Methods

## Systematic Review Methods

### Aim and Research Questions

The aim of this review was to identify plausible hypothesised causes of decreasing rates of teenage pregnancy in the UK from 1999 to 2015. This was done by answering two research questions:

RQ1: What changes in interventions, culture, policy and environment between 1999 and 2015 are hypothesised causes of reduction in rates of teenage pregnancy observed in the UK?

RQ2: What are the most likely explanations, from among those identified in RQ1?

### Search strategy

Four databases (Medline, Embase, Scopus and Cochrane Database of Systematic Reviews) were searched for reviews of adolescent/teenage pregnancy and prevention or reduction in rates. Full search strategies are presented in [Appendix x].

### Screening

#### Inclusion criteria

##### Population

The focus of this review is on adolescents, aged 13-19 years at risk of pregnancy. Populations of the UK and other, similar European and High-Income Countries (HICs) were included, as observations and evaluations of eligible interventions may have been conducted elsewhere. Populations of the 31 countries included in EURO-PERISTAT reporting (Euro-Peristat, 2018); representing European countries were eligible. Further English-speaking HICs were included, as data from these would be comparable to the UK and easily accessed without language barriers. These were Australia, Canada, New Zealand and the United States of America (see Box 1).

Austria

Australia

Belgium

Bulgaria

Canada

Croatia

Cyprus

Czech Rep.

Denmark

Estonia

Finland

France

Germany

Greece

Hungary

Iceland

Ireland

Italy

Latvia

Lithuania

Luxembourg

Malta

Netherlands

New Zealand

Norway

Poland

Portugal

Romania

Slovakia

Slovenia

Spain

Sweden

Switzerland

UK

United States of America

Box 1 A list of the 35 countries to be included in this review

(compiled from Euro-Peristat countries plus other high-income countries to compare with the UK)

##### Intervention/Exposure

Any review addressing an intervention or cultural-macro change which is hypothesised to be causative of changing teenage pregnancy rates over time. This includes reviews evaluating interventions designed to be applied to a population, or vulnerable subset of a population (whether or not details of such application is reported in the review), and observations of other changes in exposure across populations.

##### Control

To be included the review should make a comparison between an exposed and unexposed population, looking at a difference in effects. This may be presented as simultaneous differentiation between two groups, or a before and after comparison of a single group undergoing exposure to a hypothesised cause.

##### Outcome

The primary outcome was rates of pregnancy, measured or estimated from combined counts of births, stillbirths, miscarriages and abortions, or counts of recorded births. Reviews reporting behaviours known to be causally associated with risk of pregnancy – including contraception use and sexual activity – were included, and these measures treated as secondary outcomes.

##### Study type

Only reviews of other published literature were included. This was to address the expectedly large volume of published literature of all types on the topic, by selecting reviews as likely the most influential on policy and anticipating that the most important hypothesised causes would be addressed in such reviews.

#### Exclusion criteria

Several studies were excluded which only looked at pregnancy and maternity after conception.

Additionally, studies which evaluated associations of hypothesised determinants of pregnancy risk across population, with no analysis of changes of exposure over time, were not included.

### Data extraction

From each review, I extracted the included population, the interventions or exposures assessed, outcomes related to pregnancy and a summary of the findings of the effects of the interventions or exposures addressed. To assess quality of evidence, I extracted summaries of the search strategy used, whether the review was described as systematic or not

### Synthesis

In synthesising the evidence I used SWiM guidelines, developed to structure reporting of quantitative data without meta-analysis (Campbell et al., 2020).

Reviews which addressed a single defined intervention model were first grouped into broad categories determined by context of delivery and characteristics of participants. These were then sub-categorised into distinguishable intervention models. Papers which evaluated several interventions were then analysed and each model assigned to existing or newly created categories.

Cultural, environmental and policy changes which were reviewed as potentially causative of changing rates, or proposed by reviews as time-varying confounders, were collated as one broad category. These were then grouped by distinguishable hypothesised pathways.

For each intervention, data was extracted on the description of the intervention, the context of delivery, the countries and periods studied (dated by publication of papers used, as data-collection dates were not always available), targeted participants, relevant outcomes tested, and evidence presented for effects on primary and secondary outcomes.

#### Quality assessment

I used the questions developed by Aromataris et al, (2015; Table 1) in assessing the quality of each review.

### Analysis

#### Assessing feasibility using adapted questions

To assess the plausibility of each hypothesised cause, a series of questions were developed to determine the strength of evidence presented in the included reviews. The criteria developed by Bradford Hill were used as a starting point (Hill, 1965). Adaptation of these was necessary to account for the anticipated heterogeneity of evidence presented in the reviews, and the likely complexity of the proposed exposures to be tested. An updated model for this purpose is presented by Howick et al. (2009). This is intended to account for various types of evidence available in the ‘evidence hierarchy’, allowing some confidence of causation to be drawn from non-RCT evidence (Table 1).

|  |  |  |
| --- | --- | --- |
| Type of evidence | Revised, structured guidelines | Hill's original guidelines |
| *Direct* | Size of effect not attributable to plausible confounding | Experiment |
| Appropriate temporal and/or spatial proximity (cause precedes effect and effect occurs after a plausible interval; cause occurs at the same site as the intervention) | Strength |
| Dose‐responsiveness and reversibility | Temporality |
| *Mechanistic* | Evidence for a mechanism of action (biological, chemical, mechanical) | Biological gradient  Biological plausibility |
| Coherence | Coherence |
| *Parallel* | Replicability | Consistency |
|  | Similarity | Analogy |

Table 1 Bradford Hill's original guidelines and proposed revisions (Howick et al., 2009; Table 1).

'Coherence' is moved from the 'Parallel' subheading to 'Mechanistic', in accordance with the paper's summary of the guidelines

Use of the criteria for assessing non-randomised evidence for public health interventions is also discussed by Schünemann et al. (2011) and Hultcrantz et al. (2017).

In this systematic review, it was expected that multiple types of study would be present, fitting the purposes of these revised guidelines. Additionally, evidence would often be missing as to the proportions of populations exposed to each hypothesised causal pathway. The presence of such evidence would strengthen causal inference, however its absence may not be sufficient cause to eliminate the causal pathway from the final model. This would present challenges to the goal of assessing the plausibility of causation of nationally observed decreases in rates of teenage pregnancy.

To account for this, a set of questions was devised to be applied to each causal pathway. These correspond to the revised, structured guidelines of Table 1 (the category of ‘Similarity’ was omitted, as reviews being assessed tend to pre-empt this question):

1. Is data presented which shows an effect while controlling for plausible confounding? [Direct]
2. Is this hypothesis:
   1. Tested across the population of a whole nation (as included in this analysis – see Box 1); *OR*
   2. Tested amongst a large proportion of the population, or a high-risk group (e.g. targeting prevention of repeat pregnancies; low socio-economic status of participant); *OR*
   3. Noted by authors to have been applied to a population other than the observed groups, but within this review’s observation period (1990-)? [Direct]
3. Is a logic model, or narrative synthesis of mechanisms presented? [Mechanistic]
4. Is the model coherent with what is currently known? [Mechanistic - Coherence]
5. Are results seen consistently in different national contexts? [Replicability]

For interventions which are presented as effective in an experimental population, but for which no evidence is presented for national implementation which may have occurred after the reviews were conducted, further evidence was sought to determine whether this model had been widely applied. Policies and guidelines for the appropriate UK providers was consulted when needed to answer question 2. For each intervention the five questions were presented narratively, then summarised in tabular format.

#### Presenting hypothesised causes by plausibility

By comparing across hypothesised pathways, the most plausible causes were identified as those presenting evidence to answer all five questions. Further causes which presented persuasive evidence were also considered to be plausible. Hypotheses which offered no further evidence were excluded. For each included pathway, a simplified logic model was extracted from the relevant papers. This was used to produce a finalised logic model, based on the initially proposed model in [Section 2.x.x].

## References

Aromataris, E., Fernandez, R., Godfrey, C. M., Holly, C., Khalil, H., & Tungpunkom, P. (2015). Summarizing systematic reviews. *International Journal of Evidence-Based Healthcare*, *13*(3), 132–140. https://doi.org/10.1097/XEB.0000000000000055

Campbell, M., McKenzie, J. E., Sowden, A., Katikireddi, S. V., Brennan, S. E., Ellis, S., … Thomson, H. (2020). Synthesis without meta-analysis (SWiM) in systematic reviews: Reporting guideline. *The BMJ*, *368*. https://doi.org/10.1136/bmj.l6890

Euro-Peristat. (2018). Country Teams - Euro-Peristat. Retrieved April 3, 2019, from https://www.europeristat.com/index.php/our-network/country-teams.html

Hill, A. B. (1965). The Environment and Disease: Association or Causation. *Proceedings of the Royal Society of Medicine*, *58*(5), 295–300. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/14283879

Howick, J., Glasziou, P., & Aronson, J. K. (2009). The evolution of evidence hierarchies: what can Bradford Hill’s “guidelines for causation” contribute? *Journal of the Royal Society of Medicine*, *102*(5), 186–194. https://doi.org/10.1258/jrsm.2009.090020

Hultcrantz, M., Rind, D., Akl, E. A., Treweek, S., Mustafa, R. A., Iorio, A., … Guyatt, G. (2017). The GRADE Working Group clarifies the construct of certainty of evidence. *Journal of Clinical Epidemiology*, *87*, 4–13. https://doi.org/10.1016/j.jclinepi.2017.05.006

Schunemann, H., Hill, S., Guyatt, G., Akl, E. A., & Ahmed, F. (2011). The GRADE approach and Bradford Hill’s criteria for causation. *Journal of Epidemiology and Community Health*, *65*(5), 392–395. https://doi.org/10.1136/jech.2010.119933