THE CAMPBELL COLLABORATION

# Protocol for a Systematic Review: Child Protection Training for Professionals to Improve Reporting of Child Abuse and Neglect

**Ben Mathews, Kerryann Walsh, Sandra Coe, Maureen C Kenny, Dimitrios Vagenas**

Submitted to the Coordinating Group of:

| | |
|---|---|
| ☐ | Crime and Justice |
| ☐ | Education |
| | ☐ Disability |
| ☐ | International Development |
| | ☐ Nutrition |
| X | Social Welfare |
| ☐ | Other: |

Co-registration:

| | | | |
|---|---|---|---|
| ☐ | No | | |
| X | Yes | X Cochrane | ☐ Other |
| ☐ | Maybe | | |

Approval Date: 15 June 2015
Publication Date: 1 July 2015

This protocol is co-registered within both the Cochrane and Campbell Collaborations.
A version of this protocol can also be found in the Cochrane Library.

## 1. BACKGROUND

### *Description of the condition*

#### Child abuse and neglect

Child abuse and neglect includes physical abuse, sexual abuse, psychological or emotional abuse, and neglect. Most child abuse and neglect occurs in private, is inflicted or caused by parents and caregivers, and does not become known to government authorities or helping agencies. Except for sexual abuse, younger children (those aged one year and under), are the most vulnerable of all children to be abused and neglected, and younger children are generally more vulnerable to abuse and neglect than older children (US Dept HHS 2010). While its true extent is unknown, child abuse and neglect is a well-established problem worldwide (Pinheiro 2006). Numerous prevalence studies conducted at the population level have established that the various forms of child abuse and neglect are very widespread, although some forms of abuse and neglect are more common than others (Finkelhor 2005; Finkelhor 2010; Radford 2012; Sedlak 2010; Stoltenborgh 2011).

The adverse effects of child abuse and neglect are significant and can endure throughout a person's life. The most serious consequence is fatality, with an estimated 155,000 deaths globally per annum (WHO 2006). Other effects include: physical injuries; failure to thrive; impaired social, emotional and behavioural development; reduced reading ability and perceptual reasoning; depression; anxiety; post-traumatic stress disorder; low self-image; alcohol and drug use; aggression; delinquency; long-term deficits in educational achievement; and adverse effects on employment and economic status (Egeland 2009; Gilbert 2009; Hildyard 2002; Landsford 2002; Norman 2012; Paolucci 2001). For society, effects include lost productivity and cost to child welfare systems (Currie 2010; Fang 2012), and intergenerational victimisation (Draper 2008). The annual economic cost in the USA is estimated at USD 124 billion, based on a cost per non-fatal case of USD 210,012 (Fang 2012).

Although there is some variance across cultures in perceptions of what may and may not constitute child abuse and neglect (Finkelhor 1988; Korbin 1979), in recent decades there is an emerging consensus about its parameters, especially for child sexual abuse and severe physical abuse and neglect. This is reflected in the United Nations Convention on the Rights of the Child, which has been almost universally ratified, criminal prohibitions on this conduct across developed and developing countries, and scholarly research (Finkelhor 1988).

#### Professional reporting duties

In an attempt to respond to child abuse and neglect, and to enable early intervention to assist the child and his or her family, many governments require members of selected professional groups to report suspected cases of significant child abuse and neglect (Mathews 2008a). The duty to report is usually conferred on professionals who frequently

encounter children in their daily work such as teachers, nurses, doctors, and police (Mathews 2008b). In some jurisdictions, these reporting duties have been enacted in legislation (called 'mandatory reporting laws'), but in others, the duties appear only in occupational employment policy. Differences do exist between jurisdictions in some aspects of these duties, but consistency exists across jurisdictions and professions in the essential nature of the duties (Mathews 2008a). This enables comparisons to be made across key dimensions of professional training in different contexts.

Studies have found that professionals who are required to report child abuse and neglect consider they have not had sufficient training to fulfil their role (Abrahams 1992; Christian 2008; Hawkins 2001; Kenny 2001; Kenny 2004; Mathews 2011; Reiniger 1995; Starling 2009; Walsh 2008). Research has also found low levels of knowledge about the nature of the reporting duty (Beck 1994; Mathews 2009) and indicators of abuse and neglect (Hinson 2000). Studies have also found that professionals may hold attitudes which are unconducive to reporting (Feng 2005; Jones 2008; Kalichman 1993; Mathews 2009; Zellman 1990). Effective reporting is thought to be influenced by several factors, including knowledge of the duty (Crenshaw 1995; Kenny 2004), the ability to recognise abuse (Crenshaw 1995; Goebbels 2008; Hawkins 2001), and positive attitudes towards the duty (Fraser 2010; Goebbels 2008; Hawkins 2001).

Improved reporting offers the prospect of enhanced detection of child abuse and neglect, the provision of interventions and redress for victims (Kohl 2009), and engagement with parents and caregivers to establish supportive measures (Drake 1996; Drake 2007). Improved reporting by professionals should also diminish clearly unnecessary reports and avoid wasting of scarce government resources and distress to families (Ainsworth 2006). In addition, effective child protection training for professionals should also assist in developing greater understanding of legal protections conferred on professional reporters themselves, and avoidance of potential legal liability and professional discipline for noncompliance. At its best, child protection training could also enhance professional ethical identities and contribute to broader workforce professionalisation.

### *Description of the intervention*

This Cochrane review focuses on child protection training that takes place after initial qualification (i.e. post-qualification), also referred to in the literature as "continuing professional education", "continuing professional development", and "professional development". Child protection training interventions aim to improve reporting of child abuse and neglect by professionals who are required by law or policy to do so. The core aim of such training is to improve the reporting of cases where abuse or neglect exists or can reasonably be thought to exist; and to reduce the making of clearly unnecessary reports where there are insufficient grounds on which a knowledgeable reporter would make a report.

Different approaches may be taken in training professionals to improve reporting. Training may focus on increasing awareness of the indicators of each type of abuse and neglect, the nature of the reporting duty, or procedures for reporting. Training may also focus on enhancing reporters' attitudes towards the reporting duty or to child protection in general. Training can be implemented in different formats, for example, single professional development sessions or extended courses. Different delivery modes can be adopted, for example, online or face-to-face (Kenny 2001; McGrath 1987). Training may vary in duration (Donohue 2002; Hazzard 1984), and may be targeted at different skill levels (e.g. basic, advanced).

### *How the Intervention Might Work*

It can be hypothesised that, viewed as an application of adult learning (Knowles 2011), child protection training for professionals is an educational intervention through which professionals develop knowledge, skills, attitudes, and behaviours. By raising awareness, providing information and resources, developing skills and strategies, and fostering dispositions, training may change professionals' ability and willingness to engage in decision-making processes that lead to improved reporting. There is some evidence to suggest that, for some professions and for some types of abuse, exposure to training is associated with effective reporting (Fraser 2010; Walsh 2012), self-reported preparedness to report (Fraser 2010), confidence identifying abuse (Hawkins 2001), and awareness of reporting responsibilities (Hawkins 2001). Some studies have indicated that lack of adequate training is associated with low awareness of the reporting duty (Hawkins 2001), low preparedness to report (Kenny 2001), low self-reported confidence identifying child abuse (Hawkins 2001; Mathews 2008b; Mathews 2011), and low knowledge of indicators of abuse (Mathews 2011). However the specific components of training, which are responsible for improving reporting practices (e.g. reporting deserving cases, and not reporting undeserving cases), are not yet known.

### *Why it is Important to do the Review*

Child abuse and neglect results in significant costs for children and communities. As a core public health strategy, diverse professional groups are required by law and policy in many jurisdictions to report suspected cases. Numerous different training initiatives appear to have been developed and implemented for professionals but there is little evidence regarding the precise training components and mechanisms that improve reporting of child abuse and neglect both generally, for specific professions, and for distinct types of child abuse and neglect. To enhance reporting practice, designers of training programmes require detailed information about what programme features will offer greatest benefit. A systematic review which identifies the effectiveness of different training approaches will advance the evidence base and develop a clearer understanding of optimal training content and methods. In addition, it will provide policymakers with a means by which to assess whether current

training interventions are congruent with what is demonstrated to be effective. It will also inform future research, public policy, and professional practice in this field.

## 2. OBJECTIVES

To assess the effectiveness of training aimed at improving reporting of child abuse and neglect by professionals and to investigate possible components of effective training interventions.

## 3. METHODOLOGY

### *Criteria for including and excluding studies*

*Types of studies*

Randomised controlled trials (RCTs), quasi-RCTs (i.e. studies in which participants were assigned to intervention or comparison or control groups by a quasi-randomised method such as allocation by date of birth, or similar methods) and controlled before-and-after studies (i.e. studies where participants were allocated to intervention and control groups by means other than randomisation, but take into account baseline measurements of main outcomes; contemporaneous data collection for pre- and post-test intervention periods).

We will include controlled before-and-after studies because studies of educational interventions are often conducted in natural settings where randomised designs are not feasible. We will use explicit study design features rather than study design labels when deciding which types of non-randomised studies to include. We will follow the guidance on how to assess and report on non-randomised studies in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011).

All studies must evaluate the effects of training on at least one of the outcomes listed below. We will examine all designs closely for threats to validity. We will include studies irrespective of publication date, language, type, and status.

*Types of interventions*

**Included**

Child protection training interventions aimed explicitly at improving reporting of child abuse and neglect by qualified professionals, irrespective of programme type, mode, content, duration, intensity, and delivery context. These interventions will be compared with no training, wait-list control, or comparison training not related to child abuse and neglect (e.g. first aid training).

**Excluded**

Training interventions where improving professionals' reporting of child abuse and neglect is a minor focus, such as brief professional induction or orientation programmes targeting a broad range of employment responsibilities, where it would not be possible to isolate the specific effects of the child protection training component. Child protection training or education conducted before professional qualifications have been obtained, for example, as part of undergraduate college or university-level professional preparation programmes (e.g. initial teacher training, pre-service education for nurses, entry-level medical education, or basic police education).

*Types of outcome measures*

**Primary outcomes**

1. Changes in the number of reported cases of child abuse and neglect:

- as measured subjectively by participant self-reports of actual cases reported;
- as measured subjectively by participant responses to vignettes; and
- as measured objectively in official records of reports made to child protection authorities.

2. Changes in the quality of reported cases of child abuse and neglect, as measured via coding of the actual contents of reports made to child protection authorities (i.e. in government records or archives).

3. Adverse events:

- increase in failure to report cases of child abuse and neglect that warrant a report as measured subjectively by participant self-reports (i.e. in questionnaires); and
- increase in reporting of cases that do not warrant a report as measured subjectively by participant self-reports (i.e. in questionnaires).

It should be noted that studies using official records, which identify a change in relevant outcomes, such as the number of reports made and the number of reports substantiated after investigation, may indicate improved reporting effectiveness after reporter training but are not themselves determinative. Therefore, analysis and interpretation of such results would not occur in isolation but would also be informed by consideration of other important contextual factors, including: whether the aim of the training was to increase reports of a type of abuse, to decrease reports of another phenomenon, or both; and circumstances surrounding the reporter training such as the introduction of a new duty or the

implementation of training as a response to a high profile case or inquiry. Similarly, studies using official records cannot measure some aspects of reporting behaviour such as false negatives (i.e. where a case provided grounds to suspect maltreatment, and the professional suspected maltreatment but did not report).

### Secondary outcomes

These include objectively or subjectively measured outcomes closely associated with improved reporting practice, and that may help to account for how the interventions may work.

1. Knowledge of the reporting duty, processes, and procedures.
2. Knowledge of core concepts in child abuse and neglect such as the nature, extent, and indicators of the different types of abuse and neglect.
3. Skill in distinguishing between cases which should be reported from those that should not.
4. Attitudes towards the duty to report child abuse and neglect.

We will include all primary and secondary outcomes in a 'Summary of findings' table.

### Timing of outcome assessment

We will categorise primary and secondary outcomes into three time periods: short-term outcomes (assessed immediately after the intervention and up to 12 months after the intervention); medium-term outcomes (assessed between one and three years after the intervention); and long-term outcomes (assessed more than three years after the intervention).

## *Search strategy*

### Electronic searches

We will search the following databases.

1. Cochrane Central Register of Controlled Trials (CENTRAL), part of the Cochrane Library (current issue).
2. Ovid Medline® (1946 to current).
3. Embase (embase.com) (1966 to current).
4. CINAHL (EBSCOhost) (1981 to current).
5. ERIC (EBSCOhost) (1966 to current).

6. PsycINFO (EBSCOhost) (1966 to current).
7. Social Services Abstracts (ProQuest Research Library) (1966 to current).
8. Science Direct (Elsevier) (1966 to current).
9. Sociological Abstracts (ProQuest Research Library) (1952 to current).
10. ProQuest Psychology Journals (ProQuest Research Library) (1966 to current).
11. ProQuest Social Science (ProQuest Research Library) (1966 to current).
12. ProQuest Dissertations and Theses (ProQuest Research Library) (1997 to current).
13. Social Policy and Practice (Ovid) (1860 to current).
14. Lexis (Lexis.com) (1980 to current).
15. LegalTrac (GALE) (1980 to current).
16. Westlaw International (Thomson Reuters) (1980 to current).
17. Conference Proceedings Citation Index – Social Science & Humanities (Web of Science) (1990 to current).
18. Database of Abstracts of Reviews of Effects (DARE), part of the Cochrane Library (current issue).
19. Cochrane Database of Systematic Reviews, part of the Cochrane Library (current issue).
20. Violence and Abuse Abstracts (EBSCOhost) (all available years).
21. LILACS (lilacs.bvsalud.org/en/) (all available years).
22. ClinicalTrials.gov (clinicaltrials.gov/).
23. World Health Organization International Clinical Trials Registry Platform (WHO ICTRP) (www.who.int/ictrp/en/).
24. Australia and New Zealand Clinical Trials Registry (www.anzctr.org.au/).
25. OpenGrey (www.opengrey.eu/).

We will search Ovid MEDLINE using the strategy in Appendix 1, which includes the Cochrane highly sensitive search strategy for identifying randomised trials (Lefebvre 2008). No date or language limits will be applied. We will adapt the strategy for other sources as appropriate.

### Searching other resources

To identify studies not obtained by searching the databases listed above, we will carry out additional searches. We will handsearch the following journals.

1. Child Maltreatment.
2. Child Abuse and Neglect.
3. Children and Youth Services Review.
4. Trauma, Violence and Abuse.
5. Child Abuse Review.

We will modify the search strategy and apply it to each of these journals.

We will also search a number of key websites for additional studies.

1. International Society for Prevention of Child Abuse and Neglect via www.ispcan.org/.
2. US Department of Health and Human Services Children's Bureau, Child Welfare Information Gateway via https://www.childwelfare.gov/.
3. Promising Practices Network operated by the RAND Corporation via http://www.promisingpractices.net/.
4. National Resource Center for Community-Based Child Abuse Prevention (CBCAPP) via http://friendsnrc.org/.
5. California Evidence-Based Clearinghouse for Child Welfare (CEBC) via http://www.cebc4cw.org/.
6. Coalition for Evidence-Based Policy via http://coalition4evidence.org/.
7. Institute of Education Sciences What Works Clearinghouse via http://ies.ed.gov/ncee/wwc/.
8. National Institute for Health and Care Excellence (NICE) UK via www.nice.org.uk/.

Finally, we will handsearch the reference lists of included studies in order to identify further potential studies. We will also contact key researchers in this field for unpublished studies.

## *Data collection and analysis*

### Selection of studies

Using reference management software (e.g. EndNote), two review authors (BM, KW) will independently screen titles and abstracts of studies identified from the searches to determine if they meet eligibility criteria. Criteria will include: study design; participants; type of intervention; and types of comparisons. At this stage, we will reject studies if the title and abstract clearly indicate that the report does not meet these criteria. Two review authors (BM, KW) will independently retrieve and assess the full text of studies that appear to meet the eligibility criteria. If insufficient information is provided in the paper to assess eligibility for inclusion, we will contact study authors to provide missing information. We will link together multiple publications and reports on the same study. Where necessary, we will translate studies into English with the assistance of translators. We will resolve differences of opinion regarding the eligibility of a study for inclusion through discussion and consensus. If agreement cannot be reached, we will elicit the opinion of a third author (MK), whereby the final list of included and excluded studies will be decided. We will document primary reasons for study exclusion.

### Data extraction and management

We will develop and pilot test a data extraction form based upon the checklist of items from the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, Table 7.3a). We will extract data from study reports concerning details of: study design and methods, participants, setting, intervention group(s), control group(s), intervention content and processes, outcome measures, raw data, and data analysis. We will enter this information into Review Manager 2014 and present it in a 'Characteristics of included studies' table for each included study. Two review authors (BM, KW) with different disciplinary backgrounds will independently complete a data extraction form for each study. In the event that a study report lacks relevant information for the study design, we will contact study authors for further details and will record responses. To reduce the risk of overly positive responses, we will use open-ended questions such as "please describe measures used to ….". A third review author will cross-check data collection forms and we will resolve any discrepancies via discussion and consensus.

### Assessment of risk of bias in included studies

Two review authors (KW, MK) will independently assess risk of bias for included studies (i.e. the risk that studies may over- or under-estimate the intervention's actual effect) using the Cochrane revised 'Risk of bias' assessment tool (Higgins 2011, Table 8.5a). We will include the tool as a section within the data collection form described above.

The tool consists of seven domains. For randomised studies, we will add an eighth domain, reliability of outcome measures, as we anticipate that some studies may use custom-made

instruments and scales. For non-randomised studies only, we will add two further domains: group comparability and contamination.

For each included study, we will judge the relevant domains as 'low', 'high', or 'unclear' risk of bias. We will make our judgements by answering 'yes', 'no', or 'unclear' to pre-specified questions as follows.

**Sequence generation**

Description: the method used to generate the allocation sequence was described in sufficient detail to enable assessment of whether it could produce comparable baseline groups.

Question: do study authors describe a random component in the sequence generation process?

**Allocation concealment**

Description: the method used to conceal the allocation sequence was described in sufficient detail to determine whether allocations could have been predicted before or during the assignment-to-groups process.

Question: do study authors report an adequate method of concealing allocation to intervention or control groups?

**Blinding of participants and personnel**

Description: the methods used, if any, to blind study participants and personnel from knowledge of participants' group membership were described in sufficient detail to enable assessment of their effectiveness.

Question: do the study authors report an adequate method of participants and personnel from knowledge of participants' belonging to either intervention or control groups?

**Blinding of outcome assessment**

Description: the methods used to blind outcome assessors from knowledge of participants' group membership were described in sufficient detail to enable assessment of their effectiveness.

Question: do study authors note blinding of outcome assessors from knowledge of participants' belonging to either intervention or control groups?

**Incomplete outcome data**

Description: data on attrition, exclusions, and withdrawals were reported (numbers compared with the total number randomised or as a proportion of the total number randomised, or both), and reasons for incomplete outcome data were provided.

Question: do study authors report missing data, reasons for missing data, and imputation methods?

**Selective reporting**

Description: the study's pre-specified primary and secondary outcomes were reported in sufficient detail to assess their completeness.

Question: do study authors report on all pre-specified outcomes of interest?

**Other sources of bias**

Description: the study was free from other sources of bias such as fraudulence.

Question: was the study free of other problems that could put the study at risk of bias?

**Reliability of outcome measures**

Description: the study outcomes were measured using reliable instruments or scales (Cronbach's alpha of 0.6 or above), and reliability scores were reported or could be found in other publications.

Question: do the study authors report reliability data in sufficient detail to enable its assessment?

**Group comparability**

Description: information on the comparability of groups at baseline was provided in sufficient detail for each outcome measure to enable its assessment.

Question: do the study authors report group comparability at baseline for each of the outcome measures of interest?

**Contamination**

Description: the measures taken to prevent or minimise the possibility that participants in a control group might receive part or all of the intervention were described in sufficient detail to enable assessment of contamination between groups.

Question: do study authors report contamination minimisation measures or ways in which contamination may have been possible (e.g. media reports during a training intervention period)?

Wherever possible, we will use verbatim text from the study reports or correspondence with study authors (appropriately cited) as support for our risk of bias judgments. Review authors assessing risk of bias will not be blinded to the names of the authors, institutions, journals, or results of studies. We will resolve disagreements between review authors by discussion, and where consensus cannot be reached, by consulting with a third review author. For studies in which essential information is not available, we will contact study authors with an open-ended request for missing information (as noted above). We will enter the information into Review Manager 2014 and summarise it in 'Risk of bias' tables for each included study. We will also present two figures: a 'Risk of bias' graph illustrating the proportion of studies for each risk of bias criterion, and a 'Risk of bias' summary graph visually depicting our judgements across all studies. From here, our strategy will be to conduct multiple sensitivity analyses for each outcome to show how results might be affected by our inclusion/exclusion of studies at high risk of bias. We will also provide a narrative discussion of the risk of bias.

### Measures of treatment effect

We will report treatment effects for outcomes separately.

### Continuous data

We will report continuous data using means and standard deviations (SDs). We will summarise study effects as mean differences (MD) and 95% confidence intervals (CIs) for continuous data where the same scale is used to measure similar outcomes. We will use standardised mean differences (SMD) and 95% CIs where different scales are used to measure the same outcome.

### Dichotomous data

We will report dichotomous data with raw counts and rates for intervention and control groups. We will summarise dichotomous data using risk ratios (RR) with corresponding 95% CIs. For the primary outcome, this statistic could be expressed, for example, as the risk of failure to report child maltreatment in the intervention group compared with the risk of failure to report in the control group.

### Unit of analysis issues

#### Cluster-randomised trials

Cluster-randomised trials are widespread in the evaluation of health care and educational interventions (Donner 2002), but they can be poorly reported (Campbell 2004). For included studies with incorrectly analysed data that does not account for clustering, we will adjust sample sizes according to procedures outlined in *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, Section 16.3.4).

Initially, we plan to use an estimate of the intracluster correlation coefficient (ICC) from an included study that adequately accounts for a clustered design and reports an ICC. If no studies directly report an ICC, this value may be imputed from other sources such as studies in similar areas, with similar populations, or meta-analysis of other similar subjects. Recent reviews have compiled a range of empirically-based ICCs for professional development interventions with teachers (Kelcey 2013), and primary care providers (Eccles 2003), and have reported individual trial ICCs of between 0.15 and 0.21 (teachers) and 0.01 and 0.16 (primary care). A suitably conservative approach is to conduct calculations using an ICC of 0.20.

We will adjust study sample sizes according to the ICC using the procedure outlined in the Cochrane Handbook for Systematic Reviews of Interventions (Higgins 2011, Section 16.3.4). We will test the robustness of these assumptions in Sensitivity analysis, where we will use at least the two extreme values of ICC reported in the literature for each subgroup of professional. This is important as different assumed values for ICCs will affect the weights assigned to the different trials. Furthermore, we will check if the results of the cluster analyses are similar to or different from that of non-cluster trials. If the results are markedly different we will explore potential reasons and, depending on reasons and number of trials, will report the results separately.

### Studies with multiple treatment groups

In trials with multiple intervention groups, control groups, or both, also known as multi-arm studies, we will first determine which intervention groups are most relevant to the review according to the intervention type and outcomes assessed. Where appropriate, we will combine all relevant intervention groups into a single intervention group and all control groups into a single control group, to enable a single pairwise comparison. This will be undertaken using the calculator tool in Review Manager 2014. For dichotomous data, we will sum sample sizes and events across groups. For continuous data, we will combine sample sizes, means, and SDs according to the formula detailed in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, Section 7.7.3.8).

## Dealing with missing data

Missing data may be in the form of missing studies, missing outcomes, missing summary data, or missing participants. We do not anticipate missing studies, as our search strategy will be comprehensive and we will take all reasonable steps to locate the full texts of eligible studies.

In studies with missing outcomes (owing to selective reporting) or missing summary data, we will contact first-named study authors via email with a request to provide the data. For continuous data, where possible, we will calculate missing SDs from other test statistics (e.g. t values, F values). In cases where SDs are unavailable and cannot be calculated, we will impute an average SD from other included studies as this method has been found to produce

approximately correct results (Higgins 2011, Section 16.1.3.1). We will assess the extent to which this alters results in a Sensitivity analysis.

For studies with participants missing from trial analyses or incomplete outcome data (owing to attrition or exclusion), we will contact first-named study authors via email with a request for further information. If data are available, we will conduct analyses including the participants who were excluded by study authors. If data are not available, we will conduct analyses using available data only and will not impute values. We will report the extent of missing data and approaches to imputation within individual studies in the 'Risk of bias' tables.

**Assessment of heterogeneity**

To assess the extent of variation between studies, we will examine distributions of relevant participant (e.g. professional discipline), delivery (e.g. classroom), and trial (e.g. type and duration of intervention) variables. Using forest plots available in Review Manager 2014, we will visually examine CIs for the outcome results of individual studies paying particular attention to poor overlap, which can be used as an informal indicator of statistical heterogeneity (Higgins 2011). In Review Manager 2014 we will examine three estimates investigating different aspects of heterogeneity as recommended by Borenstein 2009. First, as a test of statistical significance of heterogeneity, we will examine the Q statistic and its P value. For any observed Q, a low P value provides evidence of heterogeneity of intervention effects (i.e. that studies do not share a common effect size) (Higgins 2011). Second, as an estimate of the magnitude of variation between studies, we will estimate and present $Tau^2$ along with its CIs. This will give us an estimate of the amount of between study variation. Third, we will estimate the $I^2$ statistic and its CIs, which describes the proportion of variability in effect estimates due to heterogeneity rather than chance (Higgins 2011). These three quantities (Q, $Tau^2$, and the $I^2$ statistic), along with the appropriate CIs, will give us a good picture of the presence and the degree of heterogeneity among the studies. They are viewed as complementary rather than mutually exclusive quantities. Rather than defaulting to interpretations of heterogeneity based on rules of thumb (i.e. that an $I^2$ statistic value of 30% to 60% represents moderate heterogeneity, 50% to 90% represents substantial heterogeneity, and 75% to 100% represents considerable heterogeneity), we will use all three measures of heterogeneity (Q, $Tau^2$, and the $I^2$ statistic) to fully describe the aspects of variability in the data as detailed in Borenstein 2009. For example, $Tau^2$ or the $I^2$ statistic (or both) will be used to measure the magnitude of true variation, and the P value for Q or CIs for $Tau^2$ or the $I^2$ statistic will be used as an indicator of uncertainty regarding the genuineness of the heterogeneity. This provides essential detail for judging the presence and magnitude of heterogeneity. Substantial heterogeneity may render a group of studies unsuitable for meta-analysis. We will further strive to understand the reasons for potential presence of heterogeneity.

**Assessment of reporting biases**

We will assess reporting bias in the form of selective outcome reporting as one of the domains within the 'Risk of bias' assessments. However, we do not expect to find published protocols for studies included in this Cochrane review to use for comparative purposes.

We will assess publication bias. If there are sufficient studies (at least 10, using the rule of thumb in the *Cochrane Handbook for Systematic Reviews of Interventions* (Higgins 2011, Section 10.4.3.1)), we will draw funnel plots in Review Manager 2014 to assess the relationships between effect estimates and their standard errors (on a reversed scale). We will use visual inspection in the first instance. If funnel plots are found to be asymmetrical, we will consider possible explanations and take these into account in the interpretation of the overall estimate of treatment effects, including but not restricted to small study effects. If the latter is suspected and there are enough studies present, we will perform both a fixed-effect and a random-effects analysis and we will compare the results, which will aid in the detection of such effect. We note that asymmetrical funnel plots (small study effects) are not always indicative of the presence of publication bias (Higgins 2011).

**Data synthesis**

We will assess the appropriateness of combining studies based on sufficient homogeneity with respect to: the training interventions delivered (these should be similar in content and method), the study population characteristics (such as professional group), measurement tools or scales used (these should report on similar primary or secondary outcomes), and summary points (outcomes should be measured within comparable timeframes pre- and post-intervention). We will combine data for comparable professional groups (e.g. elementary and high school teachers), and similar outcome measures. We will conduct separate analysis for training type (e.g. online or face-to-face training). We will use meta-analysis to compute pooled estimates of intervention effects for those studies for which data are available and can be appropriately combined.

For those studies for which data can be combined, we will calculate summary statistics (RR for dichotomous data, and MD or SMD for continuous data) and 95% CIs for each outcome. In the meta-analyses, we will first generate fixed-effect models for combining data where we have judged that studies are estimating the same underlying treatment effect. That is, where studies report data on training interventions with analogous contents, or with comparable professional groups, or measured in similar time frames. Fixed-effect models ignore heterogeneity, but are generally interpreted as being the best estimate of the intervention effect (Higgins 2011). If there is evidence of substantial heterogeneity, assessed as above using Q, $Tau^2$, and the $I^2$ statistic, we will also generate random-effect models, which can account for diversity among studies (by assuming that included studies may not all estimate precisely the same intervention effect), and provide a more conservative estimate of effect (Higgins 2011). We will compare the results of the fixed-effect and the random-effects models to assess the impact of statistical heterogeneity. If results converge, we will report the results of the random-effects models only. If results diverge, we will report the results of both

models. Where possible, we will report the clinical significance of the results of the meta-analysis in the results section of the review. If meta-analysis is inappropriate, we will include a narrative overview to qualitatively synthesize the data.

Two review authors (KW, MK) will independently code and categorise intervention contents and the resulting typology will be determined via consensus among all review authors (Marusic 2013).

It should be noted that studies using official records, which identify a change in relevant outcomes, such as the number of reports made and the number of reports substantiated after investigation, may indicate improved reporting effectiveness after reporter training but are not themselves determinative. Analysis and interpretation of such results would therefore not occur in isolation but would also be informed by consideration of other important contextual factors, including: whether the aim of the training was to increase reports of a type of abuse, to decrease reports of another phenomenon, or both; and circumstances surrounding reporter training, such as the introduction of a new duty, or the implementation of training as a response to a high profile case or inquiry. Similarly, studies using official records cannot measure some aspects of reporting behaviour such as false negatives (i.e. where a case provided grounds to suspect maltreatment, and the professional suspected maltreatment, but did not report it).

### 'Summary of findings' table

We will construct and present a 'Summary of findings' table and will rate the quality of evidence for all primary and secondary outcomes using methods developed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group (http://www.gradeworkinggroup.org/index.htm). The GRADE system classifies the quality of evidence in one of four categories: (i) high quality, when further research is very unlikely to change our confidence in the estimate of effect; (ii) moderate quality, when further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate; (iii) low quality when further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate; or (iv) very low quality, when we are very uncertain about the estimate of the effect. We will consider the following factors when grading the quality of evidence: study design, risk of bias, precision of effect estimates, consistency of results, directness of evidence, and magnitude of effect (Guyatt 2011).

### Subgroup analysis and investigation of heterogeneity

We will investigate any heterogeneity observed in the results of included studies. If there is enough available comparable data, that is at least 10 studies (Higgins 2011, Section 9.6.5), we will undertake the following subgroup analyses:

- training method (face-to-face or online);

- training delivered by specialist or non-specialist trainers;
- training duration (single or multiple sessions); and
- time of study (less recent or more recent studies).

We will assess the differences between subgroups by inspection of the CIs which, if not overlapping indicate a statistically significant difference in training effects between the subgroups. We will also examine and report interaction effects using analysis of variance (ANOVA).

**Sensitivity analysis**

We will perform a sensitivity analysis to test the robustness of decisions made in this Cochrane review. We will do this first by separating randomised from non-randomised studies, and second by separating cluster-randomised studies where there are concerns about failure to adjust for clustering. We will also perform a sensitivity analysis to explore the impact of study quality by removing studies from specific outcome analyses assessed to be at high risk of bias on the domains most relevant to each outcome (including other sources of bias identified for this review, such as reliability of outcome measures). Where feasible, we will treat studies judged to be at low risk of bias as a separate group. In a series of sensitivity analyses, we will explore how the results of meta-analyses might be affected by excluding unpublished studies (i.e. theses), and studies with selective outcome reporting. If the analysis of heterogeneity finds outlying studies with results that appear vastly different from other included studies, we will perform a sensitivity analysis to assess the effect on the results of meta-analyses. We will also conduct a sensitivity analysis to assess the impact of imputing missing data (e.g. SDs and ICCs as outlined above).

## 4. APPENDIX 1

### *Table 1: Search strategy*

1. exp Child Abuse/
2. Child Welfare/
3. ((baby or babies or infant$ or child$ or preschool$ or pre-school$ or teen$ or adolescen$) adj3 (abuse$ or matreat$ or mal-treat$ or neglect$)).tw.
4. ((baby or babies or infant$ or child$ or preschool$ or pre-school$ or teen$ or adolescen$) adj3 (protect$ or safeguard$ or safe-guard)).tw.
5. ((at risk or high risk) adj1 child$).tw.
6. or/1-5
7. exp Child/
8. adolescent/
9. (baby or babies or infant$ or child$ or teen$ or adolescen$).tw.
10. or/7-9
11. ((non-accidental or deliberate) adj3 injur$).tw.
12. ((emotional$ or psycholog$) adj3 (abuse$ or matreat$ or mal-treat$ or neglect$)).tw.
13. sex offenses/ or rape/
14. ((sex$ adj3 abuse$) or rape or incest$).tw.
15. or/11-14
16. 10 and 15
17. 6 or 16
18. exp education, professional/
19. exp inservice training/
20. exp Teaching/
21. education.fs.
22. Health Knowledge, Attitudes, Practice/
23. Clinical Competence/
24. ((education$ or instruction$ or teach$ or train$) adj3 (program$ or intervention$ or course$ or model$ or post-qualif$ or continuing )).tw.
25. ((education$ or instruction$ or teach$ or train$) adj3 (dentist$ or doctor$ or medic$ or midwi#e$ or nurs$ or social worker$ or social service$ or police$ or teacher$ or health professional$)).tw.
26. Mandatory Reporting/
27. (mandatory adj1 (notif$ or report$)).tw.
28. or/18-27
29. 17 and 28
30. ((child abuse or sex$ abuse) adj1 (detect$ or diagnos$ or education or training)).tw.
31. 29 or 30
32. randomized controlled trial.pt.
33. controlled clinical trial.pt.
34. randomi#ed.ab.
35. placebo$.ab.
36. drug therapy.fs.
37. randomly.ab.
38. trial.ab.
39. groups.ab.
40. or/32-39
41. exp animals/ not humans.sh.
42. 40 not 41
43. 31 and 42

## 5. REFERENCES


Abrahams N, Casey K, Daro D. Teachers' knowledge, attitudes, and beliefs about child abuse and its prevention. Child Abuse & Neglect 1992;16(2):229-38.

Ainsworth F, Hansen P. Five tumultuous years in Australian child protection: little progress. Child & Family Social Work 2006;11(1):33-41.

Beck KA, Ogloff JRP, Corbishley A. Knowledge, compliance, and attitudes of teachers toward mandatory child abuse reporting in British Columbia. Canadian Journal of Education 1994;19(1):15-29.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Introduction to Meta-Analysis. West Sussex: John Wiley & Sons, Ltd, 2009.

Campbell MK, Grimshaw JM, Elbourne DR. Intracluster correlation coefficients in cluster randomized trials: empirical insights into how they should be reported. BMC Medical Research Methodology 2004;4:9.

Christian CW. Professional education in child abuse and neglect. Pediatrics 2008;122(Suppl 1):S13-7.

Crenshaw WB, Crenshaw LM, Lichtenberg JW. When educators confront child abuse: an analysis of the decision to report. Child Abuse & Neglect 1995;19(9):1095-113.

Currie J, Widom CS. Long-term consequences of child abuse and neglect on adult economic well-being. Child Maltreatment 2010;15(2):111-20.

Donner A, Klar N. Issues in the meta-analysis of cluster randomized trials. Statistics in Medicine 2002;21(19):2971-80.

Donohue B, Carpin K, Alvarez KM, Ellwood A, Jones RW. A standardized method of diplomatically and effectively reporting child abuse to state authorities: a controlled evaluation. Behavior Modification 2002;26(5):684-99.

Drake B. Unraveling "unsubstantiated". Child Maltreatment 1996;1(3):261-71.

Drake B, Jonson-Reid M. A response to Melton based on the best available data. Child Abuse & Neglect 2007;31(4):343-60.

Draper B, Pfaff JJ, Pirkis J, Snowdon J, Lautenschlager NT, Wilson I, et al. Long-term effects of childhood abuse on the quality of life and health of older people: results from the Depression and Early Prevention of Suicide in General Practice Project. Journal of the American Geriatrics Society 2008;56(2):262-71.

Eccles M, Grimshaw J, Campbell M, Ramsay C. Research designs for studies evaluating the effectiveness of change and improvement strategies. Quality and Safety in Healthcare 2003;12(1):47-52.

Egeland B. Taking stock: childhood emotional maltreatment and developmental psychopathology. Child Abuse & Neglect 2009;33(1):22-6.

Fang X, Brown DS, Florence CS, Mercy JA. The economic burden of child maltreatment in the United States and implications for prevention. Child Abuse & Neglect 2012;36(2):156-65.

Feng JY, Levine M. Factors associated with nurses' intention to report child abuse: a national study of Taiwanese nurses. Child Abuse & Neglect 2005;29(7):783-95.

Finkelhor D, Korbin J. Child abuse as an international issue. Child Abuse & Neglect 1988;12(1):3-23.

Finkelhor D, Ormrod R, Turner H, Hamby SL. The victimization of children and youth: a comprehensive, national survey. Child Maltreatment 2005;10(1):5-25.

Finkelhor D, Turner H, Ormrod R, Hamby SL. Trends in childhood violence and abuse exposure:evidence from 2 national surveys. Archives of Pediatrics & Adolescent Medicine 2010;164(3):238-42.

Fraser JA, Mathews B, Walsh K, Chen L, Dunne M. Factors influencing child abuse and neglect recognition and reporting by nurses: a multivariate analysis. International Journal of Nursing Studies 2010;47(2):146-53.

Gilbert R, Widom CS, Browne K, Fergusson D, Webb E, Janson S. Burden and consequences of child maltreatment in high-income countries. Lancet 2009;373(9657):68-81.

Goebbels A, Nicholson J, Walsh K, De Vries H. Teachers' reporting of suspected child abuse and neglect: behaviour and determinants. Health Education Research 2008;23(6):941-51.

Guyatt G, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines 1. Introduction-GRADE evidence profiles and summary of findings tables. Journal of Clinical Epidemiology 2011;64(4):383–94.

Hawkins R, McCallum C. Mandatory notification training for suspected child abuse and neglect in South Australian schools. Child Abuse & Neglect 2001;25(12):1603-25.

Hazzard A. Training teachers to identify and intervene with abused children. Journal of Clinical Child Psychology 1984;13(3):288-93.

Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011. Available from http://handbook.cochrane.org/.

Hildyard KL, Wolfe DA. Child neglect: developmental issues and outcomes. Child Abuse & Neglect 2002 ;26(6-7):679-95.

Hinson J, Fossey R. Child abuse: what teachers in the '90s know, think, and do. Journal of Education for Students Placed at Risk 2000;5(3):251-66.

Jones R, Flaherty EG, Binns HJ, Price LL, Slora E, Abney D, et al. Clinicians' description of factors influencing their reporting of suspected child abuse: report of the Child abuse Reporting Experience Study Research Group. Pediatrics 2008;122(2):259-66.

Kalichman SC, Brosig CL. Practicing psychologists' interpretations of and compliance with child abuse reporting laws. Law and Human Behavior 1993;17(1):83-93.

Kelcey B, Phelps G. Considerations for designing group randomized trials of professional development with teacher knowledge outcomes. Educational Evaluation and Policy Analysis 2013;35(3):370-90.

Kenny MC. Child abuse reporting: teachers' perceived deterrents. Child Abuse & Neglect 2001;25(1):81-92.

Kenny MC. Teachers' attitudes toward and knowledge of child maltreatment. Child Abuse & Neglect 2004;28(12):1311-9.

Knowles MS, Holton EF III, Swanson RA. The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development. 7th edition. Oxford: Butterworths-Heinemann Ltd, 2011.

Kohl P, Jonson-Reid M, Drake B. Time to leave substantiation behind: findings from a national probability study. Child Maltreatment 2009;14(1):17-26.

Korbin J. A cross-cultural perspective on the role of community in child abuse and neglect. Child Abuse & Neglect 1979;3(1):9-18.

Landsford JE, Dodge KA, Pettit GS, Bates JE, Crozier J, Kaplow J. A 12 year prospective study of the long-term effects of early child physical maltreatment on psychological, behavioural, and academic problems in adolescence. Archives of Pediatric & Adolescent Medicine 2002;156(8):824-30.

Lefebvre C, Manheimer E, Glanville J. Chapter 6: Searching for studies. In Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011], The Cochrane Collaboration, 2011. Available from http://handbook.cochrane.org/.

Marusic A, Wager E, Utrobicic A, Sambunjak D, Anderson MS, Rothstein HR. Interventions to prevent misconduct and promote integrity in research and publication [Protocol]. Cochrane Database of Systematic Reviews 2013, Issue 2. Art. No.: MR000038. DOI: 10.1002/14651858.MR000038.

Mathews BP, Kenny MC. Mandatory reporting legislation in the United States, Canada, and Australia: a cross-jurisdictional review of key features, differences, and issues. Child Maltreatment 2008;13(1):50-63.

Mathews B, Fraser J, Walsh K, Dunne M, Kilby S, Chen L. Queensland nurses' attitudes and knowledge of the legislative duty to report child abuse and neglect: results of a state-wide survey. Journal of Law and Medicine 2008;16(2):288-304.

Mathews B, Walsh K, Rassafiani M, Butler D, Farrell A. Teachers' reporting child sexual abuse: results of a three-state study. University of New South Wales Law Journal 2009;32(3):772-813.

Mathews B. Teacher education to meet the challenges of child sexual abuse. Australian Journal of Teacher Education 2011;36(11):13-32.

McGrath P, Cappelli M, Wiseman D, Khalil N, Allan B. Teacher awareness program on child abuse: a randomized controlled trial. Child Abuse & Neglect 1987;11(1):125-32.

Norman RE, Byambaa M, De R, Butchart A, Scott J, Vos T. The long-term health consequences of child physical abuse, emotional abuse, and neglect: a systematic review and meta-analysis. PLoS Medicine 2012;9(11):e1001349.

Paolucci E, Genuis M, Violato C. A meta-analysis of the published research on the effects of child sexual abuse. Journal of Psychology 2001;135(1):17-36.

Pinheiro PS. World report on violence against children. www.unicef.org/violencestudy/reports.html (accessed 12 June 2015).

Radford L, Corral S, Bradley C, Fisher H. Trends in child maltreatment. Lancet 2012;379(9831):2048.

Reiniger A, Robison E, McHugh M. Mandated training of professionals: a means for improving reporting of suspected child abuse. Child Abuse & Neglect 1995;19(1):63-9.

Review Manager (RevMan) [Computer program]. Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Sedlak AJ, Mettenburg J, Basena M, Petta I, McPherson K, Greene A, et al. Fourth National Incidence Study of Child Abuse and Neglect (NIS4). Report to Congress, Executive Summary. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families 2010.

Starling SP, Heisler KW, Paulson JF, Youmans E. Child abuse training and knowledge: a national survey of emergency medicine, family medicine, and pediatric residents and program directors. Pediatrics 2009;123(4):e595-602.

Stoltenborgh M, Van Ijzendoorn MH, Euser EM, Bakermans-Kranenburg MJ. A global perspective on child sexual abuse: meta-analysis of prevalence around the world. Child Maltreatment 2011;16(2):79-101.

US Department of Health and Human Services, Administration for Children and Families, Administration on Children, Youth and Familes, Children's Bureau. Child Maltreatment 2009. http://archive.acf.hhs.gov/programs/cb/pubs/cm09/cm09.pdf (accessed 12 June 2015).

Walsh K, Bridgstock R, Farrell A, Rassafiani M, Schweitzer R. Case, teacher and school characteristics influencing teachers' detection and reporting of child physical abuse and neglect: results from an Australian survey. Child Abuse & Neglect 2008;32(10):983-93.

Walsh KM, Mathews BP, Rassafiani M, Farrell A, Butler DA. Understanding teachers' reporting of child sexual abuse: measurement methods matter. Children and Youth Services Review 2012;34(9):1937-46.

World Health Organization, International Society for the Prevention of Child Abuse and Neglect. Preventing child maltreatment: a guide to taking action and generating evidence. http://whqlibdoc.who.int/publications/2006/9241594365_eng.pdf (accessed 20 April 2015).

Zellman G. Child abuse reporting and failure to report among mandated reporters: prevalence, incidence, and reasons. Journal of Interpersonal Violence 1990;5(1):3-22.

## 6. REVIEW AUTHORS

**Lead review author:**

The lead author is the person who develops and co-ordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base and takes responsibility for the on-going updates of the review.


| **Name: Ben Mathews** |
|---|
| Title: Associate Professor |
| Affiliation: Queensland University of Technology |
| Address: Faculty of Law<br>Australian Centre for Health Law Research<br>GPO Box 2434 |
| City, State, Province or County: Brisbane, Queensland |
| Postal Code: 4001 |
| Country: Australia |
| Phone: +61 731 382 983 |
| Email: b.mathews@qut.edu.au |


**Co-author(s):**


- Kerryann Walsh - Faculty of Education, Queensland University of Technology, Brisbane, Australia
- Sandra Coe - Faculty of Law, Australian Centre for Health Law Research, Queensland University of Technology, Brisbane, Australia
- Maureen C Kenny - College of Education, Florida International University, Miami, Florida, USA
- Dimitrios Vagenas - Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia


## 7.ACKNOWLEDGEMENTS


We gratefully acknowledge comments received on this protocol from the external referees, and the Cochrane Developmental, Psychosocial and Learning Problems Group editorial team.

## 8.ROLES AND RESPONSIBLIITIES

BM and KW conceived and designed the project. BM and KW oversaw the project. BM wrote the text for the background sections. KW and DV wrote the text for the methodology sections. SC wrote and tested the search strategy. MK provided general advice.

## 9. SOURCES OF SUPPORT

**Internal sources**

- Australian Centre for Health Law Research, Faculty of Law, Queensland University of Technology (QUT), Australia

  BM and SC acknowledge funding support

- Faculty of Education and the Children and Youth Research Centre at Queensland University of Technology (QUT), Australia

  KW acknowledges in-kind support

## 10.DECLARATIONS OF INTEREST

Ben Mathews - none known.
Kerryann Walsh - none known.
Sandra Coe - none known.
Maureen C. Kenny - none known.
Dimitrios Vagenas - none known.

BM, KW and MK are authors of several studies included in this protocol. BM (Fraser 2010; Mathews 2008a; Mathews 2008b; Mathews 2009; Mathews 2011; Walsh 2012), KW (Fraser 2010; Goebbels 2008; Mathews 2008b; Mathews 2009; Walsh 2008; Walsh 2012), MK (Kenny 2001; Kenny 2004; Kenny 2001; Mathews 2008a). These review authors will not be involved with the assessment of eligibility, assessing risk of bias, or extracting data from any of the studies in which they are involved; this will be done by two independent review authors.