

# **The effects of Positive Youth Development interventions on substance use, violence and inequalities: systematic review of theories of change, processes and outcomes**

## ***Background***

This review will synthesise evidence on positive youth development (PYD) interventions aimed at reducing substance use and violence, and inequalities in these outcomes.

## ***Description of the problem***

Young people in the UK have among the worst health in Europe with marked inequalities across the social scale (1, 2). Health risk behaviours increase during adolescence (3) and lead to high rates of later chronic disease and other problems, and substantial economic costs (4). Child poverty is currently increasing, raising the possibility of upward trends in young people's risk behaviours, with worrying implications for future chronic disease rates and NHS costs (5). Substance use and violence are highly prevalent and damaging to young people's long-term health. Rates of youth smoking, drinking and illicit drug use, collectively termed substance use, in the UK are among the highest in Europe (6, 7). One in four 15-year-olds are regular or occasional smokers, with previous declines having ceased since 2006 and just under a third of 15-year-olds drink alcohol every week (8). One in five 15-year-olds report drug use in the past month and over 11,500 under-18s access treatment services each year (9). Youth substance use reinforces existing socioeconomic inequalities in health across the life-course: substance use is most prevalent among socially-disadvantaged young people and frequent use at a young age is strongly associated with more harmful use and chronic illness in adulthood (10-12). Aggression and violence are similarly challenging. One survey reports that by age 15-16, a quarter of young people have carried a weapon and 19% reported attacking someone with the intention to hurt them seriously (13). Violence is subject to marked health inequalities (14) and is associated with an increased risk of: physical health problems (15); engaging in health risk behaviours such as substance use (16-18); long-term emotional, behavioural and mental health problems (15, 19, 20); and self-harm and suicide (21). The economic costs associated with youth substance use and aggression are extremely high (10, 22, 23).

## ***Description of the intervention***

Interventions to reduce health risks in adolescence are potentially highly cost effective (24). There are increasing calls for adolescent health interventions to address multiple rather than single risk behaviours because such behaviours cluster together (25, 26) and because such interventions are potentially more feasible and efficient (27). PYD is one such intervention to address inter-clustered risk behaviours among young people. PYD is the dominant paradigm in youth work in the UK. The National Youth Agency (NYA), the major youth work organisation in the UK, defines such interventions as voluntary and informal educational activities aiming to bring about generalised youth development rather than merely remedying 'problem behaviours'. Such development is defined in terms of the promotion of positive skills, attitudes, relationships and identities (28). A literature review published by NYA developed a complex definition of PYD in terms of philosophy, constructs, domains and processes but similarly emphasised young people's positive attributes and competencies through structured voluntary activities (29).

Similarly, in the USA, PYD is defined in terms of its goal of developing: bonding; resilience; social, emotional, cognitive, behavioural or moral competence; self-determination; spirituality; self-efficacy; clear and positive identity; belief in the future; recognition for positive behaviour; opportunities for pro-social involvement; and/or pro-social norms (24, 40). academic, cognitive or vocational skills; confidence; connections to peers and adults; character in terms of self-control, respect and morality; and caring for others (30) Drawing on these definitions, we define PYD as involving voluntary education provided by youth workers outside of normal

school time targeting young people age 11-18 not merely the prevention of problem behaviour but addressing generalised, positive development in terms of: bonding; resilience; social, emotional, cognitive, behavioural or moral competence; self-determination; spirituality; self-efficacy; clear and positive identity; belief in the future; recognition for positive behaviour; opportunities for pro-social involvement; and/or pro-social norms. Our definition excludes PYD delivered in school time because this has been the subject of recent reviews (31, 32). It also excludes interventions delivered in custodial or probationary settings, clinical settings, employment training for school leavers or that primarily target families. PYD has the potential to reduce substance use and violence through various complex pathways. First, PYD can address some of the underlying, social determinants of these outcomes, such as disengagement from education, lack of social support and low aspirations for the future (24). Second, PYD can divert young people away from substance use and violence through engaging them in more positive forms of recreation (30). Third, PYD can promote social and emotional competences, which are an important protective factor against adolescent health-risk behaviours (24). Fourth, PYD providers can provide credible health messages and signpost health services (33).

Even in a context of public-sector cuts, there is major investment in such interventions. The UK government's Positive for Youth (34) report announced a multi-million pound investment in youth work, youth centres, the National Citizen Service and other youth volunteering projects. The most recent public health white paper (35) cited such work as a key element in promoting young people's health. The London mayor and local government across the UK are also investing millions of pounds in various PYD interventions (36). The devolved governments in Scotland and Wales also emphasise these principles and promote investment in PYD (37, 38).

However, despite this widespread investment and potential, the evidence base for the public health benefits of such interventions is unclear. While a systematic review examining non-health outcomes (39) reported benefits for self-confidence and self-esteem, school bonding, positive social behaviours, school grades and achievement test scores, the review did not systematically examine health effects. Systematic reviews of health outcomes have so far only focused on sexual health (40, 41), reporting sustained effects but with considerable unexplained variability between programmes. For example, the Children's AID Society Carrera programme reduced teenage pregnancy in some US sites but not others (33), while two evaluations of PYD interventions in the UK respectively suggested adverse and no effects on sexual health (42, 43). US researchers have argued that some youth programmes which target 'delinquent' young people and which are relatively unstructured may actually reinforce violence and anti-social behaviours via 'peer deviancy training' (44).

Others have disputed this, referring to meta-analyses of interventions addressing youth delinquency (45) which suggest that targeting and structure of sessions do not moderate effects. However, no systematic review focused on PYD interventions has examined these questions. Non-systematic review of PYD effects on violence and drug use (24, 46) have reported benefits as well as variability, but their findings must be treated with caution given that they were unsystematic and are now quite old.

### ***Rationale for current study***

This review will fill two timely and important knowledge gaps and provide important evidence to local government commissioners of youth services and public health. First, it will aim to synthesize evidence on the effectiveness and cost-effectiveness of PYD interventions in reducing substance use and violence. Second, it will aim to examine how effects vary according to the characteristics of participants (in order to assess what works for whom and estimate effects on health inequalities) and context (in order to assess what works in what

contexts and determine likely generalisability to different settings). Addressing the first gap, concerning intervention effects, is timely and important because, as described above, UK young people have among the worst health in Europe with marked inequalities across the social scale and PYD interventions are receiving significant policy attention and investment despite a lack of evidence of health benefits from systematic reviews. Addressing the second gap, concerning moderators of effects is also important given the possibility discussed above that PYD effects will vary, and given our interest in assessing the potential of PYD to reduce health inequalities.

Our synthesis of theories of change and process evaluations will inform the development of a taxonomy of interventions and of hypotheses to be tested regarding how intervention effects are moderated by characteristics of participants and context. We will then use a combination of meta-regression and qualitative comparative analysis to test these hypotheses. Thus, our synthesis will facilitate a more informed view of the likely impact of PYD on public health and health inequalities internationally and in the UK, and which approaches have the most potential for public health improvement in different settings and populations across the UK. If appropriate, our review will inform our development of a research proposal to evaluate the effectiveness of a PYD intervention in the UK. We anticipate that such an intervention would be coordinated by the National Youth Agency, which is a collaborating institution on this proposal. Our research will also make a major contribution to the rapidly developing field of "implementation science" (47) through its application of meta-regression and qualitative comparative analysis to public health interventions for young people.

### **Research aims**

To search systematically for, appraise the quality of and synthesise evidence to address the following review questions: RQ1. What theories of change inform PYD interventions delivered to young people aged 11-18 addressing substance use and violence?

RQ2. What characteristics of participants and contexts are identified as barriers and facilitators of implementation and receipt in process evaluations of PYD?

RQ3. What is the effectiveness and cost-effectiveness of PYD when compared to usual or no treatment in reducing substance use (smoking, alcohol, drugs), and violence (perpetration and victimization)?

RQ4. What characteristics of participants and contexts appear to moderate, or are necessary and sufficient for PYD effectiveness?

### **Research objectives**

(1) To conduct electronic and other searches for studies of PYD interventions by December 2013.

(2) To screen references and reports for inclusion in the review by February 2014.

(3) To extract data from and assess the quality of included studies by May 2014.

(4) To synthesise thematically theories of change of PYD interventions to produce a taxonomy and theory of change of PYD interventions by July 2014.

(5) To synthesise process evaluations of PYD interventions by September 2014.

(6) To consult with policy/practice and young people to validate the resultant taxonomy and theory of change by October 2014.

(7) To synthesise outcome and economic evaluation data and undertake meta-regression and qualitative comparative analyses by December 2014.

(8) To draw on these syntheses to draft a report addressing our review questions by February 2015.

(9) To consult with policy/practice and young people on the draft report to inform amendments and dissemination by March 2015.

(10) To submit the final report to NIHR by May 2015.

### **Research design**

Our proposal is for a multi-method systematic review of theories of change, and process, outcome and economic evaluations of PYD interventions delivered to young people age 11-18 addressing substance use (smoking, alcohol, drugs) and violence (perpetration and victimization). The review will follow existing general criteria for the good conduct and reporting of systematic reviews (e.g. Centre for Reviews and Dissemination; Preferred Reporting Items for Systematic Reviews and Meta-Analyses). The review protocol will be registered with PROSPERO International Prospective Register of Systematic Review (<http://www.crd.york.ac.uk/Prospero/>).

Our four components fit together as follows. Our review and thematic synthesis of theories of change (RQ1) will enable us to create a taxonomy of PYD interventions and theory of change of how these are implemented, and aim to achieve their effects in different contexts and subgroups. Our review and thematic synthesis of process evaluations (RQ2) will enable us to refine this theory of change to incorporate hypotheses about potential barriers and facilitators of implementation and receipt relating to characteristics of participants and contexts. Our review of outcome and economic evaluations (RQ3) will enable us to estimate the effectiveness and cost-effectiveness of PYD. Depending on the results of the taxonomy and the heterogeneity of evaluation studies found in relation to this taxonomy, this will either examine the effectiveness of PYD overall or of sub-types of PYD.

Our use of meta-regression and qualitative comparative analysis will allow us to develop hypotheses about what factors relating to participants and contexts moderate / are necessary and sufficient for intervention effects (RQ4).

### ***Size of available literature***

On 22 March 2013, we conducted a search in PubMed using the search string given in appendix 2. This was a limited search in that: (i) it searched only medical journals when it is likely that relevant studies are published in sources other than journals as well as in journals in the fields of criminology, sociology, psychology, education and public/social policy; (ii) it is not possible to use adjacency terms within PubMed; and (iii) we narrowed the search by including terms for study design which would not occur in a full search. Our search identified 2,209 references which were screened on title and abstract only. Of these, 57 studies appeared very likely to meet our inclusion criteria while 138 might possibly meet these criteria but would require screening of the full report to decide. Of the 57 studies which were deemed likely to be included, 25 were outcome evaluations, 23 process evaluations and 9 were theoretical frameworks. The preliminary search and screening confirmed the utility and applicability of our inclusion criteria. While this was by no means an exhaustive search, it does suggest that we will be able to identify a sufficiently large number of relevant studies which will enable us to answer our review questions using the methods specified.

### ***Criteria for considering studies for this review***

### *Types of participant*

We will include studies conducted where a majority of participants are age 11-18 years. We will exclude studies of populations targeted on the basis of pre-defined physical and mental health conditions but not by pre-existing risk behaviour or other forms of targeting (e.g. area-level deprivation).

### *Types of intervention*

We will include PYD interventions involving voluntary education which aim not merely to prevent problem behaviour but aim to address generalised (beyond health) and positive (beyond avoiding risk) development in terms of promoting: bonding (developing the child's relationship with a healthy adult, positive peers, school, community, or culture); resilience (strategies for adaptive coping responses to change and stress, and promoted psychological flexibility and capacity); social competence (developmentally appropriate interpersonal skills, and rehearsal strategies for practicing these skills including communication, assertiveness, refusal and resistance, conflict-resolution, and interpersonal negotiation strategies for use with peers and adults); emotional competence (identifying feelings in self or others, skills for managing emotional reactions or impulses, or skills for building the youth's self-management strategies, empathy, self-soothing, or frustration tolerance); cognitive competence (cognitive abilities, processes, or outcomes, including academic performance, logical and analytic thinking, problem-solving, decision-making, planning, goal-setting, and self-talk skills); behavioural competence (skills and reinforcement for effective verbal, non-verbal and other actions); moral competence (empathy, respect for cultural or societal rules and standards, a sense of right and wrong, or a sense of moral or social justice); self-determination (capacity for empowerment, autonomy, independent thinking, or self-advocacy, or their ability to live and grow by self-determined internal standards and values which may or may not include group values); spirituality (beliefs in a higher power, internal reflection or meditation, or supported youth in exploring a spiritual belief system, or sense of spiritual identity, meaning, or practice); self-efficacy (personal goal-setting, coping and mastery skills, or techniques to change negative self-efficacy expectancies or self-defeating cognitions); clear and positive identity (healthy identity formation and achievement in youth, including positive identification with a social or cultural sub-group that supports their healthy development of sense of self); belief in the future (belief in his or her future potential, goals, options, choices, or long range hopes and plans were classified as promoting belief in the future, including guaranteed tuition to post-secondary institutions, school-to-work linkages, future employment opportunities, or future financial incentives to encourage continued progress on a pro-social trajectory; or optimism about a healthy and productive adult life); recognition for positive behaviour (response systems for rewarding, recognizing, or reinforcing children's pro-social behaviors were classified as using recognition for positive behaviour); opportunities for pro-social involvement (activities and events in which youths could actively participate, make a positive contribution, and experience positive social exchanges); and/or pro-social norms (clear and explicit standards for behavior that minimized health risks and supported pro-social involvement). Informed by (24, 40). PYD interventions address one of these but applied to different domains (family, community, school) or more than one of these in a single domain. Our definition excludes PYD delivered in school time, or in custodial or probationary setting, clinical settings or employment training for school leavers. It also excludes interventions that target parents/carers alongside young people in order to focus on family functioning.

### *Types of outcome*

We will include studies addressing:

Substance use (smoking, alcohol and/or drug use); or Violence (perpetration and/or victimization).

Informed by existing systematic reviews focused on substance use and violence among young people (48-51), outcome measures may draw on dichotomous or continuous variables, and self-report or observational data. They may use measures of frequency (monthly, weekly or daily), the number of episodes of use or an index constructed from multiple measures. Alcohol measures may examine alcohol consumption or problem drinking. Drug outcomes may examine drugs in general or specific illicit drugs. Measures of violent and aggressive behaviour may examine the perpetration or victimization of physical violence including violent crime. Our Data Analysis section describes how we will combine measures.

#### *Types of studies*

In order to address RQ1, we will include studies describing PYD intervention theory of change. We will define theory as we did in our previous NIHR/PHR funded review of the effects of schools and school-environment interventions on health (52). Included studies may deal exclusively with theory of change or might address it alongside the reporting of empirical data. In order to address RQ2, we will include studies reporting on process evaluation of PYD intervention. This would include studies reporting on the planning, delivery, receipt or causal pathways of PYD using quantitative and/or qualitative data. These studies may report exclusively on process evaluations or report process alongside outcome or economic data. In order to address RQ3, we will include studies reporting on outcome and economic evaluations of PYD interventions. We will include experimental (randomized controlled trials) and quasi-experimental studies (employing non-randomized prospective comparison groups). Control groups will receive usual care or no treatment. Economic studies addressing RQ3 will be defined in terms of their comparison of the costs and consequences of two or more interventions or, where there is good reason to believe outcomes are similar, involve cost-minimisation analyses. In order to address RQ4, we will draw on syntheses of all of the above study types.

#### *Language*

We will only include studies published in English because these interventions have been overwhelmingly developed in English speaking countries (24, 40).

#### *Dates*

We will only search as far back as studies published in 1985 since this is when PYD interventions first began to be developed (24, 40).

#### ***Search methods for the identification of studies***

In appendix 1, we provide the search string that we have used in a preliminary search in PubMed. As explained above, this was a limited search but it will inform the development of a more sophisticated search strategy maximising sensitivity as recommended by the Cochrane Handbook for Systematic Reviews of Interventions (53).

We will also learn from previous systematic reviews focused on the effects of PYD on sexual health (40, 41) in developing our search strategy. The studies sought by this review are not likely to be reliably indexed in databases with controlled vocabularies. So we anticipate our searches involving a large number of free text terms. We will take the following essential concepts of the inclusion criteria to develop the search string: young people; and positive youth development. If these searches elicit over 30,000 hits in pilot searches on ASSIA and Medline we will consider introduce a set of terms for our outcome measures. Our searches will involve different free text and controlled vocabulary terms for each of these concepts using the Boolean operator “AND”. The concepts will be linked by the Boolean operator “OR”. The combination of these three concepts is considered specific enough to include all available studies regardless of study design. We will not restrict the searches by date, language or publication type.

### ***Electronic searches***

We will search the following databases from inception to present: ASSIA (Applied Social Science Index and Abstracts); Australian Educational Index; BiblioMap (Database of health promotion research); British Educational Index; CENTRAL (Cochrane Central Register of Controlled Trials); The Campbell Library; CINAHL; CISDOC (The Health and Safety Information Centre of the International Labour Office); Cochrane Controlled Trials Database; DARE (Database of Abstracts of Reviews of Effects); Database of Educational Research; Econlit; ERIC; Health Management Information Consortium; IBSS (International Bibliography of the Social Sciences); International Clinical Trials Registry Platform; Medline; NHS Economic Evaluation Database; OpenGrey (System for Information on Grey Literature in Europe); Proquest Dissertations and Theses; PsycInfo; Social Policy and Practice including Child Data and Social Care Online; Social Science Citation Index/Web of Knowledge; Sociological Abstracts; Dissertation Abstracts/Index to Theses; and TRoPHI (Trials Register of Promoting Health Interventions).

### ***Searching other resources***

1. We will carefully search reference lists from all studies that meet the inclusion criteria.
2. We will only hand search those journals which; i) contain studies that we include, ii) which are found only via reference checking and iii) which are not indexed on databases we have searched. We will hand search these initially for the last 5 years and if these elicit >1 new included study hand search for a further 5 years back.
3. We will search for references to relevant studies in international government reports and non-governmental organization publications via a Google search.
4. We will contact subject experts to identify unpublished or on-going research.
5. We will search all available clinical trials registers (e.g. clinicaltrials.gov) for relevant ongoing and unpublished trials.

### ***Data collection and assessment***

#### ***Selection of studies***

Search results will be downloaded to EPPI- Reviewer 4. A worksheet with the inclusion criteria operationalised into exclusion criteria and guidance notes will be prepared and piloted by two reviewers screening 100 references. Pilot screening results will be discussed by pairs of reviewers involved in screening to ensure consistency in applying the criteria. A 90% agreement rate will be required before proceeding to independent screening of the full data set. If a single reviewer cannot reach a decision regarding inclusion of a specific article, judgement for selection will be referred to a second reviewer. If both reviewers disagree and cannot reach a consensus a third reviewer will be consulted. Full reports will be obtained for those references judged as meeting our inclusion criteria or where there is insufficient information from the title and abstract to judge inclusion. A second round of screening will then occur using the same approach but based on full study reports in order to determine which studies are included in the review. We will maintain a record of the selection process for all screened material.

#### ***Data extraction and management***

Two reviewers will independently extract data from studies meeting the inclusion criteria, using a piloted data extraction form with guidance developed for this review. Where the two authors disagree, they will meet to discuss this and if possible reach a consensus. If the reviewers cannot reach consensus regarding the particulars of data extraction for a specific study, judgement will be referred to a third reviewer. Included studies will be described using the EPPI-Centre classification system for health promotion and public health research (54),

supplemented by additional codes developed for this review. For all studies where relevant, we will extract information pertaining to: basic study details (individual and organizational participant characteristics, study location, timing and duration, research questions or hypotheses); study design and methods (design, allocation, blinding, sample size, control of confounding, accounting for data clustering, data collection, attrition, analysis); intervention characteristics (including timing and duration, programme development, theoretical framework/logic model, content and activities, providers and details of any intervention offered to the control group); process evaluation of the intervention (feasibility, fidelity/quality, intensity, coverage/accessibility, acceptability, mechanism and context using an adapted version of an existing tool (55)); outcome measures at follow-up(s) (reliability of measures, effect size both overall and where available by age, sex, socio-economic status and ethnic sub-group). For economic analyses, we will extract data in order to complete a number of data tables. These will include data on the key study design elements and results for each identified study; for example, the intervention costs, price year, time horizon, base case assumptions and perspective taken regarding cost and effectiveness estimates, and estimates of cost-effectiveness such as incremental cost-effectiveness ratios and net benefit statistics. The two reviewers will independently enter data from the data extraction forms into EPPI-Reviewer 4.

Published reports may be incomplete in a wide range of ways. For example: they may not report sufficient detail about their participants for our equity analysis; they may not present information on all the outcomes that were measured (possibly resulting in outcome reporting bias); they may not provide sufficient information about the intervention for accurate characterisation; and they may not report the necessary statistical information for the calculation of effect sizes. In all cases where there is a danger of missing data affecting our analysis, we will contact authors of papers wherever possible to request additional information. If authors are not traceable or sought information is unavailable from the authors within two months of contacting them, we will record that the study information is missing on the data extraction form, and this will be captured in our risk of bias assessment of the study.

### ***Assessment of quality and risk of bias***

We will assess the quality of theoretical literature using a framework previously developed in our NIHR-funded systematic review of the health effects of schools and school –environment interventions, which assesses factors such as clarity and parsimony (52). We will assess the quality of qualitative studies using standard Critical Appraisal Skills Program and EPPI-Centre tools for qualitative studies (56). Quality tools for qualitative studies address the rigour of: sampling; data collection; data analysis; the extent to which the study findings are grounded in the data; whether the study privileges the perspectives of participants; the breadth of findings; and depth of findings. A final step in the quality assessment of qualitative studies will be to assign studies two types of ‘weight of evidence’.

First, reviewers will be asked to assign a weight (low, medium or high) to rate the reliability or trustworthiness of the findings (the extent to which the methods employed were rigorous/could minimise bias and error in the findings). Second, reviewers will also be asked to assign an additional weight (low, medium, high) to rate the usefulness of the findings for shedding light on factors relating to the review questions. Guidance will be given to reviewers to help them reach an assessment on each criterion and the final weight of evidence. For outcome evaluations, we will assess risk of bias within each included study using the tool outlined in the Cochrane Handbook for Systematic Reviews of Interventions (53). For each study, two reviewers will independently judge the likelihood of bias in seven domains: sequence generation; allocation concealment; blinding (of participants, personnel, or outcome assessors); incomplete outcome data; selective outcome reporting; and other sources of bias (e.g. recruitment bias in cluster-randomised studies); and intensity/type of comparator. Each



study will subsequently be allocated a score of 'high risk', 'low risk' or 'unclear risk' within each domain. In cases of disagreement, the reviewers will meet to establish consensus but where the two authors cannot reach consensus regarding categorisation for risk of bias for a specific study, we will refer judgement to a third reviewer. For economic evaluations, we will assess quality using the Drummond checklist (57), supplemented with the Philips checklist (58) if an evaluation contains a decision model component. The Drummond checklist asks general questions about issues such as the comprehensiveness of the descriptions regarding the interventions at hand and the quality of the effectiveness evidence underpinning the evaluation. The Philips checklist asks detailed questions regarding factors such as the rationale for the model structure and the overall quality of the modelling methodology. We will also consider adding extra questions to the Drummond checklist for issues that are particularly pertinent to PYD interventions such as costs accruing to participants' families, where we believe use of the Drummond checklist alone might miss in terms of critical appraisal (59). We will assess reporting bias according to Sterne's guidance (60). We will reduce the effect of reporting bias by including studies and not publications in order to avoid the introduction of duplicated data (i.e. two articles could represent duplicate publications of the same study). Following the Cho statement on redundant publications (61), we will attempt to detect duplicate studies and, if multiple articles report on the same study, we will extract data only once. We will prevent location bias by searching across multiple databases. We will prevent language bias by not excluding any article based on language.

### **Data analysis**

#### *RQ1 and 2: Thematic synthesis of theory and qualitative data*

Syntheses of theories of change (RQ1) and process evaluations (RQ2) will employ thematic synthesis methods (62-64) to develop and refine: a taxonomy of PYD interventions; theory of change underlying these interventions; and characteristics of participants and context acting as potential barriers and facilitators of implementation and receipt. These syntheses will not be restricted to studies judged to be of high quality but instead conclusions drawing on poorer quality evidence will be given less weight. Detailed evidence tables will be prepared to describe: the methodological quality of each study; details of the intervention examined; study site/population; and full findings. Two reviewers will read and re-read data contained within the evidence tables, apply line-by-line codes and memos to capture the content of the data, and then group and organise codes into higher-order themes. These themes will be used to generate an explanatory framework to address

#### *RQ2. RQ3: Synthesis of overall quantitative outcome and economic data*

#### *Outcome evaluations*

In order to address RQ3, we will produce pooled estimates using EPPI-Reviewer 4 and a narrative account of the effectiveness interventions. We will regard follow-up times of less than three months, three months to one year and more than one year as different outcomes. Once we know the number of studies and the extent of heterogeneity amongst the studies (as determined by a Q test), we will make a decision whether to calculate an effect size across all outcomes (i.e., combining substance use and violent behaviour outcomes), or run analyses separately for the different outcomes, or conduct a multivariate meta-analysis (65). Once that decision has been made, or each study grouping that has a sufficient number of studies to undertake meta-analysis, we will produce forest plots. A forest plot includes the point estimates and standard errors for each study, such as risk ratios for dichotomous outcomes or standardised mean differences for continuous outcomes, with the estimates weighted by a function of the sample sizes. When the results cannot be plotted, we will describe them in the 'characteristics of included studies' table, or enter the data into additional tables. We will use the chi-squared test and the  $I^2$  statistic to measure heterogeneity. The results of these statistical tests will be evaluated in accordance with the Cochrane handbook (53). If we

consider that we have unexplained statistical heterogeneity in any of our study groupings, we will investigate it further using subgroup and sensitivity analyses. When studies are found to be statistically heterogeneous, we will use a random effects model; otherwise we will use a fixed-effect model. When using the random-effects model, we will conduct a sensitivity check by using the fixed-effect model to reveal differences in results. If an indication of substantial heterogeneity is determined (e.g.  $I^2$  value of greater than 50%), that cannot be explained through analyses of variance or meta-regressions (see analysis plan for RQ4, below), then we will not produce a pooled estimate and will present a narrative summary of our findings. The narrative report will classify and present studies according to: intervention content; research design; time to outcome measurement; outcome measure; and intervention effect.

Prior to synthesis, we will check for correct analysis by cluster and report values of: intra-cluster correlation coefficients (ICC), cluster size, data for all participants or effect estimates and standard errors. Where proper account has not been taken of data clustering, we will correct for this by inflating the standard error by the square root of the design effect [63] and, for dichotomous outcomes, adjusting the numerators and denominators (51). Where ICCs are not reported we will contact authors to request this information or impute one, based on values reported in other studies. Where imputation is necessary, we will undertake sensitivity analyses to assess the impact of a range of possible values. In other instances of missing data (such as missing population information) it may not be possible to include a study in a particular analysis if, for example, it is impossible to classify the population using our equity tool.

We will use the GRADE approach as described in the Cochrane Handbook for Systematic Reviews of Interventions to present the quality of evidence and 'Summary of findings' tables. The downgrading of the quality of a body of evidence for a specific outcome will be based on five factors: limitations of study; indirectness of evidence; inconsistency of results; precision of results; and publication bias. The GRADE approach specifies four levels of quality (high, moderate, low and very low). If sufficient studies are found, we will draw funnel plots to assess the presence of possible publication bias (trial effect versus standard error). While funnel plot asymmetry may indicate publication bias, this can be misleading with a small number of studies. We will discuss possible explanations for any asymmetry in the review in light of our number of included studies.

We will undertake a sensitivity analysis to explore whether the findings of the review are robust in light of the decisions made during the review process. We will also assess the impact of risk of bias in the included studies via restricting analyses to studies deemed to be at low risk of selection bias, performance bias and attrition bias. Where meta-analysis establishes that PYD has effects on primary outcomes and where data allows, we will undertake additional exploratory meta-analyses to determine PYD effects on social outcomes to examine the plausibility that such outcomes might lie on causal pathways. Such analyses will be informed by the synthesis of process evaluation findings to avoid data dredging, and compare those findings to the results of analyses with studies of all quality levels included.

### *Economic evaluations*

Our team includes a health economist who will appraise the evidence to determine the most appropriate approach to synthesis. The overall approach to critiquing and synthesising the information on cost-effectiveness will be narrative using standard checklists (57). However, in the event that there is substantial variation across study results, particular emphasis will be placed on identifying the reasons for these differences. Cost estimates will be inflated to current prices and converted into UK currency using purchasing power parity statistics to aid the comparison of results. Where included studies report on cost effectiveness using non-monetary outcome indicators such as QALYs, we will examine this.

However, we anticipate that this will be rare and most studies will report cost-effectiveness in terms of the costs of averting our primary outcomes of substance use and violence. The main limitation of this approach is that trade-offs between different outcomes are not made explicit and results are difficult to interpret when a particular intervention is associated with 'better' outcomes on some scales, but 'poorer' scores on others. We will, if feasible, undertake supplementary decision modelling alongside reviewing to link observed trial effects to longer-term health outcomes and their associated costs. However, this may not be possible. In 2009, Hummel et al. built an economic model for the NICE public health programme to assess the cost-effectiveness of universal interventions that aimed to promote emotional and social wellbeing in secondary schools (32). While they identified a number of published longitudinal studies, they concluded that existing data-sets did not contain appropriate information to estimate these longer-term effects and emphasised the many caveats with respect to the number of assumptions made by the modelling required to estimate long-term effects. The NICE public health methods guidance clearly states that complex modelling should be avoided if it is likely to create cost-effectiveness estimates that are highly uncertain (66).

#### *RQ4: Meta-regression and qualitative comparative analysis*

If sufficient studies are found, we will employ meta-regression using Stata 12 to investigate what factors moderate intervention effects (67, 68) in order to examine RQ4. It may not be feasible to apply this method if we judge there are too many confounders, insufficient data or meta-regression is unable to account for inter-dependencies in complex interventions. Hence, we will complement meta-regression with qualitative comparative analysis, adapted for use in research synthesis (69) to assess necessary and sufficient conditions for intervention effectiveness. We will employ EPPI-Reviewer (70) and "fsQCA" software as appropriate (71) to do this. The use of initial hypotheses derived from work addressing RQ2 will protect us from 'dredging' the data for spurious statistically significant results. The required steps of 'qualitatively anchoring' outcomes in qualitative comparative analysis will ensure that changes in outcomes are meaningful and not simply statistical artefacts with little relevance for decision-making (71). These analyses will include examination of how intervention effects are moderated by characteristics of participants (for example in terms of individual socioeconomic status, sex and ethnicity) and contexts (for example in terms of area deprivation), in order to examine potential impacts on health inequalities. This will draw on existing methods involving an 'equity lens' (72) employing meta-regression to examine effects by participant sub-group. We will examine whether participant socio-economic status, sex and ethnicity, and area deprivation moderate PYD effectiveness. We should stress that meta-regression and qualitative comparative analysis will be exploratory analyses oriented towards hypothesis development rather than testing since these will draw on observational rather than experimental comparisons.

#### **Reporting**

Our aim is to provide research outputs which provide rigorous evidence on the potential effectiveness and cost-effectiveness of PYD in preventing violence and substance use, and reducing health inequalities, and to inform the future optimization of PYD. Our synthesis will enable a more informed view of the likely impact of PYD on public health and health inequalities internationally and in the UK, and which approaches have the most potential for public health improvement in different settings across the UK. If appropriate, our review will inform our development of a research proposal to evaluate the effectiveness of a PYD intervention in the UK. We anticipate that such an intervention would be coordinated by the National Youth Agency, which is a collaborating institution on this proposal.

We will produce three reports: a full technical report for NIHR; a briefing report for policy and practice audiences (particularly targeting local authority commissioners of public health and youth services); and a concise young people's report in consultation with the 'ALPHA' (Advice

Leading to Public Health Advancement) young people's group. All three reports will be published online. Stakeholders will be able to use these latter two reports as a yard-stick in assessing local services and lobbying for more effective services. The research will be launched at an event organised through the Association for Young People's Health group, which is a collaborating partner in the research. In addition, we will disseminate the research via open-access scientific journals, and via academic and policy conferences. We will undertake seminars involving representatives of the UK and devolved national government departments of health and education to present the research to discuss policy implications and next steps. We will also use stakeholder and academic networks to support dissemination as well as existing web and social media platforms based at the EPPI-Centre (Institute of Education) and DECIPHer (Cardiff University).

### ***Socioeconomic position and inequalities***

Socioeconomic status and health inequalities are central to the research. PYD interventions are generally targeted towards disadvantaged settings and populations with the aim of reducing inequalities. Our review will examine evidence for them doing so effectively. As described above, as well as examining overall effectiveness, where data allow our review will use an 'equity lens' (72), as described above. We will examine whether participant socioeconomic status, sex and ethnicity, and area deprivation moderate PYD effectiveness.

### ***Research governance and ethics***

The principal investigator is Chris Bonell, Professor of Sociology and Social Policy, Department of Childhood, Families and Health, Institute of Education, University of London who is responsible for the conduct and delivery of the work. The sponsor of the research is Steve Denton, Pro-Director of

Strategy and Organisation at the Institute of Education. The co-applicants will form an investigator committee which will meet monthly throughout the project, overseeing its conduct. These meetings will be minuted to keep a record of tasks, deadlines and responsibilities. Since the research involves no human participants and draws solely on evidence already in the public realm, there is no requirement for review by research ethics committee. While recognising that systematic reviews do not have the same potential for direct harm that primary research can have, it is essential that the team has considered relevant issues. The team will therefore follow relevant guidelines and best practice. EPPI-Centre staff follow the Social Research Association's (SRA) ethical guidelines (73) and refer also to guidance recommended by the National Coordinating Centre for Public Engagement (74). The SRA guidelines emphasise four obligations to: society; funders and employers; colleagues; and research participants. The second set of guidelines emphasise seven principles of particular relevance for work requiring public engagement: mutual respect; equality and inclusion; democratic participation; active learning; making a difference; collective action; and personal integrity.

### ***Project timetable and milestones***

- October-December 2013: Electronic and other searches for studies of PYD interventions.
- December 2013-February 2014: Screening of found references and reports for inclusion in the review.
- February-May 2014: Data extraction and quality assessment of included studies.
- May-July 2014: Thematic synthesis of theories of change of PYD interventions to produce taxonomy and theory of change.
- July-September 2014: Synthesis of process evaluations of PYD interventions.
- October 2014: Consultation with policy/practice and youth stakeholders to validate the taxonomy and theory of change.
- October-December 2014: Synthesis of outcome and economic evaluation data and

- undertake meta-regression and qualitative comparative analyses.
- December 2014-February 2015: Draft report addressing our review questions.
- March 2015: Consultation with policy/practice and youth stakeholders on the draft report to inform amendments and dissemination.
- May 2015: Submission of the final report to NIHR.

### **Expertise**

The team for this project involves investigators with recognised expertise and experience in: synthesis of theories of change, and of quantitative and qualitative evidence; evaluation of PYD and other complex interventions aiming to promote young people's health; thematic and statistical meta-analysis, meta-regression and qualitative comparative analysis, and economic evaluation. Prof. Chris Bonell (Professor of Sociology and Social Policy, Institute of Education) will direct the project, lead on searching, appraisal and qualitative synthesis, and act as a second reviewer. Prof. Bonell will provide supervision to the research officer employed to work on the project. This will involve weekly meetings with all actions minuted. Dr James Thomas (Reader in Social Policy, Institute of Education) will lead statistical meta-analyses and meta-regressions, and qualitative comparative analysis, and act as a second reviewer on some studies. Dr Adam Fletcher (Senior Lecturer in Social Science and Health, Cardiff University) will advise on synthesis of qualitative evidence, lead consultation with young people, and act as second reviewer on some studies. Prof. Rona Campbell (Professor of Public Health Research, University of Bristol) will advise on the systematic review of complex interventions. Dr Simon Murphy (Senior Research Fellow, DECIPHer, Cardiff University) will be an unfunded co-applicant, advising on stakeholder consultation, dissemination and knowledge transfer. Dr Alec Miners and Ms Claire Stansfield are not co-applicants but are nonetheless key, costed members of our team. Dr Miners is an experienced health economist who will lead our economic analyses. Ms Stansfield is an information scientist who will lead our search strategy.

### **Partner collaboration**

We have consulted with potential users of the research to gauge their interest in the work, identify additional priorities and establish whether they might sit on our stakeholder consultative body. These stakeholders include the Department of Health (Richard Sangster), Department for Education (Richard White), National Youth Agency (Alex Stutz), Action on Smoking and Health (ASH) Wales and Association for Young People's Health (John Coleman), all of whom support the work and will collaborate on it. We also consulted with the ALPHA young people's public input advisory group based in DECIPHer (Development and Evaluation of Complex Public Health Interventions for Public Health Improvement) across the universities of Cardiff, Bristol and Swansea. This consultation informed our view that synthesising evidence on PYD was a priority, and that this should include assessments of processes alongside outcomes in order to consider the acceptability of interventions as well as their potential transferability across different settings. This consultation also informed our decision to prioritise evidence of effects on substance use (smoking, alcohol and drugs) and violence (perpetration and victimization).

As with our previous evidence synthesis of the effects on health of schools and school environment interventions, funded by the NIHR Public Health Research Programme, we will consult with policy/practice and youth stakeholders in the course of the project. We will convene an advisory group of the above policy/practice stakeholders, and we will consult separately with the ALPHA young people's public input advisory group based in DECIPHer. Consultations with each of these two bodies will occur at two points: first, when we have synthesised evidence addressing theories of change and process evaluations, in order to validate and refine our theory of change; and second, during the write up of the research, to

inform how research outputs are structured and disseminated. If appropriate, our review will inform our development of a research proposal to evaluate the effectiveness of a PYD intervention in the UK. We anticipate that such an intervention would be coordinated by the National Youth Agency, a collaborating institution on this proposal.

## References

1. UNICEF. Child poverty in perspective: an overview of child well-being in rich countries. Innocenti Report Card 7. Florence: UNICEF; 2007.
2. Coleman J, Brooks F, Threadgold P. Key Data on Adolescence 2011: The Latest Information and Statistics about Young People Today London: Association for Young People's Health; 2011.
3. Viner RM, Ozer EM, Denny S, Marmot M, Resnick M, Fatusi A, et al. Adolescence and the social determinants of health. *Lancet*. 2012;379(9826):1641-52.
4. Sawyer SM, Afifi RA, Bearinger LH, Blakemore SJ, Dick B, Ezech AC, et al. Adolescence: a foundation for future health. *Lancet*. 2012;379(9826):1630-40.
5. DWP. Households Below Average Income: An Analysis Of The Income Distribution 1994/5-2010/11. London: Department for Work and Pensions; 2012.
6. Hibell B, Guttormsson U, Ahlström S, Balakireva O, Bjarnason T, Kokkevi A, et al. The 2011 ESPAD Report: Substance Use Among Students in 36 European Countries. Stockholm: The Swedish Council for Information on Alcohol and other Drugs; 2011.
7. Fuller E. Smoking, drinking and drug use among young people in England in 2011. London: National Centre for Social Research; 2012.
8. NatCen/NFER. Smoking, Drinking and Drug Use among Young People in England in 2009. London: NHS Health and Social Care Information Centre; 2010.
9. National Treatment Agency for Substance Misuse. Drug Treatment Activity in England 2006/07. London: Department of Health; 2007.
10. Donaldson L. Tackling the health of the teenage nation: Chief Medical Officer's Annual Report 2007. London: Department of Health, England; 2008.
11. Patton GC, Coffey C, Carlin JB, Degenhardt L, Lynskey M, Hall W. Cannabis use and mental health in young people: cohort study. *British Medical Journal*. 2002;325(7374):1195-8.
12. World Health Organisation. Evidence-based strategies and interventions to reduce alcohol - related harm: Global assessment of public -health problems caused by harmful use of alcohol. Geneva: World Health Organisation; 2007.
13. Beinart S, Anderson B, Lee S, Utting D. Youth at Risk?: A National Survey of Risk Factors, Protective Factors and Problem Behaviour among Young People in England, Scotland and Wales (JRF Findings 432). York: Joseph Rowntree Foundation; 2002.
14. Jansen DE, Veenstra R, Ormel J, Verhulst FC, Reijneveld SA. Early risk factors for being a bully, victim, or bully/victim in late elementary and early secondary education: the longitudinal TRAILS study. *BMC Public Health*. 2011;11((1):440.

15. Arseneault L, Walsh E, Trzesniewski K, Newcombe R, Caspi A, Moffitt T. Bullying victimization uniquely contributes to adjustment problems in young children: a nationally representative cohort study. *Pediatrics*. 2006;118(1):130-8.
16. Forero R, McLellan L, Rissel C, Bauman A. Bullying behaviour and psychosocial health among school students in New South Wales, Australia: cross sectional survey. *British Medical Journal*. 1999;319(7206):344-8.
17. Kaltiala-Heino R, Rimpelä M, Rantanen P, Rimpelä A. Bullying at school-an indicator of adolescents at risk for mental disorders. *Journal of Adolescence*. 2000;23(6):661-74.
18. Juvonen J, Graham S, Schuster MA. Bullying among young adolescents: the strong, the weak, and the troubled. *Pediatrics*. 2003;112(6):1231-7.
19. Bond L, Carlin JB, Thomas L, Rubin K, Patton G. Does bullying cause emotional problems? A retrospective study of young teenagers. *British Medical Journal*. 2001;323:480-84.
20. Hawker DS, Boulton MJ. Twenty years' research on peer victimization and psychosocial maladjustment: a meta-analytic review of cross-sectional studies. *Journal of Child Psychology and Psychiatry*. 2000;41(4):441-55.
21. Van der Wal MF, De Wit CA, Hirasing RA. Psychosocial health among young victims and offenders of direct and indirect bullying. *Pediatrics*. 2003;111:1312-7.
22. Sainsbury Centre for Mental Health. *Diversion: a better way for criminal justice and mental health*. London: Sainsbury Centre for Mental Health.; 2009.
23. Scott S, Knapp M, Henderson J, Maughan B. Financial cost of social exclusion: follow up study of antisocial children into adulthood. *BMJ*. 2001;323(7306):191.
24. Catalano RF, Berglund LM, Ryan JAM, Lonczak HS, Hawkins JD. Positive Youth Development in the United States: Research Findings on Evaluations of Positive Youth Development Programs. *Prevention and Treatment*. 2002;5(1):1-166.
25. Jackson C, Sweeting H, Haw S. Clustering of substance use and sexual risk behaviour in adolescence: analysis of two cohort studies. *British Medical Journal Open*. 2012;2(e000661).
26. Buck D. *Clustering of unhealthy behaviours over time*. London: The King's Fund; 2012.
27. Kipping RR, Campbell RM, MacArthur GJ, Gunnell DJ, M. H. Multiple risk behaviour in adolescence. *Journal of Public Health*. 2012;34(s1): i1-i2.
28. National Youth Agency. *The NYA Guide to Youth Work in England*. Leicester: National Youth Agency; 2007.
29. Schulman S, Davies T. *Evidence of the impact of the 'youth development model' on outcomes for young people - a literature review*. Leicester: National Youth Agency; 2007.
30. Roth J, Brooks-Gunn J. Promoting healthy adolescents: synthesis of youth development program evaluations. *Journal of Research on Adolescence*. 1998;8(4):423-59.
31. Durlak JA, Dymnicki AB, Taylor RD, Weissberg RP, Schellinger KB. The impact of enhancing students' social and emotional learning: a meta-analysis of school-based universal interventions. *Child Development*. 2011;82(1):405-32.

32. Hummel S, Naylor P, Chilcott J, Guillaume L, Wilkinson A, Blank L. Cost effectiveness of universal interventions which aim to promote emotional and social wellbeing in secondary schools: report for NICE Sheffield: University of Sheffield; 2009.
33. Philiber S, Kaye JW, Herrling S. The National Evaluation of the Children's Aid Society Carrera Model Program to Prevent Teen Pregnancy. New York: Philiber Research Associations; 2001.
34. Department for Education. Positive for Youth - A new approach to cross-government policy for young people aged 13 to 19. London: Stationery Office; 2011.
35. Department of Health. Healthy Lives, Healthy People. Our Strategy for Public Health in England. London: HM Government; 2010.
36. Mayor's Fund For London. Mayor's Fund for London: Annual Report 2011. London: London Mayor; 2011.
37. Scottish Government. Valuing Young People: Principles and connections to support young people achieve their potential. Edinburgh: Scottish Government.; 2009.
38. Welsh Assembly Government. The Youth Work Curriculum Statement for Wales. Cardiff: Welsh Assembly Government; 2007.
39. Durlak JA, Weissberg RP. The impact of after-school programs that promote personal and social skills. Chicago, IL: Collaborative for Academic, Social, and Emotional Learning; 2007.
40. Gavin LE, Catalano RF, David-Ferdon C, Gloppen KM, Markham CM. A review of positive youth development programs that promote adolescent sexual and reproductive health. *J Adolesc Health*. 2010;46(3 Suppl):S75-91.
41. Harden A, Brunton G, Fletcher A, Oakley A, Burchett H, Backhans M. Young people, pregnancy and social exclusion: A systematic synthesis of research evidence to identify effective, appropriate and promising approaches for prevention and support. London: Institute of Education; 2006.
42. Wiggins M, Bonell C, Sawtell M, Austerberry H, Burchett H, Allen E, et al. Health outcomes of youth development programme in England: prospective matched comparison study. *BMJ*. 2009;339
43. Bonell C, Maisey R, Speight S, Purdon S, Keogh P, Wollny I, et al. Randomized controlled trial of 'Teens and Toddlers': a teenage pregnancy prevention intervention combining youth development and voluntary service in a nursery. *Journal of Adolescence* (in press).
44. Dishion TJ, McCord J, Poulin F. When interventions harm. *American Psychologist*. 1999;54(9):755-64.
45. Weiss B, Caron A, Ball S, Tapp J, Johnson M, Weisz JR. Iatrogenic effects of group treatment for anti-social youth. *Journal of Consulting and Clinical Psychology*. 2005;73:1036-44.
46. Roth JL, Brooks-Gunn J. Youth development programs: risk, prevention and policy. *Journal of Adolescent Health*. 2003;32:170-82.



47. Gaglio B, Glasgow RE. Evaluation approaches for dissemination and implementation research. In: Brownson RC, Colditz G, Proctor E, editors. *Dissemination and Implementation Research in Health: Translating Science to Practice*: Oxford University Press; 2012.
48. Faggiano F, Vigna-Taglianti FD, Versino E, Zambon A, Borraccino A, Lemma P. School-based prevention for illicit drugs use. *Cochrane Database of Systematic Reviews Issue 2*, Art No CD003020 DOI@: 10.1002/14651858 CD 003020 pub2. 2005.
49. Foxcroft DR, Ireland D, Lowe G, Breen R. Primary prevention for alcohol misuse in young people. *Cochrane Database of Systematic Reviews Issue 2*, Art No CD003020 DOI@: 10.1002/14651858 CD 003020 pub2. 2002;3:Art. No. CD003024.
50. Hahn R, Fuqua-Whitley D, Wethington H, Lowy J, Crosby A, Fullilove M. Effectiveness of universal school-based programs to prevent violent and aggressive behavior: a systematic review. *Am J Prev Med.* 2007;33(2 Suppl):S114-S29.
51. Thomas R, Perera R. School-based programmes for preventing smoking (Cochrane Review). *Cochrane Database of Systematic Reviews 3*: CD001293. DOI:10.1002/14651858.CD001293. 2006.
52. Bonell C, Jamal F, Harden A, Wells H, Parry W, Fletcher A, et al. Systematic review of the effects of schools and school environment interventions on health: evidence mapping and synthesis. *Public Health Research.* 2013;1(1).
53. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 [updated March 2011]*. Oxford: The Cochrane Collaboration; 2011.
54. Peersman G, Oliver S, Oakley A. *EPPI-Center Review Guidelines: Data Collection for the EPIC Database*. London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London; 1997.
55. Egan M, Bambra C, Petticrew M, Whitehead M. Reviewing evidence on complex social interventions: development and testing of a new tool for appraising implementation. *Journal of Epidemiology and Community Health.* 2009;63:4-11.
56. Shepherd J, Harden A, Rees R, Brunton G, Garcia J, Oliver S, et al. *Young People and Healthy Eating: A systematic review of barriers and facilitators*. London: EPPI-Centre, Social Science Research Unit; 2001.
57. Drummond MF, Sculpher MJ, Torrance GW, O'Brien B, Stoddart GL. *Methods for the economic evaluation of health care programmes*. 3rd ed. Oxford: Oxford University Press; 2005.
58. Philips Z, Ginnelly L, Sculpher M, Claxton K, Golder S, Riemsma R, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. *Health Technology Assessment* 2004;8(36).
59. Miners AH, Yao G, Raftery J, Taylor RS. Economic evaluations of calcineurin inhibitors in renal transplantation: A literature review. *Pharmacoeconomics.* 2007;25(11):935-47.
60. Sterne J, Egger M, Moher D. Addressing reporting biases. In: Higgins JPT, Green S, editors. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester: Wiley; 2008. p. 297-334.

61. Cho BK, Rosenfeldt F, Turina MI, Karp RB, Ferguson TB, Bodnar E, et al. Joint statement on redundant (duplicate) publication by the Editors of the undersigned cardiothoracic journals. *Ann Thorac Surg.* 2000;69:663.
62. Arai L, Roen K, Roberts H, Popay J. It might work in Oklahoma but will it work in Oakhampton? Context and implementation in the effectiveness literature on domestic smoke detectors. *Inj Prev.* 2005;11:148-51.
63. Noyes J, Popay J, Garner P. What can qualitative research contribute to a Cochrane systematic review of DOT for promoting adherence to tuberculosis treatment? *Qualitative Research and Systematic Reviews workshop*; 2005 28-29 June; Continuing Professional Development Centre, University of Oxford. 2005.
64. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology.* 2008;8:45.
65. Kalaian HA, Raudenbush SW. A multivariate mixed linear model for meta-analysis. *Psychological Methods.* 1996;1:227-35.
66. NICE. *NICE Guide to the Methods of Technology Appraisal.* London: National Institute of Health and Clinical Excellence; 2008.
67. Thompson S, Sharp S. Explaining heterogeneity in meta-analysis: a comparison of methods. *Statistics in Medicine.* 1999;18(2):693-708.
68. Borenstein M, Hedges L, Higgins J, Rothstein H. *Introduction to Meta-Analysis.* Chichester: John Wiley and Sons Ltd; 2009.
69. Ragin CC. *Redesigning Social Inquiry: Fuzzy Sets and Beyond.* Chicago, IL: University of Chicago Press; 2008.
70. Thomas J, Brunton J, Graziosi S. *EPPI-Reviewer 4.0: software for research synthesis.* EPPI-Centre software. London: Social Science Research Unit, Institute of Education; 2010 Contract No.: Document Number].
71. Ragin CC, Drass KA, Davey S. *Fuzzy-Set / Qualitative Comparative Analysis 2.0.* Tucson AZ: Department of Sociology, University of Arizona; 2006.
72. Kavanagh J, Oliver S, Lorenc T, Caird J, Tucker H, Greaves A, et al. School-based cognitive behavioural interventions: a systematic review of effects and inequalities. *Health Sociology Review* 2009;18:61-78.
73. Social Research Association. *Ethical Guidelines.* London: Social Research Association; 2003.
74. Centre for Social Justice and Community Action and National Co-ordinating Centre for Public Engagement. *Community-based participatory research: a guide to ethical principles and practice, national coordinating centre for public engagement.* Bristol: National Coordinating Centre for Public Engagement; 2012.