce, & Herraiz, 2015), HIV risk behaviours (Arriola, Loudon, Doldren, & Fortenberry, 2005), eating disorders, posttraumatic stress disorder, sleep disorders, and suicide attempts (L. P. Chen et al., 2010) obesity (Danese & Tan, 2014), and psychosis (Varese et al., 2012), among others. An umbrella review, or a review of reviews, is an effective tool for synthesizing this growing body of meta-analytic literature to understand which outcomes are most affected by a history of childhood sexual abuse and where future researchers should focus their attention.

1.3 Meta-Analyses and Umbrella Reviews

A meta-analysis is defined as, “the statistical combination of results from two or more separate studies” (J. P. T. Higgins & Green, 2011). While a systematic review refers to a

standardised procedure for reviewing the literature on a topic, a meta-analysis is a specific component of a systematic review, focused on statistically comparing and pooling data from the literature. Meta-analyses allow researchers to aggregate data from multiple primary studies that address the same question, to increase power and precision and to resolve any conflicts that may arise from different results across studies. Once there exists a large enough body of primary studies addressing the effect of a specific exposure or intervention on a specific outcome (such as the effect of childhood sexual abuse on depression), a meta-analysis can be an effective way of drawing key findings and conclusions from multiple disparate analyses. However, to preserve the reliability and accuracy of their findings, meta-analyses must take into consideration the quality of the primary studies being aggregated, potential publication biases that may skew the availability of the literature included in the meta-analysis, and any heterogeneity between studies that could cause a misleading comparison of dissimilar effects.

Along with ease of comparison and increased power, a benefit of using meta-analytic techniques is the rigorous standardisation and documentation that they require. The *Cochrane Handbook for Systematic Reviews of Interventions* provides a detailed guide to procedures and best practices for systematic reviews and meta-analyses. Similarly, PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analyses) provides a checklist of the components that should be documented in any systematic review or meta-analysis (Moher, Liberati, Tetzlaff, & Altman, 2010). PRISMA also recommends that researchers publish their study protocol, describing the study rationale, hypotheses, and planned methodology, in a public online registry, such as PROSPERO, to ensure that methodological decisions are not made arbitrarily or because of known results during the research process (Shamseer, 2015).

An umbrella review is a systematic review of systematic reviews or, if it involves combining statistical results, a meta-analysis of meta-analyses. A more recent and logical extension of meta-analyses, umbrella reviews integrate the results of multiple meta-analyses and can be particularly valuable when there is more than one exposure or outcome of interest (Ioannidis, 2009). In the case of the current study, because meta-analyses have been conducted on a range of long-term outcomes associated with childhood sexual abuse, an umbrella review allows for combination and comparison of the effects of childhood sexual abuse across these different outcomes.

The rationale for using the meta-analytic methodologies of an umbrella review for the following study of childhood sexual abuse outcomes is clear: Firstly, the umbrella review will provide power and accuracy for assessing the effects of childhood sexual abuse because of the large sample sizes provided by meta-analyses and the standardised protocols and reporting guidelines required. Secondly, the umbrella review will pool and compare data across many different outcomes associated with childhood sexual abuse, enabling a clarifying overview of a large body of literature, which may be of interest to a wide audience. Thirdly, because of the rigorous and comprehensive nature of an umbrella review, it will be an effective method of identifying any existing weaknesses or content gaps in the literature and providing recommendations for future research directions.

1.4 Previous Umbrella Reviews on Outcomes of Childhood Sexual Abuse

Previous umbrella reviews have been conducted on outcomes associated with childhood sexual abuse. However, they are more limited in scope than the current review and almost all were published more than five years ago. Many important additions have been

made to the review literature in that time. Thus, an updated review is needed. Of the six identified existing umbrella reviews on childhood sexual abuse outcomes, four of them combined meta-analyses on the effects of childhood sexual abuse on single outcomes. These single outcome umbrella reviews have been conducted for depression (Maniglio, 2010), anxiety disorders (Maniglio, 2013), substance-related disorders (Maniglio, 2011a), and suicide and self-injury (Maniglio, 2011c). While each of these umbrella reviews provides a useful synthesis of meta-analyses for a single outcome, none of them combine or compare findings across outcomes.

Another umbrella review from 2009 does consider a wide range of health and psychosocial outcomes, but it does not meta-analyse those outcomes. The review reports effect size data but does not convert all of the data to the same metric, pool them, or provide any graphical representation, such as a forest plot (Maniglio, 2009). Finally, it does not distinguish between the short-term implications of childhood sexual abuse in children and the long-term implications in adults, as the current umbrella review does.

The most recent umbrella review to assess the effects of childhood sexual abuse on a range of outcomes, published in 2011, also has some limitations (Hillberg, Hamilton-Giachritsis, & Dixon, 2011). To begin with, it is a relatively small umbrella review, including only seven meta-analyses. The review also is limited to mental health outcomes, excluding potential outcomes of interest that would be categorized as physical health conditions or psychosocial outcomes. Like the 2009 umbrella review, it also does not pool or quantitatively synthesize outcomes, though it does consistently report all outcomes as correlation coefficients (Pearson's r). Finally, while the review does provide a quality assessment of the

included meta-analyses, the quality assessment tool employed was not validated, so its implications are unclear.

Because of the limitations of these previous umbrella reviews, and because numerous other meta-analyses studying childhood sexual abuse outcomes have been conducted since their publication (Amado et al., 2015; L. P. Chen et al., 2010; Danese & Tan, 2014; Devries et al., 2014; Hauser, Kosseva, Uceyler, Klose, & Sommer, 2011; Irish, Kobayashi, & Delahanty, 2010; Jespersen, Lalumiere, & Seto, 2009; Lloyd & Operario, 2012; Varese et al., 2012), a new, up-to-date, and methodologically rigorous umbrella review will contribute an important perspective on the current literature as well as directions for future research. The current umbrella review will consider health outcomes, including but not limited to mental health outcomes, as well as psychosocial outcomes. It will focus on the long-term impacts of childhood sexual abuse in adults, distinct from the short-term effects that may be observed in children, immediately following abuse. Also, it will statistically compare and pool effect size data across outcomes, and it will rigorously assess the quality of the data and methods of the included meta-analyses.

1.5 Methodological Challenges

I would like to note some of the methodological challenges and limitations of conducting an umbrella review, as well as of those of studying childhood sexual abuse. Firstly, a critical limitation of all meta-analyses is that they are, “limited by the amount, quality and comprehensiveness of available information in the primary studies” (Ioannidis, 2009). For umbrella reviews, this caveat is compounded because they are limited by the same constraints at the meta-analytic level as well. However, these challenges can be mitigated to a

certain extent by effective quality assessment, preferably employing specific rules or standardised tools, to promote transparency and reproducibility (Greco, Zangrillo, Biondi-Zoccai, & Landoni, 2013). Another challenge to meta-analytic research is accounting for and appropriately handling heterogeneity between included studies. Heterogeneity measures variation between studies beyond that which is attributable to chance. Researchers must be cautious about pooling results across studies when high levels of heterogeneity are present (Greco et al., 2013). Another challenge, publication bias, can occur when the studies included in a meta-analysis do not accurately reflect the true range of findings. This often occurs because smaller studies, studies with non-significant findings, and studies in languages other than English are less likely to be published, published in reputable journals, cited by other authors, or included in major databases. To a certain extent, publication bias can be prevented by ensuring comprehensive literature searches across multiple databases, and when publication bias does occur, it can often be identified with tools such as funnel plots (Greco et al., 2013).

One of the major challenges of researching childhood sexual abuse outcomes is the overarching reliance on retrospective self-report data. While not all studies of childhood sexual abuse use retrospective self-report data, it is the easiest data to collect and the most commonly used. Randomized controlled trials are not an ethically viable option for obvious reasons and prospective studies are difficult and time consuming, yet multiple sources have documented the inconsistency and inaccuracy of adults' reports of their own experiences of childhood sexual abuse. One study found that between two phases of a study questionnaire, conducted four to six weeks apart, 36% of adults who reported extra-familial childhood sexual abuse in phase I denied it in phase II. In a third follow-up questionnaire, the

participants cited a range of reasons for this inconsistency, such as misunderstanding (e.g. confusing intra-familial and extra-familial abuse), emotional motives (e.g. embarrassment or feeling overwhelmed), and practical considerations (e.g. lack of privacy when filling out questionnaire) (Langeland et al., 2015). Another study conducted a 20 year follow-up with adults who had experienced court substantiated cases of childhood sexual abuse and compared them to matched controls. They found significant underreporting of childhood sexual abuse, particularly for male participants, and suggested possible reasons for underreporting, including memory loss, denial, and embarrassment (Widom & Morris, 1997).

Another limitation of researching childhood sexual abuse is the potential for other variables to confound the effects of childhood sexual abuse on later outcomes. Some researchers have found that accounting for environmental factors, such as socioeconomic status or family dynamics and characteristics, significantly diminishes the effects of childhood sexual abuse on later outcomes. For instance, one 1998 study that found that family environment explained nine times more variance in later psychological adjustment than childhood sexual abuse did (Rind et al., 1998). Others, however, have found that the effects of childhood sexual abuse are still significant after adjusting for confounders. For instance, one 2008 study that found that those exposed to childhood sexual abuse had a rate of mental disorders 2.4 times higher than those not exposed to childhood sexual abuse, even after adjusting for a range of potential environmental confounders, including socioeconomic background and parental substance abuse, crime, and education, along with individual level confounders, such as IQ (Fergusson, Boden, & Horwood, 2008). Regardless of the extent to which they account for the effects of childhood sexual abuse, environmental confounders should be adjusted for in studies of childhood sexual abuse outcomes.

1.6 Aims of This Study

In consideration of the limitations of previous studies and the methodological challenges of conducting an umbrella review, this study aims to provide an up-to-date summary of the current state of the literature on the links between childhood sexual abuse and later outcomes, to assess the quality of this body of literature, to identify any gaps that should be the focus of future research, and to compare findings across different outcomes.

Chapter 2: Methods

The protocol for this umbrella review was registered on PROSPERO, the international prospective register of systematic reviews, and can be accessed at https://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016049701.

2.1 Data Sources

The initial search was conducted across four online databases: PsycINFO, PubMed, CINAHL, and Global Health. These databases were selected to maximise the likelihood of finding all relevant articles across psychological and medical journals from published and unpublished literature sources around the world. A keyword search of titles and abstracts was conducted in PsycINFO from 1806 – October 2016, in PubMed from 1809 – October 2016, in CINAHL from 1937 – October 2016, and in Global Health from 1973 – October 2016. Each database search consisted of the same keywords related to sexual abuse (sexual abuse, sexual assault, sexual trauma, victimisation, rape, incest, molestation, sexual crime, victim, maltreatment), childhood (childhood, child, youth, adolescent, young, teen, teenage, teenager), and meta-analysis (meta-analysis, meta-analytic, meta analysis, meta analytic, meta-regression, meta regression, meta-synthesis, metasynthesis, meta synthesis). As an example, the exact search phrase used in PsycINFO was ((sexual abuse OR sexual assault OR sexual trauma OR victimisation OR incest OR rape OR molestation OR sexual crime OR victim OR maltreatment) AND (child OR childhood OR youth OR adolescent OR young OR teen OR teenage OR teenager) AND (systematic review OR meta-analysis OR meta-analytic OR meta analysis OR meta analytic OR meta-regression OR meta regression OR meta-synthesis OR metasynthesis OR meta synthesis)).*ab,ti*. No language or date restrictions were used in the database searches.

A more limited secondary search was conducted for articles published from 2006-October 2016 in PsycINFO to target specific health outcomes that are major contributors to the global burden of disease. Individual searches were conducted for the top ten overall global DALY (Disability-Adjusted Life Year) risks (World Health Organization, 2014), as

well as for the top ten global mental and behavioural health DALYs (National Institute of Mental Health, 2010). The top ten global DALYs were ischaemic heart disease, lower respiratory infections, stroke, preterm birth complications, diarrhoeal diseases, chronic obstructive pulmonary diseases, HIV/AIDS, road injury, unipolar depressive disorders, and birth asphyxia and birth trauma. Searches were not conducted for preterm birth complications or birth asphyxia and birth trauma, because these outcomes necessarily occur before any experiences of childhood sexual abuse and therefore cannot be caused by them. There was also no search conducted for unipolar depressive disorders because relevant meta-analyses on this outcome had already been identified during the initial search.

The top ten behavioural and mental health DALYs were major depressive disorder, anxiety disorders, drug use disorders, alcohol use disorder, schizophrenia, bipolar disorder, dysthymia, autism and Asperger's syndrome, ADHD and conduct disorder, and eating disorders. No searches were conducted for unipolar depressive disorders, major depressive disorder, anxiety disorders, drug use disorders, and eating disorders because relevant meta-analyses on these outcomes had already been identified during the initial search.

For each targeted DALY search, titles and abstracts were searched in PsycINFO using the same sexual abuse and childhood related search terms employed during the initial broader search. However, I excluded meta-analysis-related keywords to generate broader search results and included relevant synonyms and terminology for each targeted DALY. See Table 1 for a complete list of DALY-specific search terms. For example, the exact search phrase used for ischaemic heart disease was ((sexual abuse OR sexual assault OR sexual trauma OR victimisation OR incest OR rape OR molestation OR sexual crime OR victim OR maltreatment) AND (child OR childhood OR youth OR adolescent OR young OR teen OR

teenage OR teenager) AND review AND (ischaemic heart disease OR coronary artery disease OR coronary heart disease)).*ab,ti*.

Table 1: Search terms used in targeted DALY searches

DALY	Search Terms
Ischaemic heart disease	Ischaemic heart disease, coronary artery disease, coronary heart disease
Lower respiratory infections	Lower respiratory infection, pneumonia, lung abscess, acute bronchitis
Stroke	Stroke, transient ischaemic attack
Diarrhoeal diseases	Diarrhoeal disease, diarrhoea, dysentery, cholera
Chronic obstructive pulmonary disease	Chronic obstructive pulmonary disease, COPD, chronic bronchitis, emphysema
HIV/AIDS	HIV, AIDS, human immunodeficiency virus
Road injury	Road injury, traffic injury, road injuries, traffic injuries
Alcohol use disorder	Alcohol use, alcoholism, alcoholic
Schizophrenia	Schizophrenia, schizophrenic
Bipolar disorder	bipolar disorder, manic depression, mania
Dysthymia	Dysthymia, dysthymic disorder, mild depression, chronic depression
Autism or Asperger's syndrome	Autism, autistic, Asperger, ASD
ADHD and conduct disorder	Attention deficit hyperactivity disorder, ADHD, conduct disorder

Finally, the references cited by other umbrella reviews that were identified during the search process were manually searched for additional relevant meta-analyses that may have been overlooked in the initial and targeted DALY searches.

2.2 Study Selection and Inclusion Criteria

Studies were selected based upon predetermined inclusion criteria. First, abstracts were screened for articles that clearly fell outside of the inclusion criteria. Then articles that

had been screened in as potentially relevant, based upon their abstracts, were subject to a full text review.

To be included, articles had to meet the following five criteria: (a) the study must be a meta-analysis or systematic review, (b) the study must report outcome data for victims of childhood sexual abuse, disaggregated from other forms of abuse, (c) the majority of participants included in the study must be adults (operationalized as the average participant age at time of outcome measurement is greater than 18 years for over 70% of the included primary studies), (d) the majority of participants included in the study must have been children at the time of abuse (operationalized as the average participant age at time of abuse is less than 18 years for over 70% of the included primary studies), and (e) the study must provide aggregated quantitative effect size data for health or psychosocial outcomes (e.g. odds ratio, Pearson's r correlation, Cohen's d statistic).

While multiple other umbrella reviews were identified during data collection, they were all excluded from this study in lieu of their component meta-analyses, in order to avoid duplicate data counting (Hillberg et al., 2011; Maniglio, 2009, 2010, 2011a, 2011c, 2013). When multiple meta-analyses reported effect size data for the same health or psychosocial outcome, the most recent meta-analysis was selected, and the older studies were excluded, again to avoid duplicating data. In the rare circumstance that two or more meta-analyses reported effect size data for the same outcome and were also published within one year of each other, the meta-analysis including the greatest number of primary studies was selected, and the others were excluded (Hemmingsson, Johansson, & Reynisdottir, 2014; Jumper, 1995; Li, D'Arcy, & Meng, 2016; Lindert et al., 2014; Mandelli, Petrelli, & Serretti, 2015; Roodman & Clum, 2001; Smolak & Murnen, 2002; Wang, Wu, Yang, & Song, 2015). When

the full text of an article was not available via online journals or databases, the corresponding author for the article was contacted and a request was made for a copy of the full text article. One article was excluded because it was not available online and the author could not be contacted (Dagang, 1997).

Because the meta-analyses meeting the above-listed criteria included discrete effect size statistics from nearly 100 different health and psychosocial outcomes, the remaining outcomes were narrowed down using the following criteria: Of the health outcomes, specific diagnoses (e.g. schizophrenia, fibromyalgia syndrome) and their related symptoms (e.g. psychosis, pain) were include, while sub-categories of diagnosis (e.g. social phobia as a sub-category of anxiety) and other symptoms (e.g. cardiopulmonary symptoms, chronic pelvic pain) were excluded. Of the psychosocial outcomes, the ten behavioural outcomes that included the largest number of primary studies (e.g. adult sexual revictimisation, substance abuse) were included, while non-behavioural psychosocial outcomes (e.g. self-esteem, hostility) and behavioural psychosocial outcomes including fewer primary studies (e.g. recent unprotected anal intercourse, online sexual offending compared with offline sexual offending) were excluded. Of the ten included behavioural outcomes, one outcome (homelessness) was later excluded during data extraction because the study's reported log odds ratio denoted prevalence rather than effect size data, such that the article no longer met inclusion criteria (Sundin & Baguley, 2015).

Fifteen outcomes were excluded because they were non-behavioural psychosocial outcomes (Neumann, Houskamp, Pollock, & Briere, 1996; Paquola, Bennett, & Lagopoulos, 2016; Ulrich, Randolph, & Acheson, 2005). Six behavioural psychosocial outcomes were excluded because they included fewer primary studies than the ten selected behavioural

outcomes (Babchishin, Hanson, & Hermann, 2011; Jespersen et al., 2009; Lloyd & Operario, 2012). Nine health outcomes were excluded because they were sub-categories of other diagnoses (Amado et al., 2015; Li et al., 2016; Ulrich et al., 2005). Fifteen health outcomes were excluded because they were neither diagnoses nor symptoms associated with included diagnoses (Irish et al., 2010; Jumper, 1995; Latthe, Mignini, Gray, Hills, & Khan, 2006; Neumann et al., 1996; Rind & Tromovitch, 1997; Sharpe & Faye, 2006; Ulrich et al., 2005). See Table 2 for a complete list of excluded outcomes.

Table 2: Outcomes excluded from this umbrella review by reason for exclusion

Non-behavioural psychosocial outcomes	Behavioural outcomes with fewer primary studies than the ten selected	Health outcomes that are sub-categories of other diagnoses	Health outcomes that are not diagnoses or symptoms associated with included diagnoses
Alcohol problems	Anal intercourse	Depression and/or anxiety	Cardiopulmonary symptoms
Anger	Online offending compared with normative groups	Dysthymia	Dissociation
Combined hippocampal volume	Offline offending compared with normative groups	Generalized anxiety disorder	Dysmenorrhea
Hostility	Online sex offending compared with offline sex offending	Major depressive disorder	Dyspareunia
Interpersonal problems	Paedophilic versus non-paedophilic sex offending against children	Panic disorder	Gastrointestinal symptoms
Interpersonal sensitivity	Suicide	Paranoia	General health
Left hippocampal volume		Phobia	General sequelae
Locus of control		Social phobia	Index of general symptomatology

Overall adjustment		Specific phobia	Non-cyclical chronic pelvic pain
Right hippocampal volume			Non-epileptic seizures
Self-concept impairment			Obsessive-compulsive symptomatology
Self-esteem			Obsessions and compulsions
Sexual adjustment			Psychological correlates
Sexual problems			Psychological symptomatology
Social adjustment			Psychotic symptoms

Ten health diagnosis outcomes were included in the final analysis (Amado et al., 2015; L. P. Chen et al., 2010; Fossati, Madeddu, & Maffei, 1999; Hauser et al., 2011; Lloyd & Operario, 2012). Eight corresponding health symptoms were included (Amado et al., 2015; Danese & Tan, 2014; Irish et al., 2010; Neumann et al., 1996; Quinones-Munoz, 2001; Ulrich et al., 2005; Varese et al., 2012). Also, the nine behavioural psychosocial outcomes with the largest number of primary studies were also included (after the tenth outcome, homelessness, was excluded due to lack of effect size data) (Arriola et al., 2005; Devries et al., 2014; Jespersen et al., 2009; Klonsky & Moyer, 2008; Neumann et al., 1996). See Table 6 for a complete list of included outcomes.

2.3 Data Extraction

Data extraction was conducted, following a predetermined data extraction form. Data was extracted for the following categories: general information, eligibility criteria, methodology, and study results. For any uncertain answers, a second data extractor (Dr. Rongqin Yu, Postdoctoral Research Fellow in Forensic Psychiatry, University of Oxford)

was consulted to form a consensus. See Table 3 for a complete list of data extraction questions.

Table 3: Data extracted from all meta-analyses included in this umbrella review by category of data

General information	Eligibility criteria	Methodology	Study results
First author	Meta-analysis	Outcomes studied	Number of participants with exposure
Running title	Systematic review	Whether inclusion criteria are stated	Number of participants without exposure
Journal of publication	Childhood sexual abuse disaggregated from other abuse	Number of studies	Number of participants with outcome
Year of publication	Average participant age at time of abuse < 18	Number of participants	Number of participants without outcome
Date that coding was completed	Average participant age at time of assessment > 18 Provides aggregated effect size data for health or psychosocial outcomes	Mean age of participants	Summary statistics reported
		Percentage of participants who are female	Summary statistic confidence intervals
		Percentage of participants from clinical samples	Number of participants in largest primary study
		Percentage of participants from general population	Effect size of largest primary study
		Number of databases used	Confidence interval of largest primary study
		Number of countries	Heterogeneity statistic
		Year range of review	
		Study design (prospective, retrospective, cross-sectional, or mixed)	

		Quality assessment procedures used	
--	--	------------------------------------	--

During the data extraction process, each of the included articles was also assessed for methodological quality, using the eleven-question AMSTAR (A Measurement Tool to Assess Systematic Reviews) checklist, originally developed in 2007 at the Bruyère Research Institute in Ottawa, Canada (Shea et al., 2007). See Table 4 for the list of questions. Along with the eleven basic questions, the checklist also provides additional clarifying details. For instance, question 2, about duplicate study selection and data extraction, specifies that there should be at least two study selectors and two data extractors as well as a procedure for establishing consensus in case of disagreement. For the full checklist, including clarifying details, see Appendix 1.

Table 4: AMSTAR checklist questions

AMSTAR checklist
1. Was an 'a priori' design provided?
2. Was there duplicate study selection and data extraction?
3. Was a comprehensive literature search performed?
4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?
5. Was a list of studies (included and excluded) provided?
6. Were the characteristics of the included studies provided?
7. Was the scientific quality of the included studies assessed and documented?
8. Was the scientific quality of the included studies used appropriately in formulating conclusions?
9. Were the methods used to combine the findings of studies appropriate?
10. Was the likelihood of publication bias assessed?
11. Was the conflict of interest included?

For the purposes of this umbrella review, the criteria for two questions in the AMSTAR checklist, questions 5 and 11, were modified to reflect the nature of the included

studies. In the original AMSTAR checklist, question 5 specifies that a list of both included and excluded studies should be provided for each meta-analysis (Shea et al., 2007). However, because the requirement of providing a comprehensive list of all excluded studies did not seem feasible, I adapted this criterion so that only a list of included studies was required. Question 11 in the original checklist also specifies that sources of funding and support should be indicated for both the meta-analysis itself and for each individual study included in the meta-analysis (Shea et al., 2007). While this may be an important criterion for clinical trials, which may be biased by their funding source, it seemed less relevant for the observational studies under consideration in the present umbrella review. Therefore, I modified this criterion to require only that funding and support be stated for the meta-analysis itself, not for primary studies.

One study included in this analysis (Ulrich et al., 2005) is a replication of a prior meta-analysis (Rind et al., 1998) and therefore did not independently conduct processes such as duplicate data extraction and comprehensive literature search, which the AMSTAR tool assesses. For this reason, the original meta-analysis, which was not otherwise included in the analysis of this umbrella review, was coded to supply data only for questions 1, 2, 3, and 4 of the AMSTAR checklist.

Each included meta-analysis was coded according to the AMSTAR checklist, with each of the eleven questions assigned an answer of *Yes*, *No*, *Can't answer*, or *Not applicable*. The total number of *Yes* answers was tallied up for each study to provide an overall score. Scores of 0-3 were considered low, scores of 4-7 were considered medium, and scores of 8-11 were considered high quality (Shea et al., 2007).

2.4 Data Analysis

Once effect size data were extracted, all effect size statistics and confidence intervals were converted into odds ratios to facilitate synthesis and comparison across outcomes. Ten outcome effect sizes (anxiety, anxiety symptomatology, depression, depressive symptomatology, sex trading, sex with multiple partners, adult sexual revictimisation, unprotected sexual intercourse, borderline personality disorder, and non-suicidal self-injury) were converted from Pearson's r correlations to Cohen's d statistics using the following formula:

$$d = \frac{2r}{\sqrt{1 - r^2}}$$

The confidence intervals and effect sizes for the largest primary study in each meta-analysis were converted from r to d using the same equation (Borenstein, Hedges, Higgins, & Rothstein, 2009). Following this conversion, fifteen outcome effect sizes (the nine converted from r , as well as pain as a continuous outcome, substance abuse, traumatic stress responses, psychological symptoms, and somatization) were converted from Cohen's d statistics to odds ratios using the following formula:

$$OR = e^{\frac{\pi d}{\sqrt{3}}}$$

Finally, one outcome (pain as a categorical outcome) was reported in the meta-analysis as a log odds ratio and was therefore exponentiated for the purpose of comparison with other odds ratios (Borenstein et al., 2009).

Once standardised as odds ratios, the effect sizes were analysed in Stata using the *metan* command and graphed in a stratified forest plot with their confidence intervals. To reduce the risk of duplication when pooling overall odds ratios across outcomes, health symptoms corresponding closely to another included outcome were excluded from the forest

plot (somatization, psychological symptoms, depressive symptomatology, anxiety symptomatology, psychosis, traumatic stress responses, and pain as a continuous outcome). Outcomes were stratified into the categories of psychosocial outcomes, psychiatric diagnoses, and other health outcomes.

2.5 Population Attributable Risk

A population attributable fraction is the proportion of the incidence of a given outcome in the population that can be attributed to a specific exposure. It is a common and often insightful measurement of public health risks and therefore could be a valuable statistic for quantifying the long-term harm caused by childhood sexual abuse. The population attributable fraction is calculated based on the relative risk of a given outcome due to the exposure under consideration and the proportion of the population who have experienced this exposure. It has been argued that different formulae should be used to calculate population attributable fractions, depending on whether or not confounding variables are present and factored into the relative risk (Rockhill, Newman, & Weinberg, 1998). If no confounding factors are known and the relative risk has been calculated accordingly, the following equation can be used to determine population attributable risk:

$$\frac{P_e(RR - 1)}{P_e(RR - 1) + 1}$$

RR is the relative risk of the outcome based on the exposure, and P_e is the proportion of the population under consideration that has been exposed to the risk factor (Rockhill et al., 1998). However, if other confounding variables have been considered in the calculation of a confounder-adjusted relative risk statistic, then the following equation should be used instead:

$$pd\left(\frac{RR - 1}{RR}\right)$$

Here, pd is the proportion of cases in the sample exposed to the risk factor (Rockhill et al., 1998). For the outcomes considered in this umbrella review, relative risk was calculated from odds ratio statistics using the formula:

$$RR = \frac{OR}{1 - p + (p \times OR)}$$

In this formula, p is the proportion of the overall population exposed to the risk factor (Zhang & Yu, 1998). In calculations of both relative risk and population attributable fractions, where the proportion of the overall population exposed to the risk factor was required, a 10% prevalence of childhood sexual abuse was used as a conservative estimate. This was based on the WHO estimation that 5-10% of boys and 20% of girls are exposed to childhood sexual abuse (von Ins et al., 2012).

No outcomes included in this analysis reported entirely adjusted effect size data. The nine that reported some degree of confounding or adjustment were eating disorders, posttraumatic stress disorder, schizophrenia, sleep disorders, somatoform disorders, obesity, suicidality, HIV, and somatization. Of those nine, the six outcomes that provided sufficient primary study data (total exposed and unexposed cases) to calculate the confounder-adjusted attributable fraction were eating disorders, posttraumatic stress disorder, schizophrenia, sleep disorders, somatoform disorders, and HIV. Calculating confounder-adjusted attributable fractions can also be challenging due to inconsistency in how many primary studies are adjusted, how well they are adjusted, and for what confounding variables they are adjusted. Likely due to these challenges, calculations using the confounder-adjusted formula produced attributable fractions that appeared unrealistically inflated. Therefore, the unadjusted formula

was used to calculate the population attributable fractions for all outcomes. 95% confidence intervals for each population attributable fraction were calculated using the same method, first transforming the limits of the 95% confidence interval to relative risk and then using the unadjusted formula to calculate the population attributable fraction. It should be noted that the use of the unadjusted attributable fraction formula for outcomes that are likely to be influenced at least in part by confounding variables renders the resulting risk fractions only general estimates that may not be precise.

2.6 Subgroup Analysis by Gender

A subgroup analysis allows for the comparison of effect sizes between different subsets of the overall sample included in a meta-analysis (J. P. T. Higgins & Green, 2011). The meta-analyses included in this umbrella review were assessed for potential subgroup analyses, and there were studies providing effect size statistics for ten subgroups: gender, definition of childhood sexual abuse, study design, age of abuse, type of abuser, use of force, severity of abuse, location of study, type of assessment, and type of controls. However, data for nine of these ten subgroups were only presented in one meta-analysis each. The only variable that was assessed in multiple meta-analyses, providing sufficient data for a subgroup analysis, was gender. Differential effect size statistics were provided for male and female participants for the outcomes of anxiety, anxiety symptomatology, depression, depressive symptomatology, eating disorders, schizophrenia, and suicidality. Gender-specific data was also provided for only one gender for the outcomes adult sexual revictimisation, sex trading, sex with multiple partners, unprotected sexual intercourse, fibromyalgia syndrome, HIV, sex offending versus non-sex offending, sex offending against children versus adults, and

psychological symptoms. However, none of these single-gender outcomes was included in the subgroup analysis because they provided no comparison groups for the opposite gender. The outcomes of anxiety symptomatology and depressive symptomatology were also excluded from the subgroup analysis to avoid duplication with the anxiety and depression diagnosis outcomes. Therefore, five outcomes were included in the subgroup analysis by gender: anxiety, depression, eating disorders, schizophrenia, and suicidality. When necessary, effect size data were converted from Pearson's r statistics to odds ratios using the conversion formulae listed above. The data were analysed and a stratified forest plot was created using Stata, to compare the individual outcome and pooled effect sizes by gender.

2.7 Analysis of Evidence Quality

To supplement the earlier-discussed AMSTAR quality assessment tool, which evaluates the methodological quality of systematic reviews, calculations were also done for excess statistical significance, heterogeneity, small study effects, and prediction intervals, to assess the quality of the evidence included in the meta-analyses.

Heterogeneity is a measurement of the relative consistency or inconsistency of studies pooled in a meta-analysis. When there is low overlap in the confidence intervals of primary studies, this can indicate heterogeneity. Heterogeneity statistics were reported for 19 out of the 28 outcomes in the meta-analyses included in this study. The heterogeneity statistic I^2 , used in this umbrella review, is reported as a percentage out of 100%, with 25% considered low, 50% considered moderate, and 75% considered high heterogeneity (J. P. Higgins, Thompson, Deeks, & Altman, 2003). When heterogeneity was reported by a meta-analysis as Q or X^2 it was converted to I^2 with the formula:

$$I^2 = \left(\frac{Q - df}{Q} \right) \times 100\%$$

Here, df , the degrees of freedom, is equal to $k-1$ (J. P. T. Higgins & Green, 2011).

None of the meta-analyses included in this umbrella review reported 95% confidence intervals for I^2 . However, given the potential variability in I^2 , providing confidence intervals is considered good practice (Ioannidis, Patsopoulos, & Evangelou, 2007). Therefore, I^2 95% confidence intervals were calculated using the following formula:

$$\exp (\ln(H) \pm 1.965 \times SE(\ln(H)))$$

$SE(\ln(H))$ is calculated with the equation:

$$SE(\ln(H)) = \frac{1}{2} \times \frac{\ln(Q) - \ln(k-1)}{2\sqrt{2Q} - \sqrt{2k-3}}$$

H is the transformation of I^2 , defined as:

$$H^2 = \frac{1}{1 - I^2} = \frac{Q}{k - 1}$$

Once the 95% confidence intervals for H had been calculated, they were transformed back to I^2 using the above equation (Thorlund et al., 2012).

The test for excess statistical significance is a method of comparing the expected statistical significance in a meta-analysis the statistical significance that is actually observed. Excess statistical significance was calculated as a ratio of the overall odds ratio for each outcome meta-analysis to the odds ratio of the largest primary study included in that meta-analysis. In this instance, the overall odds ratio represents the observed statistical significance and the odds ratio of the largest primary study represents the expected statistical significance. If the quality of the primary studies included in the meta-analysis is good and they have been pooled appropriately, one can expect these two odds ratios to be relatively similar. If the two odds ratios are identical, the ratio produced by the excess statistical significance test will

equal one. However, if the odds ratio of the overall meta-analysis is higher than that of the largest individual component study, resulting in a ratio greater than 1, this may be an indication that studies included in the meta-analysis have been selectively reported or analysed. In other words, it may provide evidence of publication bias (J. P. T. Higgins & Green, 2011; Ioannidis & Trikalinos, 2007).

To supplement the evaluation of excess statistical significance, another measurement of publication bias, small study effects, was calculated by plotting the effect size of each outcome against the size of the meta-analysis. When a scatter plot is created with the study size plotted on the y-axis and the effect size plotted on the x-axis, one can expect to see a roughly funnel-shaped formation of data points because analyses conducted with smaller sample sizes are likely to be less precise and yield a wider array of odds ratios. If the points on the scatter plot do not form a funnel shape, it can be an indication that some studies may be missing due to publication bias (Egger, Smith, Schneider, & Minder, 1997; Sterne et al., 2011). A comparison of different metrics of study size and effect size concluded that using standard error as a measure of study size and log odds ratio as a measure of effect size provides the most accurate and informative funnel plot (Sterne & Egger, 2001). For this reason, standard error and log odds ratios were used as the y-axis and x-axis in this study.

When sufficient data were available, prediction intervals were calculated to determine the range of likely future scores for studies of the same outcome. A prediction interval, like a confidence interval, is a range of possible scores. However, while a 95% confidence interval gives a range of scores in which the true population parameter can be said to fall with 95% certainty, a 95% prediction interval gives a range of scores in which a future sample statistic can be said to fall with 95% certainty. If the prediction interval includes the null odds ratio of

1, this suggests that future studies might find results indicating that the exposure produced no effect or the opposite effect on the outcome under consideration (Inthout, Ioannidis, Rovers, & Goeman, 2016). 95% prediction intervals were calculated using the following formula:

$$\log(OR) \pm t_{1-0.05/2, k-1} \times SD_{PI}$$

Then, the resulting upper and lower prediction interval limits were exponentiated (Inthout et al., 2016). SD_{PI} , the standard deviation of the prediction interval, is calculated:

$$SD_{PI} = \sqrt{(\tau^2 + SE^2)}$$

τ^2 is a measure of heterogeneity and SE^2 is the standard error of the log odds ratio, estimated by dividing the difference between the upper and lower limits of the 95% confidence interval by 3.92 (Inthout et al., 2016).

Because heterogeneity was generally reported by the I^2 statistic, rather than τ^2 in the included meta-analyses, the following formula was used to convert I^2 to τ^2 :

$$\tau^2 = s^2 \frac{I^2}{100 - I^2}$$

s^2 is the typical study variance, calculated by

$$s^2 = \frac{\sum w_i(k-1)}{(\sum w_i)^2 - \sum w_i^2}$$

In this formula, w_i is the inverse of the variance of a given primary study (Inthout et al., 2016). For each primary study, the odds ratio and confidence interval limits were converted to log odds, and the primary study's standard error was calculated by dividing the log of the lower confidence interval subtracted from the log of the upper confidence interval by 3.92. The standard error was squared to determine the primary study variance, the reciprocal of which provided the inverse variance, w_i . When necessary, effect sizes reported as Cohen's d at the primary study level were converted to odds ratios using the above given formula.

Once I^2 , excess statistical significance, small study effects, and prediction intervals had been calculated, a quality analysis table was created, including the above measures of data quality as well as the AMSTAR evaluation of meta-analytic methodological quality. Each outcome was assigned a score of 0 if I^2 was high ($> 75\%$) and 1 if I^2 was not high ($<75\%$) (J. P. Higgins et al., 2003). It was assigned a score of 0 if there was evidence for excess statistical significance (ratio of observed to expected odds ratios >1) and a score of 1 if there was no evidence for excess statistical significance (ratio of observed to expected odds ratios <1), a score of 0 if the prediction interval included the null and a score of 1 if it excluded the null, and a score of 0 if the meta-analysis reported evidence of small study effects and a score of 1 if the original meta-analysis reported no evidence of small study effects (for instance, via a funnel plot of primary studies). Finally, each outcome was assigned a score of 0 if the AMSTAR score was below 5 and a score of 1 if the AMSTAR score was above 5. The five binary quality analysis scores were then summed to determine an aggregate quality score within the range of 0-5, with 0 designating lowest overall quality and 5 designating highest overall quality.

Chapter 3: Results

3.1 Study Characteristics

The initial database and manual search yielded 822 articles, with no additional articles identified by the targeted DALY searches. See Figure 1. The final meta-review included 15 articles with 3,963,930 participants across 28 outcomes and 505 primary studies. However, there was some duplication of participants and primary studies for the outcomes of anxiety and anxiety symptomatology as well as for depression and depressive

symptomatology. Also, the outcomes of somatization, substance abuse, and traumatic stress responses did not provide data on number of participants. Of the included outcomes, four provided data from exclusively female samples (adult sexual revictimisation, sex trading, sex with multiple partners, and unprotected sexual intercourse) and three provided data from exclusively male samples (HIV, psychological symptoms, and sex offending against children versus adults). Another outcome, sex offending versus non-sex offending, provided almost exclusively male data, with only 22 female participants out of a total sample of 2798. All other outcomes provided data for both male and female participants.

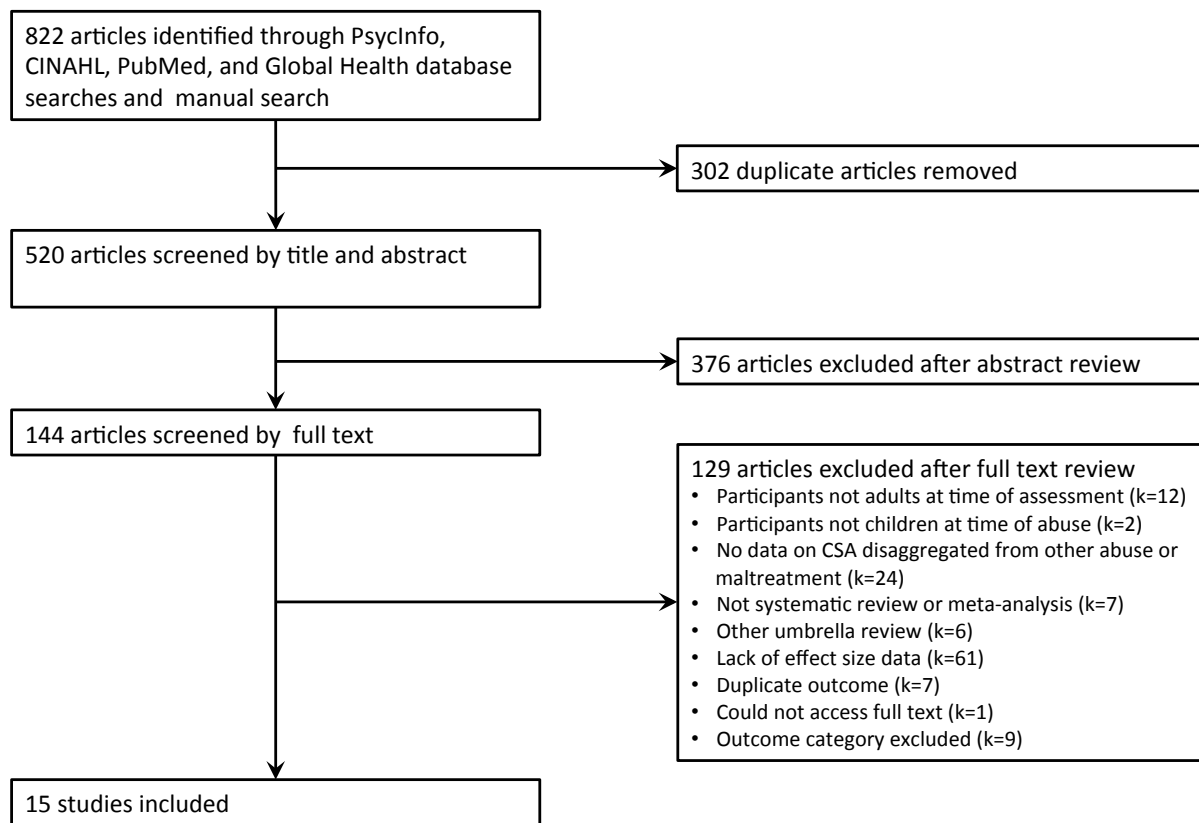


Figure 1: PRISMA flow chart of studies selected for inclusion in this umbrella review

The included meta-analyses ranged in publication date from 1996 to 2015 and included primary studies published from 1975 to 2014. While some articles only provided

data on a single childhood sexual abuse outcome, such as the meta-analyses on obesity (Danese & Tan, 2014) and psychosis (Varese et al., 2012), others provided data on multiple relevant outcomes, such as the study including eating disorders, posttraumatic stress disorder, schizophrenia, sleep disorders, and somatoform disorders (L. P. Chen et al., 2010) and the study including adult sexual revictimisation, sex trading, sex with multiple partners, and unprotected sexual intercourse (Arriola et al., 2005). One meta-analysis was based on a single primary study (sleep disorders) while the greatest number of primary studies included in a meta-analysis was 87 (depression). The smallest sample size included was 140 participants (sleep disorders) and the largest was 3,131,503 participants (schizophrenia). For more details on the characteristics of the included studies, see Table 5.

Table 5: Characteristics of all meta-analyses included in this umbrella review

Author	Year of publication	Number of databases searched	Outcome	Number of primary studies	Sample size	Year range of review
Amado et al.	2015	4	Anxiety	62	93075	1995-2014
			Anxiety symptomatology	41	24270	1995-2014
			Depression	87	123735	1988-2014
			Depressive symptomatology	59	33293	1988-2014
Arriola et al.	2005	4	Adult sexual revictimisation	21	20956	1975-2001
			Sex trading	23	14996	1975-2001
			Sex with multiple partners	23	16560	1975-2001
			Unprotected sexual intercourse	16	11770	1991-2001
Chen et al.	2010	9	Eating disorders	11	7468	1990-2008
			Posttraumatic stress disorder	3	788	1999-2006
			Schizophrenia	3	3131503	1997-2004
			Sleep disorders	1	140	1988-1988

			Somatoform disorders	3	308	1997-2008
Danese & Tan	2014	3	Obesity	26	161195	
Devries et al.	2014	20	Suicidality*	9	8733	1993-2010
Fossati et al.	1999	2	Borderline personality disorder	21	2479	1987-1994
Hauser et al.	2011	4	Fibromyalgia syndrome	10	1487	1995-2010
Irish et al.	2010	3	Pain (continuous)	9	4934	1992-2007
			Pain (categorical)	12	222893	1992-2007
Jespersen et al.	2009	3	Sex offending against children versus adults**	15	2296	1979-2003
			Sex offending versus non-sex offending	17	2798	1987-2003
Klonsky & Moyer	2008	3	Non-suicidal self-injury	45 samples from 43 studies	13687	1980-2006
Lloyd & Operario	2012	6	HIV	5	7796	1995-2009
Neuman et al.	1996	1	Substance abuse	14		
			Traumatic stress responses	4		
Quinones-Munoz	2001	3	Psychological symptoms	5	3720	1992-1998
Ulrich et al.	2005	5	Somatization	11		
Varese et al.	2012	4	Psychosis	20	53050	

*Originally reported as “suicide or suicide attempts”, but simplified for the purposes of this analysis

**Originally reported as “sex offending against adults versus children”, so the effect size data included in the following analysis is the reciprocal of the data reported in the original meta-analysis

3.2 Meta-Analysis

There was a small effect size, with $OR < 1.7$ (Chen, Cohen, & Chen, 2010), of childhood sexual abuse on eight outcomes: schizophrenia, obesity, pain (categorical), HIV, somatization, sex trading, sex with multiple partners, and unprotected sexual intercourse. The smallest effect size was unprotected sexual intercourse ($OR = 1.2$; 95% $CI = 1.1-1.4$).

Nineteen out of 28 outcomes had small to medium effect sizes of $1.7 < OR < 3.5$ (H. Chen et al., 2010). These outcomes were anxiety, anxiety symptomatology, depression, depressive symptomatology, eating disorders, posttraumatic stress disorder, somatoform disorders, borderline personality disorder, fibromyalgia syndrome, pain (continuous), traumatic stress responses, psychological symptoms, psychosis, adult sexual revictimisation, suicidality, sex offending versus non-sex offending, non-suicidal self-injury, substance abuse, and sex offending against children versus adults.

There was a medium to large effect size of $3.5 < OR < 6.7$ for only one childhood sexual abuse outcome, viz. sleep disorders ($OR=16.2$, 95% $CI=2.1-126.8$) (H. Chen et al., 2010). However, because the meta-analytic odds ratio for sleep disorders is based on a single primary study, has the smallest sample size included in this review, and also has a very large 95% confidence interval, I am considering the sleep disorders data to be inaccurate outliers, and they will be excluded from any following analyses. The next largest effect size included in this umbrella review was for the outcome sex offending versus non-sex offending ($OR=3.4$; 95% $CI=2.3-4.8$). See Table 6 for a complete list of results.

Table 6: Effect sizes of childhood sexual abuse on health and psychosocial outcomes and quality measurements

Outcome	Odds ratio	95% CI ^a	95% PI ^b	I^2	I^2 95% CI	ESS ^c ratio
Adult sexual revictimisation	1.9	[1.6, 2.4]		94%	[93%, 94%]	1.4
Anxiety	2.7	[2.5, 2.8]				0.5
Anxiety symptomatology	2.0	[1.9, 2.0]				1.5
Borderline personality disorder	2.9	[2.5, 3.3]		30%	[25%, 36%]	
Depression	2.4	[2.4, 2.6]				0.6

Depressive symptomatology	2.0	[1.9, 2.0]				1.5
Eating disorders	2.7	[2.0, 3.6]	[1.5, 4.8]	20%	[14%, 26%]	1.2
Fibromyalgia syndrome	2.0	[1.4, 2.8]	[1.0, 3.8]	20%	[13%, 26%]	1.7
HIV	1.5	[1.2, 2.0]	[0.9, 2.8]	50%	[35%, 61%]	1.1
Non-suicidal self-injury	2.4	[2.1, 2.7]		51%	[47%, 55%]	1.4
Obesity	1.4	[1.3, 1.6]		87%	[85%, 89%]	1.1
Pain (categorical)	1.7	[1.5, 1.9]				1.3
Pain (continuous)	2.0	[1.3, 3.0]				0.9
Posttraumatic stress disorder	2.3	[1.6, 3.4]	[1.0, 5.4]	0%	[0%-0%]	1.0
Psychological symptoms	1.7	[1.5, 2.0]	[0.7, 4.3]	77%	[69%, 84%]	1.1
Psychosis	2.4	[2.0, 2.9]		45%	[38%, 51%]	1.4
Schizophrenia	1.4	[0.8, 2.3]	[0.4, 4.3]	0%	[0%, 0%]	0.9
Sex offending against children versus adults	2.0	[1.4, 2.9]	[0.5, 7.1]	70%	[65%, 75%]	1.7
Sex offending versus non-sex offending	3.4	[2.3, 4.8]	[1.1, 10.3]	54%	[48%, 60%]	
Sex trading	1.6	[1.2, 2.0]		76%	[73%, 79%]	0.9
Sex with multiple partners	1.6	[1.2, 2.1]		92%	[91%, 93%]	1.4
Somatization	1.6	[1.3, 1.9]		40%	[31%, 48%]	
Somatoform disorders	1.9	[0.8, 4.5]	[0.3, 14.1]	4%	[1%, 7%]	1.9
Substance abuse	2.1	[1.8, 2.5]				
Suicidality	2.4	[1.9, 3.1]	[0.2, 37.9]	88%	[85%, 90%]	0.9
Traumatic stress responses	2.6	[2.2, 2.9]				
Unprotected sexual intercourse	1.2	[1.1, 1.4]		51%	[44%, 57%]	1.0

a: CI=confidence interval, b: PI=prediction interval, c: ESS=Excess statistical significance

Odds ratios were pooled by category for psychiatric diagnoses, other health outcomes and psychosocial outcomes (excluding symptoms, to avoid redundancy). Of these pooled effect sizes, other health diagnoses had the smallest odds ratio (OR=1.6; 95% CI=1.4-1.7).

Psychosocial outcomes provided the next largest pooled odds ratio (OR=2.0; 95% CI=1.6-2.4). And the largest pooled odds ratio was for psychiatric diagnoses (OR=2.6; 95% CI=2.4-2.8). However, these pooled odds ratios should be interpreted with caution because of high levels of heterogeneity and large p-values. The pooled odds ratios for psychiatric diagnoses and psychosocial outcomes both had relatively small *p*-values ($p=0.009$ and $p<0.001$ respectively), but both also had moderate or high levels of heterogeneity ($I^2=65.2\%$ and $I^2=90.3\%$ respectively). The pooled odds ratio for other health outcomes, conversely, had relatively low heterogeneity ($I^2=29.7\%$) but a high p-value ($p=0.234$). See Figure 2.

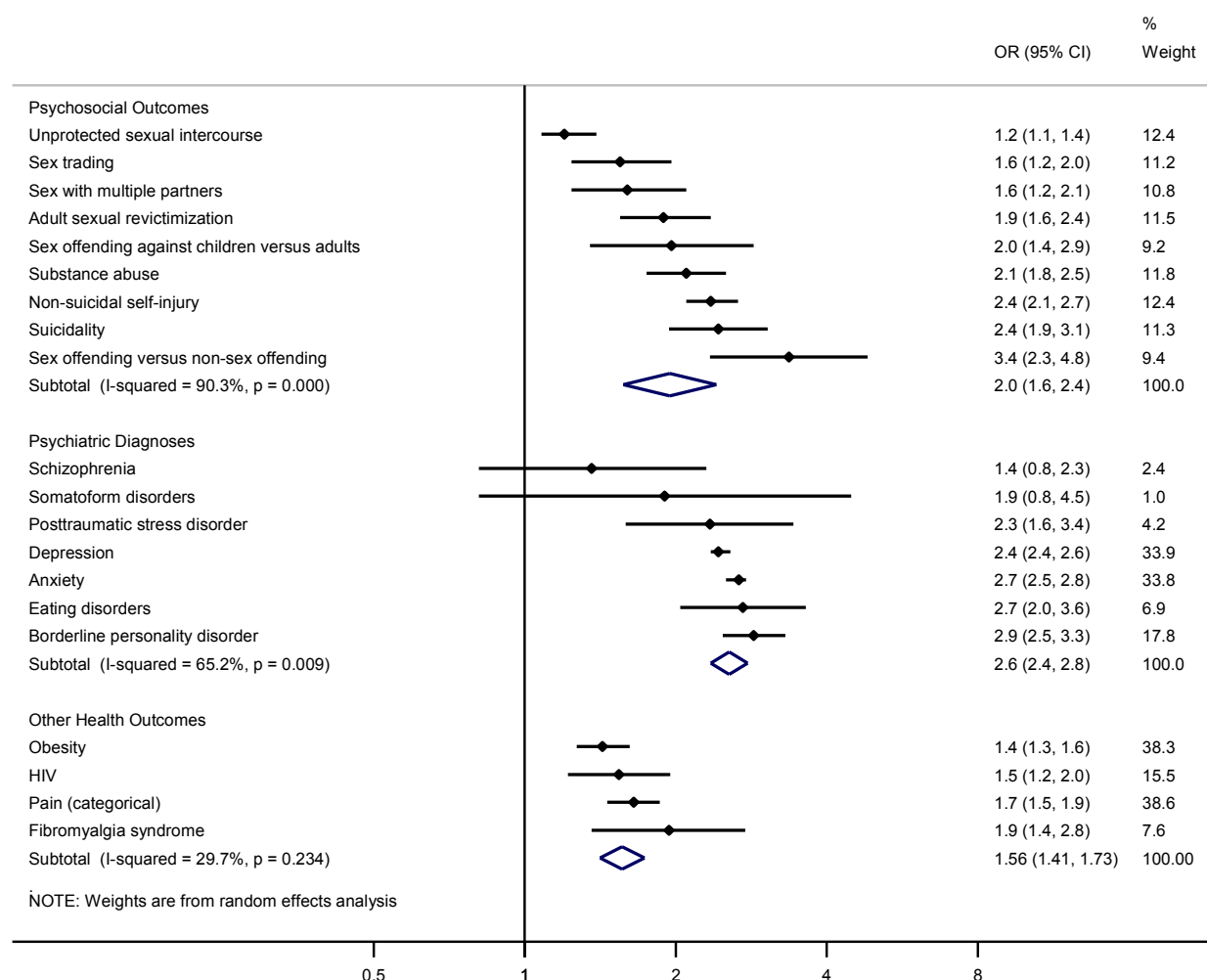


Figure 2: Odds of multiple health and psychosocial outcomes in individuals who have been sexually abused as children compared with the general population

3.3 Population Attributable Risk

Population attributable fractions were calculated using the confounder-unadjusted attributable risk formula and the conservative estimate of a 10% prevalence of childhood sexual abuse in the general population. The outcome with the lowest population attributable fraction was unprotected sexual intercourse (PAF=1.7%; 95% CI=0.7%-3.3%) and the outcome with the highest population attributable fraction was sex offending versus non-sex offending (PAF=14.7%; 95% CI=9.6%-19.9%). See Table 7 for complete results. While the

10% childhood sexual abuse prevalence estimate is conservative, based on the WHO estimate of 20% in girls and 5-10% in boys (von Ins et al., 2012), the lack of accounting for confounding variables in the data and formula may have led to inflated attributable fractions.

Table 7: Percentage of health and psychosocial outcomes in the general population that may be attributable to childhood sexual abuse

Outcome	Population Attributable Fraction	95% Confidence Interval
Sex offending versus non-sex offending	14.7%	9.6%-19.9%
Borderline personality disorder	12.4%	10.4%-14.4%
Eating disorders	11.7%	7.8%-15.8%
Anxiety	11.4%	10.6%-11.9%
Traumatic stress responses	10.9%	8.9%-12.7%
Depression	10.1%	9.7%-10.9%
Suicidality	10.1%	7.2%-13.3%
Psychosis	9.8%	7.4%-12.4%
Non-suicidal self-injury	9.7%	8.2%-11.4%
Posttraumatic stress disorder	9.6%	4.8%-15.0%
Substance abuse	8.2%	5.9%-10.6%
Pain (continuous)	7.8%	2.6%-13.1%
Anxiety symptomatology	7.3%	6.9%-7.8%
Depressive symptomatology	7.3%	6.9%-7.8%
Sex offending against children versus adults	7.3%	3.0%-12.4%
Fibromyalgia syndrome	7.2%	3.0%-11.8%
Adult sexual revictimisation	6.9%	4.5%-9.7%
Somatoform disorders	6.9%	-1.8%-18.8%
Psychological symptoms	5.7%	4.0%-7.8%
Pain (categorical)	5.2%	3.8%-6.7%
Sex with multiple partners	4.8%	2.1%-8.2%
Sex trading	4.5%	2.1%-7.3%
Somatization	4.5%	2.5%-6.6%
HIV	4.4%	1.9%-7.2%
Obesity	3.6%	2.3%-5.0%
Schizophrenia	3.0%	-1.8%-9.4%
Unprotected sexual intercourse	1.7%	0.7%-3.3%

3.4 Subgroup Analysis by Gender

There were five outcomes that provided separate effect size data for both male and female subgroups (depression, anxiety, eating disorders, schizophrenia, and suicidality). Four out of five outcomes had higher effect sizes for women than men (anxiety, depression, eating disorders, and schizophrenia). The overall pooled odds ratio for childhood sexual abuse outcomes in women (OR=2.0; 95% CI=1.9-2.1) was also higher than overall pooled odds ratio for childhood sexual abuse outcomes in men (OR=1.5; 95% CI=1.4-1.6). However, these pooled findings should be considered tentative at best because significantly more primary studies provided data on women (k=81) than on men (k=24), three outcomes (eating disorders, schizophrenia, and suicidality) only included small numbers of primary studies and had large 95% confidence intervals with substantial overlap between genders, and both the male and female pooled odds ratios had moderate to high levels of heterogeneity ($I^2=71.2\%$ and $I^2=84.6\%$ respectively). Therefore, the only gender outcomes that can be asserted with confidence are the greater effect sizes of childhood sexual abuse on depression in women (OR=2.0; 95% CI=1.9-2.0) than in men (OR=1.5; 95% CI=1.4-1.6) and on anxiety in women (OR=2.0; 95% CI=1.9-2.0) than in men (OR=1.6; 95% CI=1.4-1.7). See Figure 3.

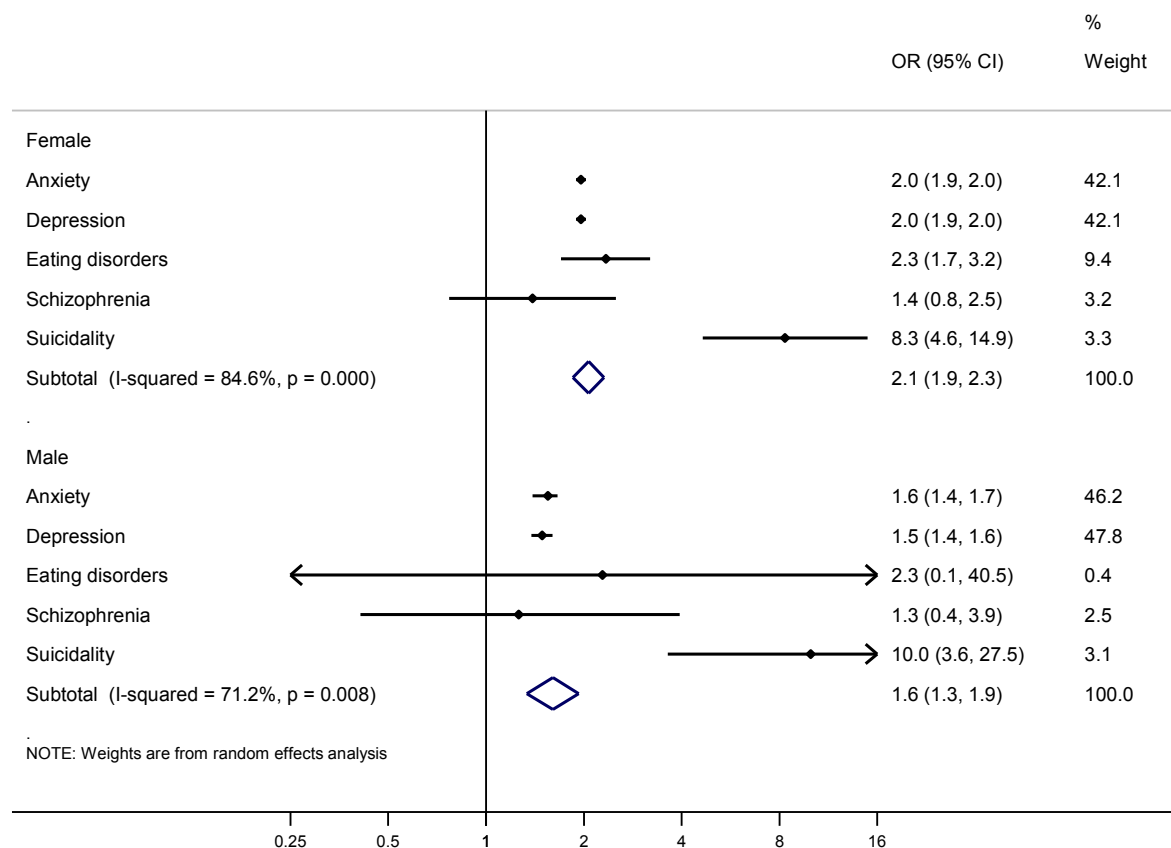


Figure 3: Comparison of odds ratios of multiple long-term outcomes of childhood sexual abuse across genders

3.5 AMSTAR

Using the eleven-question AMSTAR checklist for the methodological quality assessment of meta-analyses, each of the fifteen included studies was given a score from 0-11, with each *yes* answer contributing one point to the overall score. Scores ranged from 2-11, with the highest number of studies receiving moderate scores between 4 and 7. The outcomes with the lowest AMSTAR scores were borderline personality disorder, traumatic stress responses, and substance abuse, while the outcome with the highest score was psychosis. See Table 8 and Figure 4 for full details.

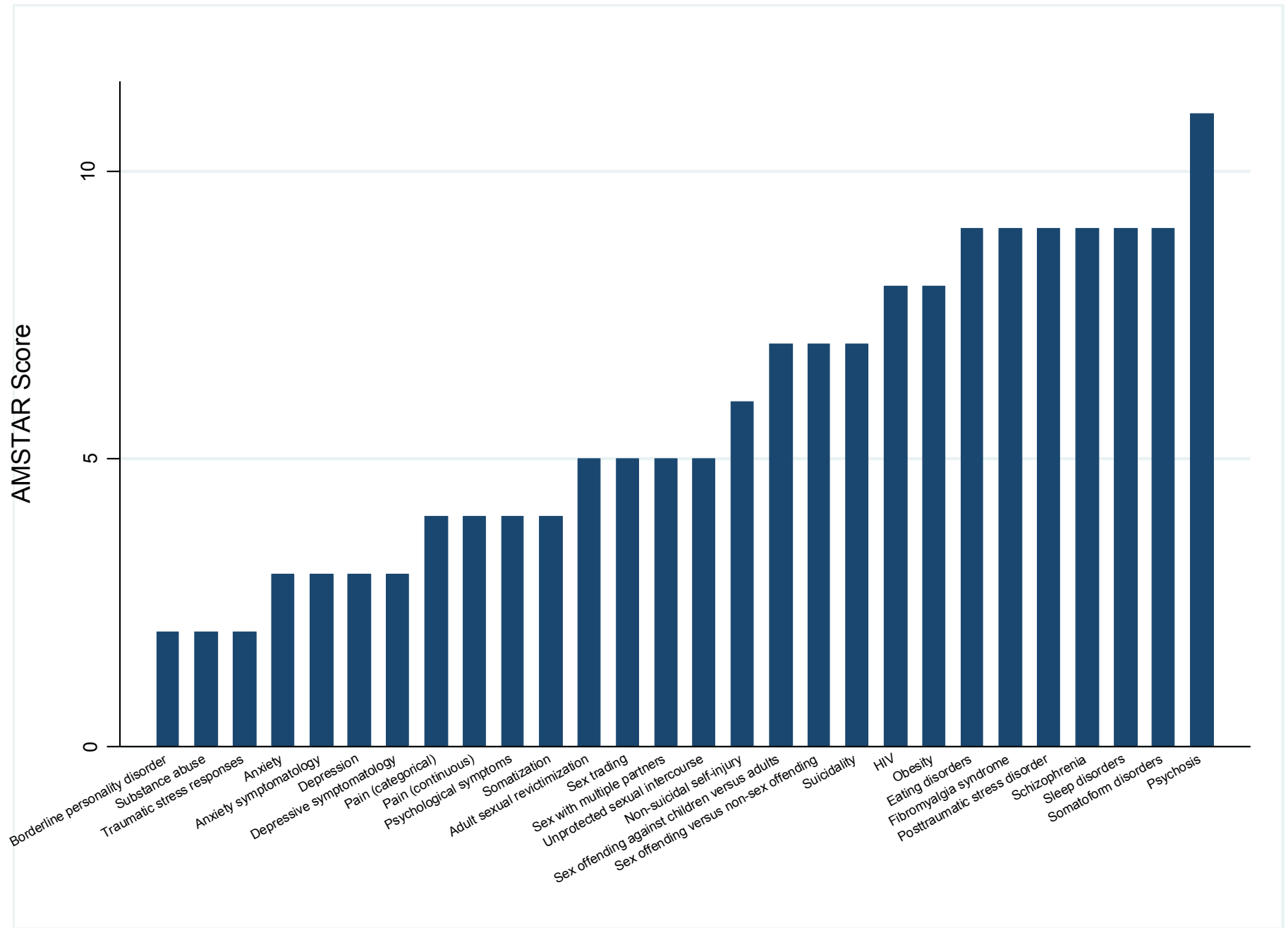


Figure 4: AMSTAR scores by outcome

3.6 Heterogeneity

There were six outcomes that exhibited a high level of heterogeneity, with I^2 statistics greater than 75% (adult sexual revictimisation, obesity, psychological symptoms, sex trading, sex with multiple partners, and suicidality). Five outcomes reported moderate levels of heterogeneity, with I^2 statistics between 50% and 75% (HIV, non-suicidal self-injury, sex offending against children versus adults, sex offending versus non-sex offending, and unprotected sexual intercourse). Eight outcomes had low levels of heterogeneity, with I^2 statistics below 50% (borderline personality disorder, eating disorders, fibromyalgia syndrome, posttraumatic stress disorder, psychosis, schizophrenia, somatization, and somatoform disorders). The eight remaining outcomes reported neither I^2 nor Q heterogeneity statistics. See Table 6 for complete I^2 data.

It is worth noting that heterogeneity calculations for meta-analyses with small samples or small numbers of primary studies have low power and may yield low I^2 statistics that do not necessarily guarantee low heterogeneity (J. P. T. Higgins & Green, 2011). In this umbrella review, the three meta-analyses producing the lowest I^2 scores (posttraumatic stress disorder, schizophrenia, and somatoform disorders) only included three primary studies each, so these results should be interpreted with caution.

3.7 Excess Statistical Significance

An outcome is considered to display evidence of excess statistical significance, here measured as the ratio of the overall odds ratio of a meta-analysis to the odds ratio of the largest component primary study in that meta-analysis, if the resulting ratio is greater than one. There were eleven outcomes in this analysis with clear evidence of excess

statistical significance with ratios ranging from 1.2 to 1.9 (adult sexual revictimisation, anxiety symptomatology, depressive symptomatology, eating disorders, fibromyalgia syndrome, non-suicidal self-injury, pain as a categorical outcome, psychosis, sex offending against children versus adults, sex with multiple partners, and somatoform disorders). Another three outcomes (HIV, obesity, and psychological symptoms) had less clear evidence of excess statistical significance, with ratios of 1.1. There were a further nine outcomes with excess statistical significance ratios less than or equal to one, implying no evidence of excess statistical significance. Finally, for the remaining four outcomes, excess statistical significance could not be calculated, due to lack of data on the largest component primary study. See Table 6 for complete data on excess statistical significance and Figure 5 for a plot of the evidence of excess statistical significance for each outcome.

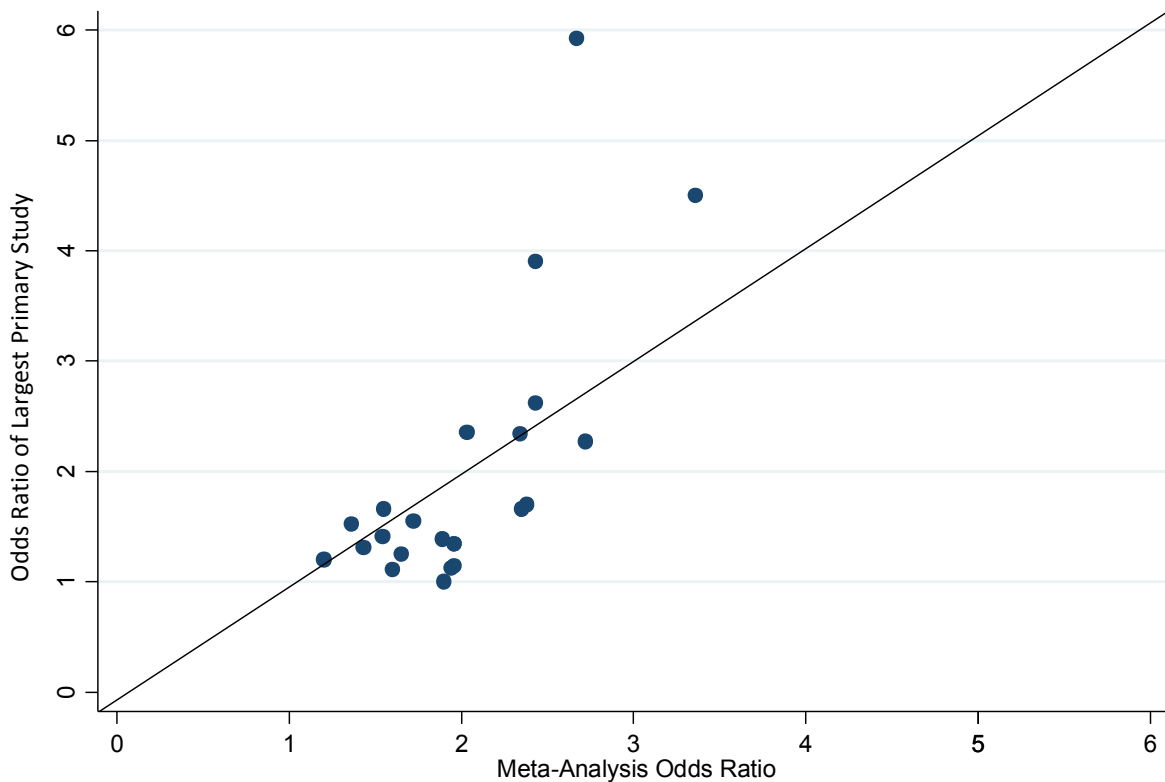


Figure 5: Evidence of excess statistical significance, shown as a ratio of the odds ratios of meta-analyses of individual childhood sexual abuse outcomes versus the odds ratios of the largest primary studies in those meta-analyses

3.8 Small Study Effects

In the funnel plot for small study effects, using the standard error for the y-axis and the log odds ratio for the x-axis, the plotted points are scattered relatively evenly around the mean in a formation expected of a funnel plot. While the individual points in the funnel plot cluster near the top due to small standard errors, this formation does not display a pattern that suggests the presence of small study effects for this umbrella review. See Figure 6.

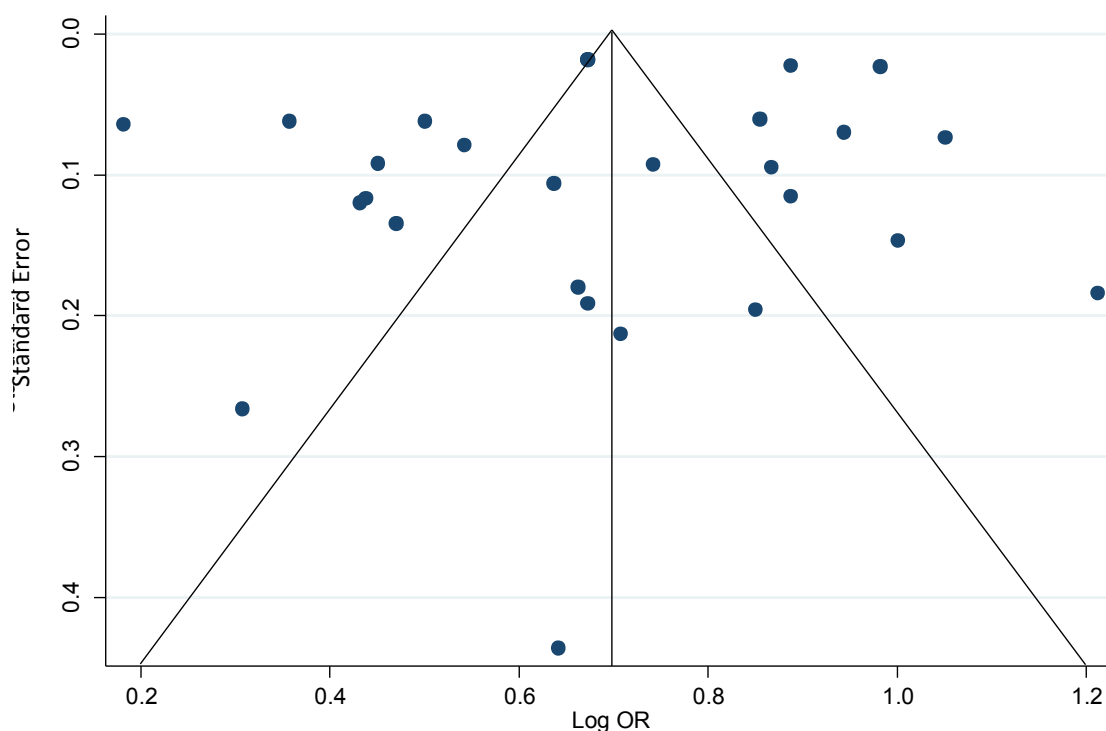


Figure 6: Funnel plot for evidence of small study effects, comparing the effect sizes of meta-analyses to the size of those meta-analyses

Each meta-analysis was also assessed for documentation of small study effects among primary studies. Six out of fifteen meta-analyses used funnel plots to determine evidence of publication bias in the primary studies. Of these six meta-analyses, evidence for small study effects was found for only two outcomes (non-suicidal self-injury and obesity).

3.9 Prediction Intervals

There was sufficient data to calculate prediction intervals for ten outcomes (schizophrenia, HIV, psychological symptoms, somatoform disorders, fibromyalgia

syndrome, sex offending against children versus adults, posttraumatic stress disorder, suicidality, eating disorders, and sex offending versus non-sex offending). See Figure 7. Of these ten outcomes, seven prediction intervals (schizophrenia, HIV, psychological symptoms, somatoform disorders, fibromyalgia syndrome, sex offending against children versus adults, and suicidality) included the null, with the lower bound of the prediction interval less than or equal to one. This suggests that a future study could find that childhood sexual abuse exposure has either no effect or even a protective effect on the outcome in question. However, three outcomes (posttraumatic stress disorder, eating disorders, and sex offending versus non-sex offending) had prediction intervals excluding the null hypothesis, with lower limits greater than one. For these three outcomes, there is at least a 95% chance that a given future study will find that exposure to childhood sexual abuse increases the likelihood of the outcome.

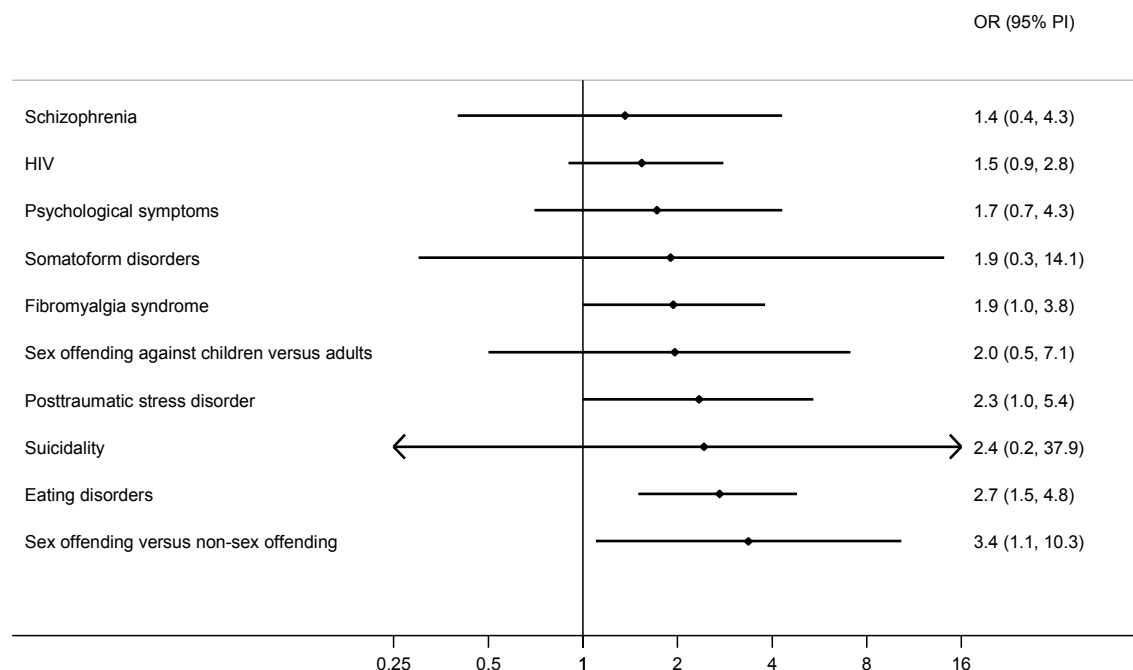


Figure 7: 95% prediction intervals for the potential odds ratios of future studies of childhood sexual abuse outcomes

3.10 Overall Quality Analysis

In the overall quality analysis, equally weighting prediction intervals, heterogeneity, excess statistical significance, small study effects, and AMSTAR scores, each outcome was assigned a quality score from 0-5. Eight outcomes received the lowest score of 0, meaning that their prediction intervals included the null, they had high heterogeneity scores, there was evidence of excess statistical significance, and they had low AMSTAR scores, or there was insufficient data provided in the original meta-analysis to calculate these quality measures. The eight low-scoring outcomes were adult sexual revictimisation, anxiety symptomatology, depressive symptomatology, pain (categorical), psychological symptoms, sex with multiple partners, substance abuse, and traumatic stress responses. Two outcomes received high overall quality scores of 4. These outcomes were eating disorders and schizophrenia. And two outcomes received the

highest scores of 5. These outcomes were posttraumatic stress disorder and sex offending versus non-sex offending. The remaining sixteen outcomes all received moderate quality assessment scores ranging from 1 to 3. See Table 9 for more details.

Table 9: Overall quality analysis results and scores for childhood sexual abuse outcomes

Outcome	Prediction interval excludes null value	Heterogeneity	Evidence of excess statistical significance	Evidence of small study effects	AMSTAR	Overall score
Adult sexual revictimisation		High	Yes		Low	0
Anxiety			No		Low	1
Anxiety symptomatology			Yes		Low	0
Borderline personality disorder		Low			Low	1
Depression			No		Low	1
Depressive symptomatology			Yes		Low	0
Eating disorders	Yes	Low	Yes	No	High	4
Fibromyalgia syndrome	No	Low	Yes	No	High	3
HIV	No	Low	Yes		High	2
Non-suicidal self-injury		Low	Yes	Yes	High	2
Obesity		High	Yes	Yes	High	1
Pain (continuous)			No		Low	1
Pain (categorical)			Yes		Low	0
Posttraumatic stress disorder	Yes	Low	No	No	High	5
Psychological symptoms	No	High	Yes		Low	0
Psychosis		Low	Yes	No	High	3
Schizophrenia	No	Low	No	No	High	4
Sex offending against children versus adults	No	Low	Yes		High	2

Sex offending versus non-sex offending	Yes	Low	No	No	High	5
Sex trading		High	No		Low	1
Sex with multiple partners		High	Yes		Low	0
Somatization		Low			Low	1
Somatoform disorders	No	Low	Yes	No	High	3
Substance abuse					Low	0
Suicidality	No	High	No	No	High	2
Traumatic stress responses					Low	0
Unprotected sexual intercourse		Low	No		Low	2

Chapter 4: Discussion

4.1 Effect Size Findings

In this umbrella review, all assessed childhood sexual abuse outcomes yielded small to moderate odds ratios, ranging from 1.2 (unprotected sexual intercourse) to 3.4 (sex offending versus non-sex offending). None of the outcomes had odds ratios less than or equal to 1. Furthermore, only two outcomes, schizophrenia and somatoform disorders, had 95% confidence intervals that included the null (95% CI=0.8-2.3, and 95% CI=0.8-4.5 respectively). The only outcome that produced a large odds ratio, sleep disorders (OR=16.2), was discounted as an outlier due to its significantly smaller sample size and large 95% confidence interval. However, three outcomes had both significant and high quality findings: eating disorders, posttraumatic stress disorder, and sex offending versus non-sex offending. These three outcomes had high overall quality analysis scores of 4 or 5, as well as confidence intervals and prediction intervals that excluded the null hypothesis (an odds ratio less than or equal to 1).

One conclusion that could be drawn from the consistently small to moderate effect sizes for all outcomes is that childhood sexual abuse is a nonspecific risk factor for negative health and psychosocial outcomes. In other words it is associated with generally worse health and wellbeing, which increases the risk of all negative outcomes. This finding is consistent with other existing reviews. Both previous umbrella reviews that compared effect size data across a range of outcomes also found small to moderate effect sizes for all outcomes and concluded that childhood sexual abuse was a nonspecific risk factor for later psychopathology. In one umbrella review, the authors found a range of small to moderate effect sizes for 25 childhood sexual abuse-related mental health

symptoms and disorders across seven meta-analyses and concluded that childhood sexual abuse was a nonspecific risk factor for a variety of psychopathological outcomes in adulthood (Hillberg et al., 2011). Another umbrella review of 14 meta-analyses also found that childhood sexual abuse was a nonspecific risk factor for a wide range of psychological outcomes, including negative physical and behavioural outcomes that may derive from psychological causes, such as revictimisation, anger, and non-epileptic seizures. However, the review found less robust evidence for the effect of childhood sexual abuse on physical outcomes such as sexually transmitted diseases (Maniglio, 2009).

Another possible explanation for the relative lack of variability in odds ratios across outcomes is the presence of confounding variables that have not been accounted for. For instance, there could be other environmental risk factors such as low socioeconomic status or unstable family environments that both increase the risk of childhood sexual abuse and have a small to moderate effect on a wide range of other negative health and psychosocial outcomes. In this case, the other underlying risk factor could lead to the appearance of childhood sexual abuse as a general risk factor.

4.2 Quality Analysis Findings

If there is evidence for small study effects or excess statistical significance in a meta-analysis, it may be due to publication bias, suggesting that the reported and pooled effect sizes do not accurately reflect the full range of findings, potentially skewing or misrepresenting the overall pooled effect size (Egger et al., 1997; Ioannidis & Trikalinos, 2007). High levels of heterogeneity suggest that results that have been combined in a

meta-analysis are inconsistent and may not have been appropriately pooled (J. P. Higgins et al., 2003). A prediction interval including the null hypothesis implies that the findings of a current meta-analysis cannot be confidently anticipated to be replicable in future studies, diminishing the certainty of clinical recommendations (Inthout et al., 2016). Low AMSTAR scores suggest that meta-analytic methodologies may have been limited by factors such as conflict of interest or incomplete literature searches (Shea et al., 2007). Meta-analyses that perform poorly across all or many of these quality analyses are more likely to report inaccurate, incomplete, or inappropriately synthesized findings, and any clinical implications drawn from such results should be done cautiously and judiciously.

There was a wide range of data and methodological quality in the studies included in this umbrella review. In the overall quality analysis, in which outcomes were scored on a scale of 0 to 5, there were multiple outcomes that achieved each score. However, significantly more outcomes received low scores than high scores. Eight outcomes received the lowest score of 0, and 20 out of 28 outcomes received scores from 0-2, while 8 outcomes received scores from 3-5 and only 2 received the highest score of 5. This suggests that, on the whole, the quality of meta-analyses included in this umbrella review were low to moderate.

Given the overall low quality of the study data and methodology included in this umbrella review, most findings should be interpreted cautiously. However, the reported effect sizes for outcomes with higher overall quality scores can be considered more robust and reliable. Of the four outcomes receiving the highest quality scores (schizophrenia, eating disorders, posttraumatic stress disorder, and sex offending versus non-sex offending) three of them had both 95% confidence intervals and 95% prediction

intervals that excluded the null, suggesting that a true population effect of childhood sexual abuse on these outcomes can be asserted with greater certainty. These three outcomes were eating disorders, posttraumatic stress disorder, and sex offending versus non-sex offending. The fourth outcome, schizophrenia, while receiving a high quality analysis score, also had quite large 95% confidence intervals and 95% prediction intervals, both of which included the null. Therefore, the evidence for the effect of childhood sexual abuse on schizophrenia is not significant.

4.3 Potential Outcome Mechanisms

Given the overall robustness of the effect size findings and quality analysis for the three outcomes of eating disorders, posttraumatic stress disorder, and sex offending versus non-sex offending, the following discussion of potential mechanisms linking childhood sexual abuse and later outcomes will focus primarily on them.

The meta-analysis on sex offending versus non-sex offending that was included in this umbrella review discusses the idea of a victim to perpetrator cycle as an explanation for the effect of childhood sexual abuse on sexual offending. This theory holds that children who experience sexual abuse are at increased risk of later perpetrating the same kind of abuse in adulthood and is supported by the article's finding that sex offenders only differed from non-sex offenders in their experiences of childhood sexual abuse, not in any other kinds of abuse (Jespersen et al., 2009). The authors also discuss possible mechanisms that might explain the victim to perpetrator cycle. One explanation that they explore, based on learning theory, suggests that the cycle is perpetuated through imitation, conditioning, and reinforcement. A different explanation suggested in the

article asserts that the experience of childhood sexual abuse impacts a young person's subsequent psychosexual development. For instance, childhood sexual abuse could lead to increased paedophilic sexual attraction.

The authors also note a few possible third variables that could explain both the experience of childhood sexual abuse and the perpetration of childhood sexual abuse. One such variable is a genetic link. If a family member sexually abuses a child, it is possible that both the child and the family member share a genetic predisposition towards childhood sexual abuse. Another possible third variable is a predisposition towards psychopathology. It is possible that those who are prone towards psychopathology are also more likely to both experience and perpetrate childhood sexual abuse. Finally, another possible third variable explanation is that childhood sexual abuse results in psychopathology, and psychopathology then results in perpetration. In this case, psychopathology would be a mediating variable between childhood sexual abuse victimisation and perpetration (Jespersen et al., 2009).

Other researchers have implicated different variables in the victim to perpetrator cycle as well, including one article that suggests that the type of relationship between the victim and the abuser may be critical. For instance, if the perpetrator is a family member or another close adult, the victim may come to associate sexual violence with nurturing relationships. The same study lists as possible confounding or mediating variables the higher rates of neglectful parents, criminality, substance abuse, and psychiatric history, among perpetrators who were assaulted as children, compared to perpetrators who were not assaulted as children (Seghorn, Prentky, & Boucher, 1987).

Another review article on the victim to perpetrator cycle offers a cognitive behaviour explanation, suggesting that the cycle is moderated by post-abuse masturbation, fantasy, and pleasure (Thomas & Fremouw, 2009). This article also discusses other broad categories of moderating variables to consider, including victim characteristics, perpetrator characteristics, abuse characteristic, post-abuse factors, and family factors. The reviewers found that, along with the post-abuse cognitive behavioural factors listed above, the moderating factors with the most support in the literature are familial factors, such as parental loss or support, and multiple kinds of abuse. The authors also found that the victim to perpetrator cycle is reported much more frequently in men than in women.

A different study supports this gender difference as well, finding that, among men at a forensic psychotherapy centre, sexual offending was positively correlated with a history of childhood sexual abuse, while the same was not true for women. They also found that men who had been sexually abused as children by women were more likely than those abused by men to become perpetrators later (Glasser et al., 2001). However, a recent primary study on pathways from childhood sexual abuse to later victimisation and perpetration complicates this account of gender differences, finding that childhood sexual abuse increased the risk of perpetrating sexual aggression not only in men but also in women. In men, this pathway was mediated by sexual self-esteem, and in women, it was mediated by risky sexual behaviour. The study also found that childhood sexual abuse indirectly increased victimisation in both women and men, with risky sexual behaviour as a mediating factor in men and sexual self-esteem as a mediating factor in women. The findings of this study suggest that gender differences in the pathways from sexual

victimhood to offending may be more complicated than previously understood (Krahé & Berger, 2017).

The meta-analysis on posttraumatic stress disorder that was included in this umbrella review also discusses possible mechanisms for the effect of childhood sexual abuse on this outcome. The authors suggest the possibility of a gene-environment interaction, to account for the association between childhood sexual abuse and various psychiatric outcomes. In particular, they note that certain genotypes of the 5HTTLPR serotonin transporter gene have been implicated as a link between childhood maltreatment and depression, and there is some evidence that these same 5HTTLPR polymorphisms influence the development of posttraumatic stress disorder in trauma survivors. They also note that the FK506 binding protein, FKBP5, has been associated with adult posttraumatic stress symptoms in patients with a history of child abuse (L. P. Chen et al., 2010). However, the types of studies that produced these findings, candidate gene studies that target the relationship between a specific gene as a hypothesized risk factor, have been found to have poor replicability and, as such, are considered obsolete (Ioannidis, Ntzani, Trikalinos, & Contopoulos-Ioannidis, 2001). The current preferred methodology for identifying genetic risk factors involves genome-wide association studies, which look for genetic associations to specific outcomes across the whole genome, without an *a priori* hypothesis.

Another possible underlying mechanism in the relationship between childhood sexual abuse and later posttraumatic stress disorder is the Hypothalamic-Pituitary-Adrenal (HPA) axis. One review on the relationship between childhood maltreatment and the HPA axis found strong evidence for HPA axis dysregulation as a link between child

maltreatment and later major depressive disorder. The same review found some evidence for HPA axis dysregulation as a factor in posttraumatic stress disorder, though the evidence is less convincing, and the findings suggest that the outcome may be influenced by other factors such as the timing and type of trauma (Shea et al., 2007). A different review suggests that the direct link between childhood trauma and adult posttraumatic stress disorder is tenuous and that, instead, childhood abuse may be a risk factor for adult traumatization which subsequently leads to the development of posttraumatic stress disorder, perhaps in combination with HPA axis sensitization (Pratchett & Yehuda, 2011).

Finally, the HPA axis has also been implicated as a potential mechanism in the relationship between childhood sexual abuse and adulthood eating disorders. Patients with both anorexia nervosa and bulimia nervosa who have experienced childhood maltreatment have been found to have lower cortisol awakening responses than eating disorder patients who have not experienced childhood maltreatment. Therefore, dysregulation of the HPA axis stress response system may be a common outcome of childhood sexual abuse as well as a risk factor for a range of adulthood mental health problems, including depression, posttraumatic stress disorder, and eating disorders (Monteleone et al., 2015).

4.4 Analysis of Gender Differences

In the gender subgroup analysis conducted for this umbrella review, small sample sizes and large 95% confidence intervals limited the extent to which strong conclusions could be drawn for the majority of outcomes. There were larger effect sizes for women

than men for all of the internalizing outcomes in the gender analysis (anxiety, depression, eating disorders, and schizophrenia), and the effect size of childhood sexual abuse on the only externalizing outcome (suicidality) was larger in men than in women. However, there was overlap in the 95% confidence intervals for the male and female scores for eating disorders, schizophrenia, and suicidality, making these gender differences statistically insignificant. Also, in particular the male effect sizes were based on small numbers of primary studies,

However, two outcomes, depression and anxiety, had no overlap in the 95% confidence intervals between men and women, and they both were based on relatively large numbers of primary studies. Therefore, for at least two psychiatric diagnoses, there is evidence for a small but significant gender difference in the effect of childhood sexual abuse. While the article included in this umbrella review that provided the meta-analyses on both depression and anxiety does not discuss possible explanations for this finding (Amado et al., 2015), gender differences in effect size, and specifically the greater effect of childhood sexual abuse on women than men, is supported by a number of existing studies. For instance, the controversial 1998 article, condemned by the United States Congress, that found no evidence for intense and pervasive harm caused by childhood sexual abuse also found that female participants were more likely than male participants to experience psychological harm (Rind et al., 1998). While no articles addressing gender differences in anxiety and depression outcomes for childhood sexual abuse were identified in the current review, other studies have found similar trends for posttraumatic stress disorder. One review found that the prevalence of posttraumatic stress disorder in adults who experienced childhood sexual abuse was significantly higher in women than

in men (OR=2.6; 95% CI=1.3-5.3). Although, this article did not find the same significant gender difference for posttraumatic stress disorder in children or adolescents who had experienced childhood sexual abuse (Tolin & Foa, 2006). Similarly, a prospective study of sex differences in posttraumatic stress disorder among adults who experienced childhood abuse or neglect found that, across all types of childhood abuse and neglect, adult women were twice as likely to develop posttraumatic stress disorder as men were. The authors also found that a significant proportion of this gender difference (39%) was explained by greater revictimisation in adulthood (Koenen & Widom, 2009). This suggests that one potential explanation for gender differences in adult psychopathology among childhood sexual abuse victims is the greater tendency for women to experience other forms of trauma or victimisation later on.

Another possible explanation for the greater risk of adult psychopathology in women than men who have experienced childhood sexual abuse involves HPA axis dysregulation. As discussed earlier, HPA axis dysregulation has been associated with increased risk of disorders such as depression, posttraumatic stress disorder, and eating disorders. In particular, depressed adults who were abused as children have been found to have a blunted cortisol response (Heim, Newport, Bonsall, Miller, & Nemeroff, 2001), and a blunted cortisol response has been found to increase the risk of developing posttraumatic stress disorder (Miller, Chen, & Zhou, 2007). Differences in HPA axis functioning have also been found between genders for adults who experienced childhood maltreatment, with women showing a more substantial cortisol blunting response (Doom, Cicchetti, Rogosch, & Dackis, 2013). Therefore, this greater cortisol blunting response in

abused women could at least partially explain the mechanisms for greater risk of later psychopathology among sexually abused girls than boys.

4.5 Mediating, Moderating, and Confounding Factors

Across the literature, a range of other factors that may affect child sexual abuse outcomes have been suggested. One 1998 meta-analysis claimed that family environment accounted for a much greater proportion of harm than childhood sexual abuse did (Rind et al., 1998), and more recent studies have also found a moderating effect of the characteristics of the abuse itself, such as age at time of abuse, form of abuse, use of violence or force during abuse, and duration of abuse (Amado et al., 2015; Arriola et al., 2005; Tyler, 2002). This research suggests that all instances of childhood sexual abuse are not same. Many other conditions in a child's environment, including pre-abuse, during abuse, and post-abuse factors, can influence the nature and the severity of the outcome.

While the majority of the meta-analyses included in this umbrella review adjusted or conducted subgroup analyses for potential mediating, moderating, and confounding variables, some focused primarily on methodological variables such as sample source or study design, rather than factors of the environment or characteristics of the abuse (Hauser et al., 2011; Irish et al., 2010; Jespersen et al., 2009; Klonsky & Moyer, 2008; Varese et al., 2012), and, critically, there was not enough overlap in the analysis of environmental variables across meta-analyses to enable statistical comparison. Therefore, at the umbrella review level, little could be concluded about the nature and intensity of their effects. These variables that are not accounted for limit the extent to which findings

about childhood sexual abuse outcomes can be generalized in clinical populations. It is important, then, that future primary studies and meta-analyses on childhood sexual abuse control, adjust, or conduct separate analyses for as many of these potentially relevant variables as possible, to allow for more nuanced analysis of childhood sexual abuse outcomes in different contexts and environments. These variables could include familial factors such as dysfunctional families, parental criminality, parental substance abuse, single or divorced parents, or parental absence. They could also include social factors, such as socioeconomic status. Finally, characteristics of the abuse itself such as duration, use of violence or force, intra-familial versus extra-familial abuse, and age of onset should be considered.

A further challenge will be determining what relationship these other factors have to the exposure and the outcome, because in many cases it is unclear whether a variable is mediating, moderating, confounding, or a combination of the above. For instance, consider alcohol abuse as a variable and the ways that it could potentially interact with childhood sexual abuse and an outcome such as depression. Alcohol abuse could be a mediating factor, if childhood sexual abuse increases the risk of alcohol abuse and alcohol abuse increases the risk of depression. It could be a moderating factor, if those who experience childhood sexual abuse and also develop an alcohol abuse problem tend to have more severe depression outcomes than those who don't abuse alcohol. It could also be a confounding factor, if adolescents who abuse alcohol are more likely to be sexually abused and are also more likely to develop depression later on. More than one of these pathways could be occurring simultaneously, as well. While the roles of different kinds of variables may be difficult to tease apart, consistent reporting and analysis of

these variables will facilitate the process, along with carefully designed prospective studies and techniques such as Mendelian randomisation.

More consistent adjusting for potential confounding variables would also allow for improved accuracy in calculations of population attributable fractions. In the current umbrella review, the calculated attributable fractions must be regarded as tentative estimates because of the lack of adjusted effect size data provided in component meta-analyses and the resulting use of an attributable risk formula that assumes a lack of confounding variables. Given the high likelihood that there are some confounding variables in the effects of childhood sexual abuse on subsequent health and psychosocial outcomes, more accurate population attributable risk estimates will only be achieved with more extensive and consistent adjustment for confounders across primary studies and meta-analyses. Which variables have been adjusted for should also be explicitly stated in each study, to facilitate comparison and pooling.

4.6 Gaps in the Literature

There are currently gaps in the childhood sexual abuse outcome literature, regarding outcomes of interest. For instance, posttraumatic stress disorder, while intuitively a relatively likely psychiatric outcome in adults who have experienced childhood sexual abuse, was only addressed in one identified meta-analysis, and that meta-analysis included only three primary studies (L. P. Chen et al., 2010). There was also one other meta-analysis that analysed traumatic stress responses more broadly and another that reported on gender differences in posttraumatic stress disorder prevalence (Neumann et al., 1996; Tolin & Foa, 2006). In comparison, six identified studies included meta-

analyses of anxiety as an outcome of childhood sexual abuse (Amado et al., 2015; L. P. Chen et al., 2010; Li et al., 2016; Lindert et al., 2014; Neumann et al., 1996; Ulrich et al., 2005), and eight identified studies included meta-analyses of depression as an outcome of childhood sexual abuse (Amado et al., 2015; L. P. Chen et al., 2010; Jumper, 1995; Li et al., 2016; Lindert et al., 2014; Mandelli et al., 2015; Neumann et al., 1996; Ulrich et al., 2005).

Other common psychiatric diagnoses were notably underrepresented in the meta-analytic literature on childhood sexual abuse outcomes as well. For instance, while there were two meta-analyses that addressed obsessive-compulsive symptomatology, there was no relevant study on obsessive-compulsive disorder specifically (Neumann et al., 1996; Ulrich et al., 2005). Similarly, there was a study on the effects of childhood sexual abuse on a range of bipolar disorder outcomes, such as early onset of bipolar disorder, suicidality in individuals with bipolar disorder, and substance abuse in individuals with bipolar disorder (Daruy-Filho, Brietzke, Lafer, & Grassi-Oliveira, 2011), but there was no meta-analysis on the effect of childhood sexual abuse on bipolar disorder in general. Finally, only one childhood sexual abuse meta-analysis was identified each for schizophrenia and sleep disorders (L. P. Chen et al., 2010). The schizophrenia meta-analysis only included three primary studies and the sleep disorders meta-analysis only included one, limiting the power and generalizability of the findings from these studies. Future childhood sexual abuse research and clinical practice would benefit from addressing these gaps in the literature with more primary studies and meta-analyses on posttraumatic stress disorder, obsessive-compulsive disorder, bipolar disorder, schizophrenia, and sleep disorders.

Among psychosocial outcomes, there is a dearth of meta-analytic work on violent and sexual offending in adults who experienced sexual abuse as children. In this umbrella review, the included meta-analysis on adulthood sexual abuse perpetration focused narrowly on a comparison between sexual offenders versus offenders whose crimes were not sexual in nature (Jespersen et al., 2009). While this study produced interesting and significant results, more meta-analytic work should also be done to compare the effects of childhood sexual abuse on sexual offenders and violent offenders versus those who do not go on to commit any kind of crime. Homelessness is another outcome that is important from a public health perspective but has not been extensively covered in the literature so far. One meta-analysis considered for inclusion in this umbrella review studied homelessness as an outcome of childhood sexual abuse but provided only prevalence data, not effect size data (Sundin & Baguley, 2015). Because of the lack of effect size data, the homelessness outcome could not be pooled or compared with other outcomes and was therefore excluded from the umbrella review. No other relevant meta-analyses on childhood sexual abuse and adulthood homelessness were identified. More primary studies and meta-analyses should be conducted in future on public health and policy relevant outcomes such as homelessness and criminal offending.

4.7 Umbrella Review Limitations

Along with the limitations posed by the included material, such as low quality evidence and methodology, lack of accounting for mediating, moderating, and confounding factors, and gaps in the literature (see above), the present study also faces certain limitations in virtue of being an umbrella review. The first such limitation

involves its exclusive reliance upon meta-analyses. Because only meta-analytic data was included in this study, informative and high-quality primary study data that has not been incorporated into a meta-analysis may have been overlooked. Thus, this umbrella review is limited to commenting on the nature and quality of the existing body of literature on childhood sexual abuse outcomes only insofar as that literature has already been meta-analysed.

A further limitation of umbrella reviews in general, because they systematically include all relevant meta-analyses that meet inclusion criteria, is that they may not be very discriminating in terms of the quality of studies they include. I have attempted to mitigate this limitation through extensive quality assessment, both evaluating the quality of the individual included meta-analyses with the AMSTAR checklist and assessing the quality of the primary data, through the various statistical analyses discussed above.

4.8 Research Implications

The above identified limitations and gaps in the current literature on long-term outcomes of childhood sexual abuse provide valuable direction for future research in this field. Firstly, more consistent subgroup analysis of mediating and moderating variables and adjusting for confounding variables will increase the accuracy, generalizability, and potential for comparison across future studies. Secondly, accuracy and generalizability will also be improved by maintaining high methodological quality in primary studies and in meta-analyses, following guidelines such as PRISMA and Cochrane. Finally, more primary studies and meta-analyses should be conducted for outcomes, such as those listed

above, that currently have either no meta-analyses or meta-analyses including only small numbers of primary studies.

4.9 Clinical Implications

Given the evidence in this umbrella review for the long-term harm caused by childhood sexual abuse, preventing and mitigating that harm should be a clinical priority. There are multiple potential points of clinical intervention for childhood sexual abuse, namely primary prevention, secondary prevention, and tertiary prevention. The first point of intervention, primary prevention, aims to stop the abuse from happening in the first place. One of the most prevalent childhood sexual abuse interventions is primary prevention in schools. School based interventions, generally personal safety programs for children, focus on increasing children's knowledge about sexual abuse and improving their ability to avoid risky situations and recognize and report abuse to adults when it occurs (Lalor & McElvaney, 2010; Smallbone, Marshall, & Wortley, 2008). While these programs have been shown to be effective at increasing children's knowledge and skills, there does not seem to be evidence that they decrease childhood sexual abuse among participants, and they have received criticism for placing the burden of preventing abuse on children (Smallbone et al., 2008). Other potential targets of primary prevention include community-based programs, such as media campaigns, parental education programs, and safeguards at the institutional level. Institutional safeguards can include screening and training staff at organizations that work with children and ensuring that staff are not left alone with children, to limit the kinds of situations that facilitate abuse (Smallbone et al., 2008). Monitoring online content can also serve as form of primary

prevention, because the Internet is a frequent platform for cyber stalking, arranging inappropriate meet-ups between children and adults, and promoting sex tourism, all of which can lead to subsequent abuse (Smallbone et al., 2008).

Secondary prevention focuses on minimising harm, immediately after childhood sexual abuse has occurred. This can involve decreasing the likelihood of revictimisation among abused children and minimising the immediate psychological harm caused by abuse. Secondary prevention that aims to minimise the immediate risk of revictimisation is not straightforward but can focus on characteristics of the victim and environment that may have facilitated the initial instance of abuse. For instance, there is some evidence that children with low self-esteem and children in dysfunctional families are more likely to be victimised, so secondary preventative measures could focus on improving the self-esteem or the family environment of victimised children (Smallbone et al., 2008). Decreasing the risk of revictimisation also often involves reporting cases of abuse and removing perpetrators from a household or other environment where they are likely to reoffend or removing a child from an environment where they are likely to be harmed again. Another aim of secondary prevention may be mitigating the immediate psychiatric symptoms and other negative outcomes in children, after abuse. Psychological treatment for children who have been sexually abused has been found to be effective, particularly among young children, and the strongest evidence in randomised controlled trials has been for cognitive behavioural therapy (Lalor & McElvaney, 2010; Ramchandani & Jones, 2003). Finally, multiple sources have documented that an important factor in resilience among children who have experienced sexual abuse is social support, particularly from parents, but also from other members of their household or community

(Lalor & McElvaney, 2010; Smallbone et al., 2008). Therefore, secondary prevention could also involve educating parents of abused children and other adults on behaviours that are supportive for victims.

The aim of tertiary prevention is to manage the long-term consequences of childhood sexual abuse. This often consists of treatment and rehabilitation for sexual offenders and therapy with adults who experienced abuse in childhood. Studies have shown that therapeutic treatment programs with offenders can effectively reduce recidivism, particularly when the programs are tailored to the individual participant, maximise therapeutic engagement, and target factors that are directly relevant to recidivism. Group based cognitive behavioural therapy has also been found to be effective (Lalor & McElvaney, 2010; Smallbone et al., 2008). Tertiary prevention can also focus on the long-term psychological needs of adults who experienced childhood sexual abuse. One study of a Danish cohort found that one year of therapy for adults with a history of childhood sexual abuse improved outcomes related to general distress, sense of self-worth, and posttraumatic stress disorder symptomatology (Elklit, 2015). Another study found moderate beneficial effects of psychotherapy for adults, with the effectiveness moderated by characteristics of the therapy (Taylor & Harvey, 2010). Finally, a different study found that posttraumatic stress disorder treatment for adults who had experienced childhood sexual abuse was effective, with individual treatments having higher effect sizes than group treatments (Ehring et al., 2014).

Another finding of this same study was a higher effect size of trauma-focused treatments than other forms of treatment for posttraumatic stress disorder in adults who had experienced childhood sexual abuse (Ehring et al., 2014). Support for the

effectiveness of trauma-focused treatment has been noted elsewhere, as well. One study of trauma-focused cognitive behavioural therapy for children in Zambia found that the effectiveness of the therapy was moderated by a history of sexual abuse. Those children who had experienced sexual abuse showed greater reductions in trauma symptomatology over the course of treatment than those children who had experienced other kinds of trauma (Kane et al., 2016). The effectiveness of trauma-focused cognitive behavioural therapy was also noted in a different study of sexually abused children who had posttraumatic stress disorder (Cohen, Deblinger, Mannarino, & Steer, 2004). While the literature on trauma-focused therapies for childhood sexual abuse is still relatively limited and focused primarily on the treatment of posttraumatic stress disorder, initial findings are promising, and future research should consider whether this kind of therapy for victims of childhood sexual abuse can be adapted to other disorders.

In reviewing the literature, preliminary evidence supports the effectiveness of clinical interventions to prevent long-term outcomes of childhood sexual abuse at the primary, secondary, and tertiary levels, and some patients may benefit specifically from trauma-focused treatments.

4.10 Conclusion

The primary findings of this umbrella review support the prevailing theory that childhood sexual abuse increases the risk of long-term negative health and psychosocial outcomes. In particular, there is good quality evidence for the effect of childhood sexual abuse on adulthood eating disorders, posttraumatic stress disorder, and sex offending versus non-sex offending. The current meta-analytic literature on the topic of childhood

sexual abuse outcomes covers a wide range of outcomes but would benefit from further research on specific outcomes such as bipolar disorder, posttraumatic stress disorder, obsessive-compulsive disorder, offending behaviours, and homelessness. Future research should also follow methodological best practice guidelines such as PRISMA and Cochrane and should account for mediating, moderating, and confounding factors as thoroughly and consistently as possible. This study has implications for the importance of primary, secondary, and tertiary prevention measures for childhood sexual abuse as well as the potential benefit of trauma-focused treatments for childhood sexual abuse victims.

References

- Amado, B. G., Arce, R., & Herraiz, A. (2015). Psychological injury in victims of child sexual abuse: A meta-analytic review. *Psychosocial Intervention*, 24(1), 49-62.
- Arriola, K. R., Loudon, T., Doldren, M. A., & Fortenberry, R. M. (2005). A meta-analysis of the relationship of child sexual abuse to HIV risk behavior among women. *Child Abuse & Neglect*, 29(6), 725-746.
- Babchishin, K. M., Hanson, R., & Hermann, C. A. (2011). The characteristics of online sex offenders: A meta-analysis. *Sexual Abuse: Journal of Research and Treatment*, 23(1), 92-123.
- Benhold, K. (2016). Child Sexual Abuse Scandal Rocks U.K. Soccer. *The New York Times*. Retrieved from https://www.nytimes.com/2016/12/13/world/europe/soccer-uk-sexual-abuse-andy-woodward.html?_r=2
- Blow, C. M. (2015). Surviving Child Sexual Abuse. *The New York Times*. Retrieved from <https://www.nytimes.com/2015/06/01/opinion/charles-blow-surviving-child-sexual-abuse.html>
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to Meta-Analysis*: John Wiley & Sons.
- Browne, A., & Finkelhor, D. (1986). Impact of Child Sexual Abuse. A Review of the Research. *Psychological Bulletin*, 99(1), 66-77. doi: 10.1037/0033-2909.99.1.66
- Chen, H., Cohen, P., & Chen, S. (2010). How Big is a Big Odds Ratio? Interpreting the Magnitudes of Odds Ratios in Epidemiological Studies. *Communications in Statistics - Simulation and Computation*, 39(4), 860-864. doi: 10.1080/03610911003650383
- Chen, L. P., Murad, M. H., Paras, M. L., Colbenson, K. M., Sattler, A. L., Goranson, E. N., . . . Zirakzadeh, A. (2010). Sexual Abuse and Lifetime Diagnosis of Psychiatric Disorders: Systematic Review and Meta-analysis. *Mayo Clinic Proceedings*, 85(7), 618-629. doi: 10.4065/mcp.2009.0583
- Civic Impulse. (2017). *H.Con.Res 107 - 106th Congress: Expressing the sense of Congress rejecting the conclusions of a recent article published by the...* Retrieved from <https://www.govtrack.us/congress/bills/106/hconres107>.
- Cohen, J. A., Deblinger, E., Mannarino, A. P., & Steer, R. A. (2004). A Multisite, Randomized Controlled Trial for Children with Sexual Abuse-Related PTSD Symptoms. *Journal of the American Academy of Child and Adolescent Psychiatry*, 43(4), 393-402. doi: 10.1097/00004583-200404000-00005
- Dagang, S. A. (1997). A meta-analytic review of studies examining the effects of childhood sexual abuse in women. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 57(7-B), 4701.
- Danese, A., & Tan, M. (2014). Childhood maltreatment and obesity: Systematic review and meta-analysis. *Molecular Psychiatry*, 19(5), 544-554.
- Daruy-Filho, L., Brietzke, E., Lafer, B., & Grassi-Oliveira, R. (2011). Childhood maltreatment and clinical outcomes of bipolar disorder. *Acta Psychiatrica Scandinavica*, 124(6), 427-434.

- Devries, K. M., Mak, J. Y., Child, J. C., Falder, G., Bacchus, L. J., Astbury, J., & Watts, C. H. (2014). Childhood sexual abuse and suicidal behavior: A meta-analysis. *Pediatrics*, 133(5), e1331-e1344.
- Doom, J. R., Cicchetti, D., Rogosch, F. A., & Dackis, M. N. (2013). Child maltreatment and gender interactions as predictors of differential neuroendocrine profiles. *Psychoneuroendocrinology*, 38(8), 1442-1454.
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629-634.
- Ehring, T., Welboren, R., Morina, N., Wicherts, J. M., Freitag, J., & Emmelkamp, P. M. G. (2014). Meta-analysis of psychological treatments for posttraumatic stress disorder in adult survivors of childhood abuse. *Clinical Psychology Review*, 34(8), 645-657. doi: 10.1016/j.cpr.2014.10.004
- Elklit, A. (2015). Treatment of Danish Survivors of Child Sexual Abuse-A Cohort Study. *Behavioral sciences (Basel, Switzerland)*, 5(4), 589. doi: 10.3390/bs5040589
- Fergusson, D. M., Boden, J. M., & Horwood, L. J. (2008). Exposure to childhood sexual and physical abuse and adjustment in early adulthood. *Child Abuse & Neglect*, 32(6), 607-619.
- Finkelhor, D. (1979). What's Wrong with Sex between Adults and Children? Ethics and the Problem of Sexual Abuse. *American Journal of Orthopsychiatry*, 49(4), 692-697. doi: 10.1111/j.1939-0025.1979.tb02654.x
- Fossati, A., Madeddu, F., & Maffei, C. (1999). Borderline Personality Disorder and childhood sexual abuse: A meta-analytic study. *Journal of Personality Disorders*, 13(3), 268-280.
- Glasser, M., Kolvin, I., Campbell, D., Glasser, A., Leitch, I., & Farrelly, S. (2001). Cycle of child sexual abuse: Links between being a victim and becoming a perpetrator. *British Journal of Psychiatry*, 179(6), 482-494. doi: 10.1192/bjp.179.6.482
- Greco, T., Zangrillo, A., Biondi-Zoccai, G., & Landoni, G. (2013). Meta-analysis: pitfalls and hints. *Heart, lung and vessels*, 5(4), 219.
- Hauser, W., Kosseva, M., Uceyler, N., Klose, P., & Sommer, C. (2011). Emotional, physical, and sexual abuse in fibromyalgia syndrome: a systematic review with meta-analysis. *Arthritis Care Res (Hoboken)*, 63(6), 808-820. doi: 10.1002/acr.20328
- Heim, C., Newport, D. J., Bonsall, R., Miller, A. H., & Nemeroff, C. B. (2001). Altered pituitary-adrenal axis responses to provocative challenge tests in adult survivors of childhood abuse. *Am J Psychiatry*, 158(4), 575-581. doi: 10.1176/appi.ajp.158.4.575
- Hemmingsson, E., Johansson, K., & Reynisdottir, S. (2014). Effects of childhood abuse on adult obesity: a systematic review and meta - analysis (Vol. 15, pp. 882-893).
- Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses.

- Higgins, J. P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions*. Retrieved from www.handbook.cochrane.org
- Hillberg, T., Hamilton-Giachritsis, C., & Dixon, L. (2011). Review of meta-analyses on the association between child sexual abuse and adult mental health difficulties: A systematic approach. *Trauma, Violence, & Abuse*, 12(1), 38-49.
- Inthout, J., Ioannidis, J., Rovers, M. M., & Goeman, J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*, 6(7). doi: 10.1136/bmjopen-2015-010247
- Ioannidis, J. P. A. (2009). Integration of evidence from multiple meta-analyses: a primer on umbrella reviews, treatment networks and multiple treatments meta-analyses. *CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne*, 181(8), 488. doi: 10.1503/cmaj.081086
- Ioannidis, J. P. A., Ntzani, E. E., Trikalinos, T. A., & Contopoulos-Ioannidis, D. G. (2001). Replication validity of genetic association studies. *Nat Genet*, 29(3), 306-309.
- Ioannidis, J. P. A., Patsopoulos, N. A., & Evangelou, E. (2007). Uncertainty in heterogeneity estimates in meta-analyses. *BMJ*, 335(7626), 914.
- Ioannidis, J. P. A., & Trikalinos, T. A. (2007). An exploratory test for an excess of significant findings. *Clinical trials (London, England)*, 4(3), 245.
- Irish, L., Kobayashi, I., & Delahanty, D. L. (2010). Long-term physical health consequences of childhood sexual abuse: A meta-analytic review. *Journal of Pediatric Psychology*, 35(5), 450-461.
- Jespersen, A. F., Lalumiere, M. L., & Seto, M. C. (2009). Sexual abuse history among adult sex offenders and non-sex offenders: A meta-analysis. *Child Abuse & Neglect*, 33(3), 179-192.
- Jumper, S. A. (1995). A meta-analysis of the relationship of child sexual abuse to adult psychological adjustment. *Child Abuse & Neglect*, 19(6), 715-728.
- Kane, J. C., Murray, L. K., Cohen, J., Dorsey, S., Skavenski Van Wyk, S., Galloway Henderson, J., . . . Bolton, P. (2016). Moderators of treatment response to trauma - focused cognitive behavioral therapy among youth in Zambia. *Journal of Child Psychology and Psychiatry*, 57(10), 1194-1202. doi: 10.1111/jcpp.12623
- Klonsky, E., & Moyer, A. (2008). Childhood sexual abuse and non-suicidal self-injury: Meta analysis. *The British Journal of Psychiatry*, 192(3), 166-170.
- Koenen, K. C., & Widom, C. S. (2009). A prospective study of sex differences in the lifetime risk of posttraumatic stress disorder among abused and neglected children grown up. *Journal of Traumatic Stress*, 22(6), 566-574. doi: 10.1002/jts.20478
- Krahé, B., & Berger, A. (2017). Gendered pathways from child sexual abuse to sexual aggression victimization and perpetration in adolescence and young adulthood. *Child Abuse & Neglect*, 63, 261-272. doi: 10.1016/j.chiabu.2016.10.004
- Lalor, K., & McElvaney, R. (2010). Child Sexual Abuse, Links to Later Sexual Exploitation/High-Risk Sexual Behavior, and Prevention/Treatment Programs (Vol. 11, pp. 159-177).

- Langeland, W., Smit, J. H., Merckelbach, H., de Vries, G., Hoogendoorn, A. W., & Draijer, N. (2015). Inconsistent retrospective self-reports of childhood sexual abuse and their correlates in the general population. *Soc Psychiatry Psychiatr Epidemiol*, 50(4), 603-612. doi: 10.1007/s00127-014-0986-x
- Latthe, P., Mignini, L., Gray, R., Hills, R., & Khan, K. (2006). Factors predisposing women to chronic pelvic pain: systematic review. *BMJ*, 332(7544), 749. doi: 10.1136/bmj.38748.697465.55
- Li, M., D'Arcy, C., & Meng, X. (2016). Maltreatment in childhood substantially increases the risk of adult depression and anxiety in prospective cohort studies: Systematic review, meta-analysis, and proportional attributable fractions. *Psychological Medicine*, 46(4), 717-730.
- Lindert, J., von Ehrenstein, O. S., Grashow, R., Gal, G., Braehler, E., & Weisskopf, M. G. (2014). Sexual and physical abuse in childhood is associated with depression and anxiety over the life course: Systematic review and meta-analysis. *International Journal of Public Health*, 59(2), 359-372.
- Lloyd, S., & Operario, D. (2012). HIV risk among men who have sex with men who have experienced childhood sexual abuse: Systematic review and meta-analysis. *AIDS Education and Prevention*, 24(3), 228-241.
- Mandelli, L., Petrelli, C., & Serretti, A. (2015). The role of specific early trauma in adult depression: A meta-analysis of published literature. Childhood trauma and adult depression. *European Psychiatry*, 30(6), 665-680.
- Maniglio, R. (2009). The impact of child sexual abuse on health: A systematic review of reviews. *Clinical Psychology Review*, 29(7), 647-657.
- Maniglio, R. (2010). Child sexual abuse in the etiology of depression: A systematic review of reviews. *Depression and Anxiety*, 27(7), 631-642.
- Maniglio, R. (2011a). The role of child sexual abuse in the etiology of substance-related disorders. *Journal of Addictive Diseases*, 30(3), 216-228.
- Maniglio, R. (2011c). The role of child sexual abuse in the etiology of suicide and non-suicidal self-injury *Acta Psychiatr. Scand.* (Vol. 124, pp. 30-41).
- Maniglio, R. (2013). Child sexual abuse in the etiology of anxiety disorders: A systematic review of reviews. *Trauma, Violence, & Abuse*, 14(2), 96-112.
- Miller, G. E., Chen, E., & Zhou, E. S. (2007). If It Goes up, Must It Come Down? Chronic Stress and the Hypothalamic-Pituitary Adrenocortical Axis in Humans. *Psychological Bulletin*, 133(1), 25-45. doi: 10.1037/0033-2909.133.1.25
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement.
- Monteleone, A. M., Monteleone, P., Serino, I., Scognamiglio, P., Genio, M., & Maj, M. (2015). Childhood trauma and cortisol awakening response in symptomatic patients with anorexia nervosa and bulimia nervosa. *International Journal of Eating Disorders*, 48(6), 615-621. doi: 10.1002/eat.22375
- National Institute of Mental Health. (2010). *Global DALYs Contributed by Mental and Behavioral Disorders*. Retrieved from <https://www.nimh.nih.gov/health/statistics/global/global-dalys-contributed-by-mental-and-behavioral-disorders.shtml>.

- Neumann, D. A., Houskamp, B. M., Pollock, V. E., & Briere, J. (1996). The long-term sequelae of childhood sexual abuse in women: A meta-analytic review. *Child Maltreatment*, 1(1), 6-16.
- Ondersma, S., Berliner, L., Chaffin, M., & Cordon, I. (2001). Sex with children is abuse: Comment on Rind, Tromovitch, and Bauserman (1998)--Comment/Reply. *Psychological Bulletin*, 127(6), 707-758. doi: 10.1037//0033-2909.127.6.707
- Paquola, C., Bennett, M. R., & Lagopoulos, J. (2016). Understanding heterogeneity in grey matter research of adults with childhood maltreatment-A meta-analysis and review. *Neurosci Biobehav Rev*, 69, 299-312. doi: 10.1016/j.neubiorev.2016.08.011
- Pereda, N., Guilera, G., Forn, M., & Gomez-Benito, J. (2009). The prevalence of child sexual abuse in community and student samples: A meta-analysis. *Clinical Psychology Review*, 29(4), 328-338.
- Pratchett, L. C., & Yehuda, R. (2011). Foundations of posttraumatic stress disorder: Does early life trauma lead to adult posttraumatic stress disorder? *Dev Psychopathol*, 23(2), 477-491. doi: 10.1017/S0954579411000186
- Quinones-Munoz, S. L. (2001). Psychological effects of childhood sexual abuse in adult males. *Dissertation Abstracts International: Section B: The Sciences and Engineering*, 62(4-B), 2074.
- Ramchandani, P., & Jones, D. P. H. (2003). Treating psychological symptoms in sexually abused children: from research findings to service provision. *The British journal of psychiatry : the journal of mental science*, 183, 484.
- Rind, B., & Tromovitch, P. (1997). A meta-analytic review of findings from national samples on psychological correlates of child sexual abuse. *Journal of Sex Research*, 34(3), 237-255.
- Rind, B., Tromovitch, P., & Bauserman, R. (1998). A meta-analytic examination of assumed properties of child sexual abuse using college samples. *Psychological Bulletin*, 124(1), 22-53.
- Rockhill, B., Newman, B., & Weinberg, C. (1998). Use and misuse of population attributable fractions. *American Journal of Public Health*, 88(1), 15.
- Roodman, A. A., & Clum, G. A. (2001). Revictimization rates and method variance: A meta-analysis. *Clinical Psychology Review*, 21(2), 183-204.
- Seghorn, T. K., Prentky, R. A., & Boucher, R. J. (1987). Childhood Sexual Abuse in the Lives of Sexually Aggressive Offenders. *Journal of the American Academy of Child & Adolescent Psychiatry*, 26(2), 262-267. doi: 10.1097/00004583-198703000-00025
- Shamseer, L. M., D. (2015). Planning a systematic review? Think protocols. Retrieved from <http://www.prisma-statement.org/Protocols/WhyProtocols.aspx>
- Sharpe, D., & Faye, C. (2006). Non-epileptic seizures and child sexual abuse: A critical review of the literature. *Clinical Psychology Review*, 26(8), 1020-1040.
- Shea, B. J., Grimshaw, J. M., Wells, G. A., Boers, M., Andersson, N., Hamel, C., . . . Bouter, L. M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Medical Research Methodology*, 7, 10-10. doi: 10.1186/1471-2288-7-10

- Smallbone, S., Marshall, W. L., & Wortley, R. (2008). *Preventing child sexual abuse : evidence, policy and practice*. Portland, OR: Willan Publishing.
- Smolak, L., & Murnen, S. K. (2002). A meta-analytic examination of the relationship between child sexual abuse and eating disorders. *International Journal of Eating Disorders*, 31(2), 136-150.
- Spiegel, D. (2000). The price of abusing children and numbers. *Sexuality & Culture: An Interdisciplinary Quarterly*, 4(2), 63-66.
- Sterne, J. A., & Egger, M. (2001). Funnel plots for detecting bias in meta-analysis: Guidelines on choice of axis. *Journal of Clinical Epidemiology*, 54(10), 1046-1055. doi: 10.1016/S0895-4356(01)00377-8
- 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.
- Summit, R., & Kryso, J. (1978). Sexual Abuse of Children: A Clinical Spectrum. *American Journal of Orthopsychiatry*, 48(2), 237-251. doi: 10.1111/j.1939-0025.1978.tb01312.x
- Sundin, E. C., & Baguley, T. (2015). Prevalence of childhood abuse among people who are homeless in Western countries: A systematic review and meta-analysis. *Social Psychiatry and Psychiatric Epidemiology*, 50(2), 183-194.
- Taylor, J. E., & Harvey, S. T. (2010). A meta-analysis of the effects of psychotherapy with adults sexually abused in childhood. *Clinical Psychology Review*, 30(6), 749-767.
- Thomas, T. A., & Fremouw, W. (2009). Moderating variables of the sexual "victim to offender cycle" in males. *Aggression and Violent Behavior*, 14(5), 382-387. doi: 10.1016/j.avb.2009.06.006
- Thorlund, K., Imberger, G., Johnston, B. C., Walsh, M., Awad, T., Thabane, L., . . . Wetterslev, J. (2012). Evolution of Heterogeneity (I²) Estimates and Their 95% Confidence Intervals in Large Meta-Analyses. *PLoS ONE*, 7(7), e39471. doi: 10.1371/journal.pone.0039471
- Tolin, D., & Foa, E. (2006). Sex differences in trauma and posttraumatic stress disorder: A quantitative review of 25 years of research *Psychol. Bull.* (Vol. 132, pp. 959-992).
- Tyler, K. A. (2002). Social and emotional outcomes of childhood sexual abuse: A review of recent research. *Aggression and Violent Behavior*, 7(6), 567-589. doi: 10.1016/S1359-1789(01)00047-7
- Ulrich, H., Randolph, M., & Acheson, S. (2005). Child sexual abuse: A replication of the meta-analytic examination of child sexual abuse by Rind, Tromovitch, and Bauserman (1998). *The Scientific Review of Mental Health Practice: Objective Investigations of Controversial and Unorthodox Claims in Clinical Psychology, Psychiatry, and Social Work*, 4(2), 37-51.
- University of Minnesota, & U.S. Department of Justice Office on Violence Against Women. (2009). The Facts about Sexual Violence *Violence against women online resources*.
- Varese, F., Smeets, F., Drukker, M., Lieveverse, R., Lataster, T., Viechtbauer, W., . . . Bentall, R. P. (2012). Childhood adversities increase the risk of psychosis: A

- meta-analysis of patient-control, prospective- and cross-sectional cohort studies. *Schizophrenia Bulletin*, 38(4), 661-671.
- von Ins, F. R., Mikton, C., Finkelhor, D., Barth, J., Pereda, N., Devries, K., & Morse, M. M. (2012). *Promoting Research to Prevent Child Maltreatment*. Paper presented at the ISPCAN International Congress on Child Abuse and Neglect.
- Walsh, S. (2016). Speaking Out About Sexual Abuse: The Seed Of Change. Retrieved from http://www.huffingtonpost.co.uk/samantha-walsh/sexual-abuse_b_13516954.html
- Wang, Y., Wu, B., Yang, H., & Song, X. (2015). The effect of childhood abuse on the risk of adult obesity. *Ann Clin Psychiatry*, 27(3), 175-184.
- Whittenburg, J. A., Tice, P. P., Baker, G. L., & Lemmey, D. E. (2001). A critical appraisal of the 1998 meta-analytic review of child sexual abuse outcomes reported by Rind, Tromovitch, and Bauserman. *Journal of Child Sexual Abuse: Research, Treatment, & Program Innovations for Victims, Survivors, & Offenders*, 9(3-4), 135-156.
- Widom, C. S., & Morris, S. (1997). Accuracy of adult recollections of childhood victimization, Part 2: Childhood sexual abuse. *Psychological Assessment*, 9(1), 34-46. doi: 10.1037/1040-3590.9.1.34
- World Health Organization. (2003). *Guidelines for medico-legal care for victims of sexual violence*. Retrieved from http://www.who.int/violence_injury_prevention/publications/violence/med_leg_guidelines/en/.
- World Health Organization. (2014). Global Health Estimates 2014 Summary Tables.
- Zhang, J., & Yu, K. F. (1998). What's the Relative Risk?: A Method of Correcting the Odds Ratio in Cohort Studies of Common Outcomes. *JAMA*, 280(19), 1690-1691. doi: 10.1001/jama.280.19.1690

Appendix 1: AMSTAR Checklist

1. Was an 'a priori' design provided?

The research question and inclusion criteria should be established before the conduct of the review.

Note: Need to refer to a protocol, ethics approval, or pre-determined/a priori published research objectives to score a “yes.”

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

2. Was there duplicate study selection and data extraction?

There should be at least two independent data extractors and a consensus procedure for disagreements should be in place.

Note: 2 people do study selection, 2 people do data extraction, consensus process or one person checks the other's work.

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

3. Was a comprehensive literature search performed?

At least two electronic sources should be searched. The report must include years and databases used (e.g., Central, EMBASE, and MEDLINE). Key words and/or MESH terms

must be stated and where feasible the search strategy should be provided. All searches should be supplemented by consulting current contents, reviews, textbooks, specialized registers, or experts in the particular field of study, and by reviewing the references in the studies found.

Note: If at least 2 sources + one supplementary strategy used, select “yes” (Cochrane register/Central counts as 2 sources; a grey literature search counts as supplementary).

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

4. Was the status of publication (i.e. grey literature) used as an inclusion criterion?

The authors should state that they searched for reports regardless of their publication type. The authors should state whether or not they excluded any reports (from the systematic review), based on their publication status, language etc.

Note: If review indicates that there was a search for “grey literature” or “unpublished literature,” indicate “yes.” SIGLE database, dissertations, conference proceedings, and trial registries are all considered grey for this purpose. If searching a source that contains

both grey and non-grey, must specify that they were searching for grey/unpublished lit.

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

5. Was a list of studies (included and excluded) provided?

A list of included and excluded studies should be provided.

Note: Acceptable if the excluded studies are referenced. If there is an electronic link to the list but the link is dead, select "no."

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

6. Were the characteristics of the included studies provided?

In an aggregated form such as a table, data from the original studies should be provided on the participants, interventions and outcomes. The ranges of characteristics in all the studies analyzed e.g., age, race, sex, relevant socioeconomic data, disease status, duration, severity, or other diseases should be reported.

Note: Acceptable if not in table format as long as they are described as above.

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

7. Was the scientific quality of the included studies assessed and documented?

'A priori' methods of assessment should be provided (e.g., for effectiveness studies if the author(s) chose to include only randomized, double-blind, placebo controlled studies, or allocation concealment as inclusion criteria); for other types of studies alternative items will be relevant.

Note: Can include use of a quality scoring tool or checklist, e.g., Jadad scale, risk of bias, sensitivity analysis, etc., or a description of quality items, with some kind of result for EACH study ("low" or "high" is fine, as long as it is clear which studies scored "low" and which scored "high"; a summary score/range for all studies is not acceptable).

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

8. Was the scientific quality of the included studies used appropriately in formulating conclusions?

The results of the methodological rigor and scientific quality should be considered in the analysis and the conclusions of the review, and explicitly stated in formulating

recommendations.

Note: Might say something such as “the results should be interpreted with caution due to poor quality of included studies.” Cannot score “yes” for this question if scored “no” for question 7.

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

9. Were the methods used to combine the findings of studies appropriate?

For the pooled results, a test should be done to ensure the studies were combinable, to assess their homogeneity (i.e., Chi-squared test for homogeneity, I²). If heterogeneity exists a random effects model should be used and/or the clinical appropriateness of combining should be taken into consideration (i.e., is it sensible to combine?).

Note: Indicate “yes” if they mention or describe heterogeneity, i.e., if they explain that they cannot pool because of heterogeneity/variability between interventions.

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

10. Was the likelihood of publication bias assessed?

An assessment of publication bias should include a combination of graphical aids (e.g., funnel plot, other available tests) and/or statistical tests (e.g., Egger regression test, Hedges-Olken).

Note: If no test values or funnel plot included, score “no”. Score “yes” if mentions that publication bias could not be assessed because there were fewer than 10 included studies.

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

11. Was the conflict of interest included?

Potential sources of support should be clearly acknowledged in both the systematic review and the included studies.

Note: To get a “yes,” must indicate source of funding or support for the systematic review AND for each of the included studies.

- ☐ Yes
- ☐ No
- ☐ Can't answer
- ☐ Not applicable

Shea et al. *BMC Medical Research Methodology* 2007 7:10 doi:10.1186/1471-2288-7-10

Additional notes (in italics) made by Michelle Weir, Julia Worswick, and Carolyn Wayne based on conversations with

Bev Shea and/or Jeremy Grimshaw in June and October 2008 and July and September 2010.

Appendix 2: PRISMA Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	5
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	11
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	16
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	17
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	20-23
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	17-20
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	17-19
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	20
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	24
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	24-25
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	25-27
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	27
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	31-32

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	33-34
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	28-31
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	36-37
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	38-39
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	45-47
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	40-41
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	42
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	48-51
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	42-45
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	56
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	70-71
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	71-72
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	2