

# Handling Complexity in Evidence from systematic reviews and meta-analyses of Public Health Interventions (CEPHI project) – Protocol for a mixed methods and mixed synthesis project

**Investigators:** Dylan Kneale\*<sup>1</sup>, Alison O'Mara-Eves\*<sup>1</sup>, Katy Sutcliffe<sup>1</sup>, Sandy Oliver<sup>1</sup>, James Thomas<sup>1</sup> with Niccola Hutchinson-Pascal<sup>2</sup> and Co-Production Collective<sup>2</sup>

\*joint leads

<sup>1</sup>EPPI-Centre, Social Science Research Unit, Social Research Institute, University College London

<sup>2</sup>Co-Production Collective, UCL Culture, University College London

## Abstract

**Aim:** To explore the feasibility of adjunct methods that can be used to explore context in evidence synthesis and to develop guidelines around their use.

**Background:** Many public health challenges appear intractable, with substantial investments made into interventions that have little impact. Childhood obesity is one such example, where levels of childhood obesity have broadly stabilised, but with no apparent declines in prevalence despite substantial investment. The development of evidence-based strategies has been widely recommended. Despite their high regard, the ability to utilise evidence from meta-analyses and systematic reviews is hampered by the lack of connection between the contexts in which interventions were conducted and the context in which the evidence is to be applied. In this proposal we seek to develop methods of exploring and enhancing the generalisability of meta-analysis using secondary data analysis and additional synthesis.

**Design:** Mixed methods design involving stakeholder engagement, designing systems-based logic models, secondary data analysis, and the **secondary synthesis** of existing systematic review evidence.

**Methods:** The proposed methodological research uses childhood obesity as a case example and is spread over three work packages (WPs):

WP1: We will co-produce a systems-based logic model with stakeholders that helps to identify contextual features of interest and how they interact with a potential intervention. This is an output in its own right, but will also be used to guide later stages.

WP2: This work package explores the utility of two approaches in examining the generalisability of evidence, drawing on WP1. First we will assess how using existing observational data and employing statistical methods (namely reweighting of effect sizes and latent class analysis) in novel ways can help to create an overall measure of effect from meta-analysis that is more applicable to a defined

population. Next we will explore the utility of Qualitative Comparative Analysis (QCA) in examining the influence of context. Drawing on set-theory, QCA will be used to examine configurations of different contextual features that align with more successful interventions.

WP3: Given that the primary motivation of this project is that end-users are not utilising review evidence because of its disconnect with their particular local circumstances, we need to explore the utility of the proposed enhancements to meta-analytic evidence. We will conduct two workshops and further meetings to evaluate the utility of the approaches trialled above and develop guidance around their usage.

**Discussion:** The impact of this work will (i) identify and evaluate new methods for exploring the generalisability of meta-analytic evidence; and (ii) allow systematic reviewers and/or knowledge translators to replicate the approaches taken through issuing guidance and case studies.

**Keywords:** context, generalisability, evidence, systematic review, secondary data analysis, QCA, meta-analysis, recalibration, logic model.

## **Background and Scientific Rationale**

Many public health challenges appear intractable in nature, with substantial investments made into interventions that appear to make little impact on a population level. Childhood obesity is one such example, where levels of childhood obesity have broadly stabilised, but with no apparent declines in prevalence despite substantial investment (van Jaarsveld and Gulliford, 2015, NHS Digital, 2020).

The development of evidence-based strategies has been widely recommended as a means of improving public health outcomes (Brownson et al., 2009). Evidence from systematic reviews and meta-analyses is considered to represent one of the most robust sources available for public health decision-making, particularly in assessing the effectiveness of interventions (Berlin and Golub, 2014). Despite their high regard, the opportunity to utilise evidence from meta-analyses and systematic reviews is hampered by the lack of connection between the contexts in which interventions were evaluated and the context in which the evidence is to be applied; this may be a particular issue where addressing health inequalities is a concern.

Researchers examining the use of evidence in decision-making have emphasised the importance of the salience of evidence in determining its perceived usefulness among decision-makers (Oliver et al., 2018). In the current landscape of public health decision-making in England, this has often meant prioritising evidence that is generated locally over evidence that is generated elsewhere (Kneale et al., 2018, Kneale et al., 2019a). The extent to which evidence that is generated in various specified settings can be applied in another defined setting reflects its contextual generalisability, and in the absence of evidence that is generalisable to local contexts, public health decision-makers typically base their decisions on smaller local evaluation studies that are not always methodologically robust, and on anecdotal input (Kneale et al., 2017). This means that the systematic process of evidence synthesis is undermined by idiosyncratic patterns of evidence use, further compounding issues around wastage in research (Ioannidis et al., 2014).

The challenge of accounting for context is much more acute in the case of public health decision-making compared to clinical decision-making. Here, contextual factors influence the design and delivery of interventions substantially in terms of governance structures, delivery bodies, epidemiological factors and populations at risk, as well as numerous other factors – particularly population characteristics – that can influence outcomes including the acceptability, reach and adherence to interventions. Where an intervention works for one population or setting, there is no guarantee that it will work for others, which has implications for rolling out interventions to reduce health inequalities that have not been evaluated with the target population. Also, fundamentally,

naturally occurring conditions of 'usual care' also differ substantially between settings. It is this form of 'context', and accounting for its influence, that is of interest in this research as we seek to develop methods of exploring and enhancing the generalisability of meta-analysis using secondary data analysis and further synthesis.

### How is contextual generalisability currently treated in systematic reviews?

In the case of meta-analysis, generalisability refers to the extent to which a meta-analytic finding can be directly utilised to answer the 'local' question at hand, ideally supported by a judgment about the degree of confidence for this specific 'local' utilization (Evans et al., 2013). This involves understanding the extent to which a single aggregate 'effect', derived by combining the results of several independent studies – each of varying internal and external validity – can be used to inform decision-making in a specific situation. Generalising meta-analytic findings to a specific situation is dependent on having a well-defined target (inference) population to which a researcher intends to extrapolate (O'Muircheartaigh and Hedges, 2014, Hedges, 2013, Glass, 2000). However, systematic reviews assemble evidence from all eligible studies, which will vary from the inference population with respect to a range of contextual factors, thus limiting the applicability of review evidence. Particular concern has been raised regarding the under-representation or under-reporting of evidence relating to disadvantaged groups, which undermines the potential of systematic reviews to inform approaches to health inequalities (e.g. incentives for obesity prevention (Paul-Ebhohimhen and Avenell, 2008)).

The generalisability of meta-analytic findings is also highly related to the external validity of the included studies. Most meta-analysts currently consider the external validity of primary studies individually, using an array of different checklists and frameworks, composed of items that may have little methodological justification (Ahmad et al., 2010). Similarly, there may be a wide gulf between the factors considered important in assessing generalisability by the systematic reviewer, and those features perceived as important among local stakeholders, including decision-makers and the public. Furthermore, there is a need to recognise generalisability as a multidimensional construct encompassing both the applicability of evidence (reflecting whether an intervention is feasible) and its transferability (whether the intervention would have the same sort of impact, reflective of several dimensions including the interplay between intervention and context) (Wang et al., 2006).

Having attempted to assess the external validity of the individual studies, the meta-analyst then needs to consider generalisability at the review level. Many meta-analysts attempt to assess generalisability statistically, by trying to identify and explain any heterogeneity (i.e., variation between studies) through sub-group or regression analyses. Decisions regarding these analyses, however, are just as likely to be data-driven as they are to be driven by concerns about applicability or transferability. Additionally, they often do not consider multiple factors that may aid in the assessment of generalisability simultaneously, and are conducted without consideration of a specific situation in which the evidence is intended to be applied. As such, current approaches to the assessment of generalisability may not focus on the population or contextual factors relevant to public health decision-making and therefore fail to consider fully the applicability of findings. These current approaches therefore rarely lead to recommendations about the appropriateness of applying review evidence to different settings, which impedes the utility of review evidence to local situations where its application may be most instrumental (Kneale et al., 2018). Moreover, exploring contextual generalisability may address some of the current limitations of many systematic reviews in addressing health inequalities.

### How might generalisability be better addressed and what is needed?

The near exclusive current focus on establishing the internal validity of studies has led many systematic reviewers to fail to consider the applicability and transferability of their findings to differing contexts, even those working in public health topics where issues of generalisability could be

considered particularly challenging (Ahmad et al., 2010), and consequently, methods to enhance the generalisability of evidence are underdeveloped. Related to this, there have been calls for better methods to examine health inequalities in systematic reviews, as these are often not reported or not even considered in the evaluation literature or by systematic reviewers (Maden, 2016). To date, few strategies or methods have been proposed that explicitly seek to utilise the wealth of routinely collected data, longitudinal cohort and panel data, and detailed cross-sectional survey data collected in the UK to investigate issues around generalisability (referred to as observational data from here on in). These data could be used to: determine the representativeness of the meta-analytic sample relative to the real-world population, especially where health inequalities are of interest; identify important contextual factors for analysis in the review; understand similarities between 'control' conditions in studies and in secondary data; explore contextual factors not available in the meta-analytic dataset; or be used to check the findings of a meta-analysis. The relatively small number of studies that have examined both meta-analysis and primary datasets reveal that substantially different results can be obtained and that clear advantages exist in refining and improving analytical strategies through contrasting both sets of data (e.g., (Marsh et al., 2009)).

As a research team we have been considering issues in the generalisability of meta-analytic and systematic review evidence (Kneale et al., 2019b, Oliver et al., 2015, Oliver et al., 2018, Burchett et al., 2018); this proposal builds on ideas we have proposed but not yet fully tested and evaluated (Kneale et al., 2019b) and substantive knowledge and networks that we have developed in this area (Sutcliffe et al., 2017). In summary, this proposal aims to develop consensus around methods to explore and enhance the generalisability of meta-analytic evidence, with integrated co-production throughout.

## The study

### Research Objectives and research questions

This protocol has three main objectives spread over three work packages (WPs). Using childhood obesity as a case example we will: (1) identify – in collaboration with stakeholders – the study features that should be examined when considering the generalisability of evidence; (2) explore the contribution of adjunct analyses in exploring and enhancing the generalisability of meta-analytic evidence; and (3) produce a deployable approach to making review evidence more relevant to local usage. The main methodological innovations involve the reanalysis of existing meta-analytic evidence, in combination with adjunct secondary data analysis and additional synthesis to assess and potentially enhance the generalisability of the evidence.

1. **WP1: Assessing contextual generalisability in a defined setting:** Can local knowledge of contextual features (in terms of people, intervention, usual care conditions, or other features) be harnessed through co-production of a logic model, and be applied in innovative generalisability analyses?
2. **WP2: Adjusting for generalisability and examining the influence of context:** Which methods can assist in *statistically* adjusting for the generalisability of a completed meta-analysis? How might Qualitative Comparative Analysis (QCA) assist in identifying particular contextual features that trigger more successful interventions, or hinder positive effects? Do the methods offer a robust way to consider health inequalities in the systematic review?
3. **WP3: Evaluating and disseminating:** How should the methods from WP2 be used in review production and decision-making? How are the findings viewed by stakeholders?

## Research Plan and Methods Organised Across Work Packages

We will apply an equity lens throughout all work packages in this project. Obesity/overweight among children and young people are highly influenced by the social determinants of health. We anticipate that social determinants, and the factors that determine how they are distributed and experienced and generate inequalities, will be strongly represented organically through the work conducted in WP1; however, we will also draw on the PROGRESS-Plus framework, and ask advisory group and workshop participants to consider the impacts of these characteristics. We will also seek representation of a breadth of participants and perspectives in the workshops and advisory group to ensure the research takes into account how different groups have different needs, barriers and facilitators to take part in or benefit from interventions. In addition, WP2 and WP3 will focus on considering different elements of generalisability including whether interventions are likely to benefit geographic areas and social groups equally. The following sections detail the three work packages.

### WP1: Assessing contextual generalisability

*Summary: Contextual generalisability refers to the extent to which the key factors represented in the review match those in the target population or setting, such as the overlap in characteristics of the population, intervention, usual care, or other features. Without understanding this, review findings represent an average of the contexts evaluated, which may not fit the context/s in which the findings are to be applied. The features of an area that determine how well an intervention for childhood obesity may 'fit' is a form of local knowledge held by stakeholders that we will incorporate into later stages, using a systems-based logic model as a vehicle for engagement and co-production. A systems-based logic model sets out to theorise aspects of complexity around relationships between intervention and broader context and how these interact.*

#### **Advisory group to guide the scope and direction of the work and advisory group meeting #1:**

This work package begins with the recruitment of an advisory group reflecting different sets of expertise and perspectives (e.g. teachers, parents, citizens, public health practitioners, clinicians etc.). This will be facilitated by Co-Production Collective who will help to identify members and run a first advisory group session to inform the group of the purpose of the study, provide an opportunity for members to get to know each other, and help clarify expectations for the workshops and determine how they should be organised. The Co-Production Collective is a community of researchers, patients, carers, practitioners, students and anyone else interested in co-production (in health or more generally), with experience in co-producing research and other knowledge. The advisory group will number approximately 8 people who will meet throughout the research.

This work is based on the re-analysis of existing systematic review evidence on the prevention of obesity/overweight among children and young people and/or the management of weight among overweight/obese children. A starting point will be the identification of an existing meta-analysis requiring further analysis to explore the salience for action in a local area. We will work with the advisory group in the first meeting to identify 1-2 suitable meta-analyses that can act as case studies. Potential target meta-analyses of interest include those within a recently updated Cochrane review on the prevention of child obesity (153 RCTs (Brown et al., 2019)) as well as a smaller review that included 38 studies (Ho et al., 2012).

**Stakeholder workshops to identify important features of local obesity systems:** Two workshops with stakeholders will be planned with Co-Production Collective and will be organised in two parts, with a first part allowing for open discussion, and a second part developing a consensus on emergent themes and how they should be organised. Each workshop will include 10-20 people so we can draw on the expertise of a range of different stakeholders (e.g. teachers, parents, citizens, public health practitioners, clinicians etc.). Part of this work will be undertaken through an online workshop if needed. As is recommended, the initial two workshops will be scheduled to take 'place at key decision points of setting the scope and the scale' (with further workshops planned around the interpretation of the findings in WP3) (Oliver et al., 2018). We will utilise the TRANSFER tool conversation guide



(Munthe-Kaas et al., 2020) to help structure workshop conversations and support participants considering aspects of generalisability.

The object of the first two workshops is to develop a systems-based logic model on child obesity/overweight that identifies important features of local obesity systems. A systems-based logic model depicts the system as 'the interaction between the participants, the intervention, and the context takes place' (Rehfuess et al., 2017, p15). Interventions for the prevention of child obesity and weight management are inevitably complex and sensitive to the contexts in which they are conducted. An intervention that is effective in one type of setting may be ineffective (or even harmful) in another (Greene et al., 2017). A useful systems-based logic model would depict the types of population characteristics, contextual, and intervention features and the combinations of these features that could influence the effectiveness of interventions. A systems-based logic model will provide a graphical way of engaging stakeholders to consider features in local systems (e.g., existing health infrastructure or fast-food outlets) that could raise or reduce the risk of obesity/overweight and facilitate or hinder the effectiveness of an intervention. The research team recently developed a systems-based logic model with stakeholder input (Kneale et al., 2020), and aim to implement a fully co-produced design in the creation of this research. The model will be used to structure the analysis in the next work packages.

**Advisory group meeting #2:** We will reconvene the advisory group again after the first two workshops in order to discuss the model and the implications for the next work packages.

*Key output and contribution:* a systems-based logic model that depicts which factors are viewed by stakeholders as important local influencers of child obesity and which may influence generalisability.

## WP2: Adjusting for generalisability and examining the influence of context

*Summary: We will explore the utility of two approaches in examining the generalisability of evidence. First, we will assess how using existing observational data and employing statistical methods (reweighting of effect sizes and latent class analysis) in novel ways can help to create an overall measure of effect from meta-analysis that is more applicable to a defined population. That is, we will explore methods for adjusting (recalibrating) the results of meta-analyses to better match the circumstances of the target context. Next, we will explore the utility of Qualitative Comparative Analysis (QCA) in examining the influence of context. QCA has been used to examine configurations of intervention components that align with successful interventions, although the use of the approach to understand how different contextual features trigger success has not been explored systematically.*

The methods described below aim to examine the degree to which information about naturally occurring phenomena/populations, can be used to (i) 'recalibrate' the effect size estimate to approximate to inference conditions, and (ii) identify discrete sets of studies based on contextual factors and discrete sets of studies based on intervention effectiveness, and measure the overlap between these sets. This builds on other work exploring the degree to which the findings from one geography are applicable to another (Hedges, 2013). We will examine two distinct approaches.

**The first approach we will trial is to analyse existing data about a given context or group using existing secondary datasets and explore the extent to which the evidence within the studies contained in a meta-analysis is similar or different to this context.** We will explore contextual information about studies (as extracted in a systematic review/meta-analysis and through incorporating other sources of secondary data as necessary) and measure their similarity to an inference population. By incorporating evidence beyond that reported in the primary studies of the review, we hope to overcome some of the limitations experienced in many systematic reviews exploring health inequalities: that the data reported in primary studies are not broken down by population groups, or that disadvantaged populations are simply not examined in the primary evaluation literature. We will prioritise exploration of features identified by stakeholders in the systems-based logic model as being important to consider in generalisability assessments. Secondary Data Analyses (SDA) of existing datasets (e.g., British Birth Cohorts, Census data) will be used to assess these characteristics for specific inference populations.

We will then use multivariate methods to develop measures of how closely studies in the meta-analysis resemble the observed characteristics of the inference population (as identified in the SDA). A pilot undertaken of this approach (Kneale et al., 2019b) involved recalibrating the effect estimate of community coalitions in reducing health disparities for ethnic minorities. Data common in all primary studies on the populations involved were extracted and were also calculated for three Local Authorities using census data. A dissimilarity matrix was created, with the Euclidean distance between study sites and three Local Authorities (separately) measured. We then progressed to include the dissimilarity matrix values within study weights in estimating the pooled effect size in the meta-analysis (we termed this approach 'recalibration'). The example helped to illuminate future avenues of enquiry and this part of the work package is intended to advance these early explorations around the choice of data used, the methods employed for measuring dissimilarity, as well as the weighting strategy for accounting for generalisability in the meta-analysis.

We will also use the model derived in WP1 to implement a form of taxonomic driven meta-analyses (Hedges et al., 2020) through identifying subgroups of studies that feature similar contextual characteristics and use these taxonomies as the basis of subgroup analyses. Here, we will also explore the utility of latent class analysis (LCA), a multivariate technique to identify groups of cases or studies based on categorical variables, to derive study-level latent classes. After removing the inference population data, the resulting latent classes will be used as the basis of subgroup analysis in a (random-effects) meta-analysis, assigning each study to a class based on their highest predicted probability. The resulting classes will reflect the degree of similarity based on observed characteristics of study populations to an inference population.

Alternatively, rather than using the full set of classes, we will also model the probability of being in the same class as the inference population as a covariate within meta-regression models, or as the basis for weighting models. This approach could overlook the variance associated with aggregated data and estimates (both of the inference data and within the study data; although we could explore ways of using this information within LCA classes or within meta-analytic weights). LCA analyses and meta-analyses will be carried out using STATA, with potential further meta-analyses carried out in R. Other challenges and potential limitations will be in the choice of variables, their influence on the outcome, and distinctiveness in defining the inference population.

**The second approach is to explore the use of Qualitative Comparative Analysis (QCA) in examining the influence of context.** QCA is increasingly employed as a solution to the challenge of analysing data containing a small number of cases, each with an extensive array of factors that may trigger a given outcome (Ragin, 2008). This "small N-many variables" challenge is similar to that often faced by systematic reviewers, and Thomas and colleagues provide one of the first examples where QCA was utilised within a systematic review to understand configurations of intervention components that were aligned with "successful" (most effective) interventions (Thomas et al., 2014). The goals of QCA have been described as integrating the best features of the case-oriented approach, involving developing an in-depth knowledge of individual studies, with the best features of a 'variable-oriented' approach, where the focus is on comparing studies and identifying cross-case patterns in the data (Rihoux, 2009).

QCA is being used within systematic reviews both to further understand the results of meta-analyses (for example (Sutcliffe et al., 2016, Brunton et al., 2015, Brunton et al., 2014)), to develop theories to test within meta-analyses (for example (Harris et al., 2019)), and occasionally as a synthesis method in its own right (Sutcliffe et al., 2019). It has been used by members of the project team in the context of a review on health inequalities, which highlighted the potential for further application (Brunton et al., 2014). Two principles that can mark interventions as being complex, which are directly related to the conduct of QCA, include equifinality and conjunctural causation. Equifinality refers to the principle that different pathways or combinations of components can lead to the same outcome. This principle of equifinality can be interrogated further, drawing on our understanding of complexity, so that we could theorise where differences in contexts could support distinct pathways towards effective interventions.

A second principle, conjunctural causation, sees that particular intervention components or contextual circumstances may only trigger a successful outcome in the presence of another component or contextual factor. We can consider that observation of the outcome is dependent on jointly observing the presence or absence of two or more factors. For example, in a review of adult weight management interventions conducted by investigators included in this proposal (Melendez-Torres et al., 2018), the importance of establishing a supportive relationship with the intervention provider in triggering a successful intervention was identified through QCA. However, this component did not work in isolation, but only as part of a configuration of other conditions (conjunctural causality).

Currently, QCA is often used to explore intervention processes and components within an intervention, but less attention has been paid to the potential of QCA as a means of understanding how different contextual factors influence the effectiveness of interventions. QCA is reliant on the development of a theory to help to guide the selection of conditions (akin to variables) to prioritise within models, and we will draw on the logic model produced in Work Package 1 as our main theoretical framework. However, a QCA approach also encourages further engagement with the studies, and we intend to also draw on Intervention Component Analysis (ICA) (Sutcliffe et al., 2015) as a way of identifying further critical contextual features of interventions based on trialists formal and informal observations and using this information within QCA models.

The use of QCA will allow us to develop a deep understanding about how different features of context and populations (or the absence of these features) could influence the effectiveness of the intervention and develop our understanding of the generalisability of the intervention. The method is based on set theory, and the application intended here will draw on effect sizes and precision as a way of determining allocation of studies into an effective set, ineffective set, or as being partially in/effective. QCA usually does not incorporate estimates of precision of the effect size (e.g. confidence intervals) in determining if interventions are un/successful; a further methodological innovation of this work is to examine how precision can be incorporated in identifying successful interventions. The application of QCA to an existing meta-analysis as a means of exploring heterogeneity is an approach first developed by members of the team (Thomas et al., 2014) and applied in subsequent reviews (Harris et al., 2019, Sutcliffe et al., 2017, Boulton et al., 2020).

**Advisory group meeting #3:** We will reconvene the advisory group again during this work package to keep the group updated on progress and consider whether the new methods are methodologically and theoretically sound.

**Challenges to explore:** The aim of this work package is to explore methods to understand the generalisability of meta-analysis to a defined population. The methods outlined share a number of challenges; these include (but are not limited to): the suitability of the current review literature; subjectivity in the choice of inference population and the extent to which secondary data support inferences to the target population; choice of variables (or conditions in QCA) used in analyses (although will be guided by WP1); harmonisation of measures across studies and populations; and the treatment of uncertainty and heterogeneity in estimates including weighting strategies in meta-analysis. A large portion of the work will involve investigating these challenges and evaluating the feasibility/potential of these avenues.

*Key outputs: Methodological guidance and the results of a case study on methods that may be suitable for using SDA to recalibrate effect sizes derived from a meta-analysis and for using QCA to examine the influence of context.*

### WP3: Evaluating and disseminating the developments

*Summary: Two further workshops with stakeholders will be planned with Co-Production Collective. They will evaluate the most promising approaches developed, the content of which will be used to form guidance on their usage. Many of the outputs are technical and primarily targeted at the academic community. However, given that the primary motivation of this project is that end-users are*



*not utilising review evidence because of its disconnect with their particular local circumstances, we need to explore the utility of the proposed enhancements to meta-analytic evidence with stakeholders.*

The different approaches trialled will be based on differing epistemologies and will generate different types of findings. Understanding how these findings are interpreted and which types of findings are useful for decision-making is the object of this final work package, and we will use two workshops and a final advisory group meeting to help to facilitate this.

We will conduct two workshops (potentially held online depending on the progress of the COVID-19 pandemic). Attendees will be a mixture of academic experts in systematic reviews and meta-analysis and—critically—public health professionals, decision-makers, teachers, citizens and parents. Participants may overlap with the first workshops in WP1, but not necessarily—particularly because stakeholders at this stage will additionally include research methodologists. The aim of the workshops will be to develop consensus around the most promising approaches and to understand how the findings from different approaches are understood. The workshops will use a mixture of presentations and discussion, with an iterative process of placing draft guidance and recommendations before attendees for critique. All participants will be contacted before the workshop and their initial responses to a number of open-ended questions will be summarised in the presentations. Importantly, we will ask public health decision-makers to rate the extent to which they view the added value in using meta-analytic evidence in their decision-making practices following the implementation of our methods.

**Advisory group meeting #4:** As data will be collected during both workshops, further synthesis will be undertaken, particularly with regards to refining the approach for assessing generalisability, with final results disseminated to attendees and a wider audience. A final advisory group meeting will be held as part of this work package to help the research team to interpret workshop/survey findings and to develop guidelines. A focus of this final advisory group meeting will be to consider how the methods outlined could be scaled up. The discussions from the workshop and the interpretation of the advisory group will form the basis of guidance that will be communicated to relevant audiences.

*Key outputs: non-academic co-produced briefing papers and a methodological guidance document that are co-developed with the input and consensus of workshop attendees. This will be supported by an interactive webinar to enable deployment of the methods.*

### Summary of key outputs

- A systems-based logic model of local influencers on child obesity, developed through stakeholder engagement. (WP1)
- Methodological guidance and the results of a case study on methods for using SDA to recalibrate effect sizes derived from a meta-analysis, targeted at the academic community. (WP2)
- Methodological guidance and the results of a case study on using QCA to examine the influence of context, targeted at the academic community. (WP2)
- Co-produced briefing papers and a methodological guidance document, targeted at non-academic stakeholders—particularly public health decision-makers. (WP3)
- An interactive webinar on the methods. (WP3)

### Ethical considerations

This research will be conducted following the Economic and Social Research Council's research ethics framework. Ethical approval will be sought from the UCL Institute of Education Research Ethics Committee. Although we do not anticipate substantial ethical issues, particularly as we will not be collecting data directly from children and young people, issues around voluntary participation and informed consent, particularly around participation in the workshops will be explored and addressed in the research. In our first advisory group meeting, we will seek the involvement of group members to co-produce a framework of ethical standards to implement in subsequent stages, including workshop

stages. The ethical framework will help us to make decisions and gain input on issues such as information/consent sheets, how data are recorded/stored and how outputs will be co-produced.

### Organisational affiliation of the project and funding

This research is conducted by researchers at the EPPI-Centre, UCL in conjunction with Co-Production Collective. The research is funded by National Institute of Health Research under the Public Health Research stream (PHR Project: NIHR133736 - Handling complexity in evidence from systematic reviews and meta-analyses of public health intervention).

### Conflict of interests

No conflict of interest has been declared by the authors of this protocol.

### References

- AHMAD, N., BOUTRON, I., DECHARTRES, A., DURIEUX, P. & RAVAUD, P. 2010. Applicability and generalisability of the results of systematic reviews to public health practice and policy: a systematic review. *Trials*, 11, 20.
- BERLIN, J. A. & GOLUB, R. M. 2014. Meta-analysis as evidence: building a better pyramid. *JAMA*, 312, 603-606.
- BOULTON, E., KNEALE, D., STANSFIELD, C., HERON, P., SUTCLIFFE, K., HAYANGA, B., HALL, A., BOWER, P., CASEY, D., JEFFERSON, L., CRAIG, D., GILBODY, S., HANRATTY, B., MCMILLAN, D., THOMAS, J. & TODD, C. 2020. Rapid review of reviews: what remotely delivered interventions can reduce social isolation and loneliness among older adults? London & Manchester, UK: EPPI-Centre & NIHR Older People and Frailty PRU.
- BROWN, T., MOORE, T. H., HOOPER, L., GAO, Y., ZAYEGH, A., IJAZ, S., ELWENSPOEK, M., FOXEN, S. C., MAGEE, L. & O'MALLEY, C. 2019. Interventions for preventing obesity in children. *Cochrane Database of Systematic Reviews*.
- BROWNSON, R. C., FIELDING, J. E. & MAYLAHN, C. M. 2009. Evidence-based public health: a fundamental concept for public health practice. *Annual review of public health*, 30, 175-201.
- BRUNTON, G., CAIRD, J., KNEALE, D., THOMAS, J. & RICHARDSON, M. 2015. Community engagement for health via coalitions, collaborations and partnerships: a systematic review and meta-analysis. London: EPPI-Centre, UCL Institute of Education.
- BRUNTON, G., O'MARA-EVES, A. & THOMAS, J. 2014. The 'active ingredients' for successful community engagement with disadvantaged expectant and new mothers: a qualitative comparative analysis. *Journal of advanced nursing*, 70, 2847-2860.
- BURCHETT, H., BLANCHARD, L., KNEALE, D. & THOMAS, J. 2018. Assessing the applicability of public health intervention evaluations from one setting to another: A methodological study of the usability and usefulness of assessment tools and frameworks *Health Research Policy and Systems*.
- EVANS, B. A., SNOOKS, H., HOWSON, H. & DAVIES, M. 2013. How hard can it be to include research evidence and evaluation in local health policy implementation? Results from a mixed methods study. *Implement Sci*, 8, 17.
- GLASS, G. V. 2000. *Meta-analysis at 25* [Online]. Available: <http://www.gvglass.info/papers/meta25.html> [Accessed 26th November 2014].
- GREENE, M. C., JORDANS, M. J., KOHRT, B. A., VENTVOGEL, P., KIRMAYER, L. J., HASSAN, G., CHIUMENTO, A., VAN OMMEREN, M. & TOL, W. A. 2017. Addressing culture and context in humanitarian response: preparing desk reviews to inform mental health and psychosocial support. *Conflict and Health*, 11, 21.

- HARRIS, K., KNEALE, D., LASSERSON, T. J., MCDONALD, V. M., GRIGG, J. & THOMAS, J. 2019. School-based self-management interventions for asthma in children and adolescents: a mixed methods systematic review. *Cochrane Database of Systematic Reviews*.
- HEDGES, L. V. 2013. Recommendations for practice: justifying claims of generalizability. *Educational Psychology Review*, 25, 331-337.
- HEDGES, L. V., SAUL, J. A., CYR, C., MAGNUS, M., SCOTT-SHELDON, L. A., YOUNG-HYMAN, D. & KHAN, L. K. 2020. Childhood Obesity Evidence Base Project: A rationale for taxonomic versus conventional meta-analysis. *Childhood Obesity*, 16, S2-1-S2-6.
- HO, M., GARNETT, S. P., BAUR, L., BURROWS, T., STEWART, L., NEVE, M. & COLLINS, C. 2012. Effectiveness of lifestyle interventions in child obesity: systematic review with meta-analysis. *Pediatrics*, 130, e1647-e1671.
- IOANNIDIS, J. P., GREENLAND, S., HLATKY, M. A., KHOURY, M. J., MACLEOD, M. R., MOHER, D., SCHULZ, K. F. & TIBSHIRANI, R. 2014. Increasing value and reducing waste in research design, conduct, and analysis. *The Lancet*, 383, 166-175.
- KNEALE, D., O'MARA-EVES, A., REES, R. & THOMAS, J. 2020. School closure in response to epidemic outbreaks: Systems-based logic model of downstream impacts *F1000Research*, 9, 352-352.
- KNEALE, D., ROJAS-GARCÍA, A., RAINE, R. & THOMAS, J. 2017. The use of evidence in English local public health decision-making. *Implementation Science*, 12, 53.
- KNEALE, D., ROJAS-GARCÍA, A. & THOMAS, J. 2018. Exploring the importance of evidence in local health and wellbeing strategies. *Journal of Public Health*, 40, i13-i23.
- KNEALE, D., ROJAS-GARCÍA, A. & THOMAS, J. 2019a. Obstacles and opportunities to using research evidence in local public health decision-making in England. *Health research policy and systems*, 17, 61.
- KNEALE, D., THOMAS, J., O'MARA-EVES, A. & WIGGINS, R. D. 2019b. How can additional secondary data analysis of observational data enhance the generalisability of meta-analytic evidence for local public health decision-making? . *Research synthesis methods*, 10, 44-56.
- MADEN, M. 2016. Consideration of health inequalities in systematic reviews: a mapping review of guidance. *Systematic reviews*, 5, 202.
- MARSH, H. W., BORNHANN, L., MUTZ, R., DANIEL, H.-D. & O'MARA, A. 2009. Gender effects in the peer reviews of grant proposals: A comprehensive meta-analysis comparing traditional and multilevel approaches. *Review of Educational Research*, 79, 1290-1326.
- MELLENDEZ-TORRES, G., SUTCLIFFE, K., BURCHETT, H. E., REES, R., RICHARDSON, M. & THOMAS, J. 2018. Weight management programmes: re-analysis of a systematic review to identify pathways to effectiveness. *Health Expectations*, 21, 574-584.
- MUNTHE-KAAS, H., NØKLEBY, H., LEWIN, S. & GLENTON, C. 2020. The TRANSFER Approach for assessing the transferability of systematic review findings. *BMC Medical Research Methodology*, 20, 11.
- NHS DIGITAL 2020. National Child Measurement Programme, England 2019/20 School Year. In: POPULATION HEALTH TEAM, N. D. (ed.). London: NHS Digital.
- O'MUIRCHARTAIGH, C. & HEDGES, L. V. 2014. Generalizing from unrepresentative experiments: a stratified propensity score approach. *Journal of the Royal Statistical Society: Series C*, 63, 195-210.
- OLIVER, S., DICKSON, K. & BANGPAN, M. 2015. Policy-relevant systematic reviews to strengthen health systems: models and mechanisms to support their production. London: EPPI-Centre.
- OLIVER, S., ROCHE, C., STEWART, R., BANGPAN, M., DICKSON, K., PELLIS, K., CARTWRIGHT, N., GOUGH, D. & HARGREAVES, J. 2018. Stakeholder Engagement for Development Impact

- Evaluation and Evidence Synthesis *CEDIL Inception Paper*. London: Centre of Excellence for Development Impact and Learning (CEDIL), London International Development Centre
- PAUL-EBHOHIMHEN, V. & AVENELL, A. 2008. Systematic review of the use of financial incentives in treatments for obesity and overweight. *Obesity Reviews*, 9, 355-367.
- RAGIN, C. C. 2008. *Redesigning social inquiry: Fuzzy sets and beyond*, Wiley Online Library.
- REHFUESS, E. A., BOOTH, A., BRERETON, L., BURNS, J., GERHARDUS, A., MOZYGEMBA, K., OORTWIJN, W., PFADENHAUER, L. M., TUMMERS, M. & WILT, G. J. 2017. Towards a taxonomy of logic models in systematic reviews and health technology assessments: a priori, staged and iterative approaches. *Research synthesis methods*, 9, 13-24.
- RIHOUX, B. 2009. Qualitative Comparative Analysis (QCA) and related techniques: Recent advances and challenges. *Methoden der vergleichenden Politik-und Sozialwissenschaft*. Springer.
- SUTCLIFFE, K., BURCHETT, H., REES, R., MELENDEZ-TORRES, G., STANSFIELD, C. & THOMAS, J. 2017. What are the critical features of successful Tier 2 lifestyle weight management programmes for children aged 0-11 years? A systematic review to identify the programme characteristics, and combinations of characteristics, that are associated with successful outcomes. London: EPPI-Centre.
- SUTCLIFFE, K., KNEALE, D. & THOMAS, J. 2019. Qualitative Comparative Analysis (QCA) is used in diverse ways in systematic reviews to identify critical intervention components. *Cochrane Colloquium*. Santiago, Chile.
- SUTCLIFFE, K., RICHARDSON, M., REES, R., MELENDEZ-TORRES, G., STANSFIELD, C. & THOMAS, J. 2016. What are the critical features of successful Tier 2 weight management programmes. *A systematic review to identify the programme characteristics, and combinations of characteristics, that are associated with successful weight loss*. London: EPPI-Centre, UCL.
- SUTCLIFFE, K., THOMAS, J., STOKES, G., HINDS, K. & BANGPAN, M. 2015. Intervention Component Analysis (ICA): a pragmatic approach for identifying the critical features of complex interventions. *Systematic reviews*, 4, 1.
- THOMAS, J., O'MARA-EVES, A. & BRUNTON, G. 2014. Using qualitative comparative analysis (QCA) in systematic reviews of complex interventions: a worked example. *Systematic reviews*, 3, 1-14.
- VAN JAARSVELD, C. H. & GULLIFORD, M. C. 2015. Childhood obesity trends from primary care electronic health records in England between 1994 and 2013: population-based cohort study. *Archives of disease in childhood*, archdischild-2014-307151.
- WANG, S., MOSS, J. R. & HILLER, J. E. 2006. Applicability and transferability of interventions in evidence-based public health. *Health promotion international*, 21, 76-83.